Master Thesis Business Informatics

Business Intelligence as a Service: A Vendor's Approach

Author:

Tim de Boer BSc
Student Nr : 0237884
Thesis Nr : INF-0237884
Email : timdeboer@xs4all.nl

Department of Information and Computing Sciences
Utrecht University
Princetonplein 5, 3584 CC Utrecht, NL

Supervisors:

Utrecht University : Dr. Marco Spruit & Dr. Slinger Jansen

Avanade : Drs. Joris Valkonet & Drs. Gijs Ramakers

Keywords: Business Intelligence, Cloud Computing, Capability, Model, Maturity, development, roadmap, BIaaS, BI, SaaS

ABSTRACT

Demand for business intelligence (BI) applications continues to grow even at a time when demand for most information technology (IT) products is low, showing the importance of BI products for a modern organization. However, globalization changes the way organizations use BI, where geographic location and time independency is becoming more and more important. Gartner's hype-cycle on BI depicts the technology of BI as a Service as being almost on top of the hype cycle, indicating there are high expectations of this new technology.

This thesis research advances on existing literature on business intelligence and cloud computing from a development perspective by introducing the concept of business intelligence as a service (BIaaS). The most important deliverable in the creation of the BIaaS concept is the BIaaS capability maturity model (CMM) that is introduced. The BIaaS CMM explains the conceptual model of BIaaS by the creation of the first BIaaS capability model containing key capabilities of BIaaS. The capability model is further enhanced with maturity levels (depicting the importance of each BIaaS capability), a maturity matrix (suggesting a roadmap for BIaaS solution development) and a BIaaS assessment model (introducing a tool for finding problem areas in existing BIaaS solutions). The developed BIaaS CMM ought to support (starting) BIaaS vendors to develop BIaaS solutions by providing an assessment tool BIaaS solutions. The assessment outcome provide the current maturity of the BIaaS solution and also include problem areas for solution improvement.

The introduction of the capability maturity positioning method (CAMP) for the development of a maturity matrix, which result in the BIaaS maturity model, is significantly different from conventional maturity modeling. To calculate the weight of each capability from the BIaaS capability model, a thorough product review of existing business intelligence and cloud computing products is performed. Analysis of the results and normalizing the outcome of that analysis together with the introduction of a calculation mapping, is input for the creation of the maturity matrix. The maturity matrix is the essential foundation for the developed business intelligence as a Service capability maturity model, which is the biggest deliverable of this thesis research.

ACKNOWLEDGEMENTS

Utrecht, July 2012

This master thesis on business intelligence as a service is the final result of my two year master (MSc) education on Business Informatics at Utrecht University.

First of all I like to thank the company and the people at Avanade to give me the opportunity to do my research making exclusive use of all their resources. In particular I like to thank two people at Avanade, Joris Valkonet and Gijs Ramakers. Joris Valkonet supported me with insights, new ideas and knowledge about Business Intelligence and the contents of my research. Gijs Ramakers accompanied me in doing technical and methodical sound research and was of great support when overcoming difficulties during research and always being available when needing support. Secondly, I am grateful to my supervisors at Utrecht University, Marco Spruit and Slinger Jansen for their guidance and constructive critics during my thesis project. Last but not least, I would like to thank my parents for providing me the opportunity to study and their support during this period.

Tim de Boer

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

BI	*Business Intelligence* is the combination of data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to planners and decision makers in organizations.
BIaaS	*Business Intelligence as a Service* is the merger of Business Intelligence (BI) with Cloud Computing to introduce BI offered as a service in the Cloud, or Business Intelligence as a Service.
BPM	*Business Performance Management* sometimes referred to as Corporate Performance Management (CPM) or Enterprise Performance Management (EPM) is a set of processes that help organizations optimize business performance by encouraging process effectiveness as well as efficient use of financial, human, and material resources.
CEP	*Complex Event Processing* technology which is used in event-driven BI applications with goal of identifying meaningful patterns, relationships and data abstractions from among seemingly unrelated events and trigger immediate response actions.
CMM	*Capability Maturity Model* are models to evaluate and compare people, process or object capability as basis for improvement and in order to derive an informal approach for increasing the capability if a specific focus area.
ETL	*Extraction Transformation Load* is a process to extract data from a data source, transform the data into usable data batches and load the usable data into a target data source (i.e. the data warehouse). With other words, the objective is to transform operational data from a source database to subject data in the data warehouse.
GIS	*Geographic Information Systems* is a software package that links databases and electronic geographical maps to analyze spatial phenomena.
IaaS	*Infrastructure as a Service* is the concept of offering a complete infrastructure to customers which is available on-demand via the internet. The infrastructure is scalable to the customer's demand and is a virtual environment.
KPI	*Key Performance Indicator* is a performance measurement used by organizations to evaluate its success in particular business activities.
OLAP	*Online Analytical Processing* is a tool for fast and user friendly analysis of the multidimensional data in the data warehouse. The key operations available in OLAP include rollup and drill-down along one or more dimension hierarchies, slice-and-dice, and pivot.
PaaS	*Platform as a Service* is the concept of offering a platform to customers which is available on-demand via the internet. Customers can use the platform to develop their own SaaS solution using the platform to deploy their solutions on. The platform is scalable on-demand and is a virtual environment.

PDD	*Process Deliverable Diagram* is a diagram to visualize processes by showing the activities and deliverables of that activity, based on a meta-modeling method which integrates two diagrams. The diagram has two sides, a left- and right-hand side. The activities are on the left-hand side of the diagram and are based on a UML activity diagram. The deliverables are on the right-hand side and is based on a UML class diagram.
SaaS	*Software as a Service* is the concept of offering complete software solutions to customers which are available on-demand via the internet. The solution is particularly scalable to the customer's use, is time and location independent, has a competitive pricing model and is a virtual environment.
SLA	*Service Level Agreement* is a contract between the provider and the user that specifies the level of service that is expected during its term. SLAs are used by vendors and customers, as well as internally by IT shops and their end users. They can specify bandwidth availability, response times for routine and ad hoc queries and response time for problem resolution (network down, machine failure, etc.). SLAs can be very general or extremely detailed, including the steps taken in the event of a failure. For example, if the problem persists after 30 minutes, a supervisor is notified; after one hour, the account representative is contacted.

1 RESEARCH INTRODUCTION

1.1 Introduction

Analysts at Gartner's research institute (Gartner, 2011) predict a shift in the way organizations use Business Intelligence (BI) nowadays (Chandler, 2011). This prediction is driven by "Software as a Service" (SaaS) solutions IT companies offer for their customers. Therefore Gartner predicts the shift towards BI as a Service (BIaaS) solutions. Although a lot of research is done on BI and also SaaS is becoming a more and more popular topic in research, the merge of both topics into BIaaS is a new research area. This thesis research elaborates on BIaaS from both scientific and business perspectives to help readers better understand BIaaS and support BIaaS vendors form their strategy to develop BIaaS solutions. This thesis proposes a capability model to make a conceptual model of BIaaS, which explains what BIaaS entails. The capability model is used to develop a BIaaS capability maturity model. Maturity models are normally used to assess processes, like the capability maturity model introduced by Paulk et al. (1993) to improve the software process, however in this thesis research the capability maturity model is specifically developed to assess BIaaS solutions. The BIaaS capability maturity model will assist vendors in developing mature BIaaS solutions and provide a guideline in their software process and software capability evaluation.

From a scientific point of view, the produced capability model and maturity model for BIaaS solutions certainly add value to existing BI, SaaS and BIaaS literature, as this research present the first BIaaS capability and maturity model available. The capability maturity model also contribute from a practical business perspective, where vendors can use the model to assess their own developed BIaaS solutions and get detailed information to improve their existing products to make them more BIaaS mature. The BIaaS capability maturity model is aimed to be used by BIaaS solution vendor's as a foundation for their BIaaS product development strategy, whereas the BIaaS capability maturity model can be used as an assessment tool for BIaaS solutions and provide improvement steps for further development.

1.2 Company Profile

This Master thesis research is performed as a collaboration between Utrecht University and Avanade. Avanade is a joint venture with Accenture (one of the world biggest business consultancy companies) and Microsoft (one of the world biggest supplier of technological business solutions). Avanade is the bridge between business consultancy (Accenture) and business technology (Microsoft), by developing and implementing technological business solutions and offering managed services with the use of Microsoft technologies.

Avanade is incepted in 2000 in the USA. Since then they are growing rapidly with an average yearly growth of twenty percent. Today Avanade is grown into an international company with more than sixty locations in over twenty countries worldwide (on every continent). Avanade has a revenue of $953 million in fiscal year 2010 and staff more than 7000 employees that serve over 700 customers.

In the Dutch market, Avanade employs over 300 professionals, whereof one-tenth is working in the Business Intelligence service line (i.e. group of professionals serving customers with BI related issues). During this research, I was part of Avanade's BI service line where I collaborated with the professionals in that group. The collaboration provided extensive knowledge about BI that was input for this research. Moreover, the company's world-wide network provided connections with professionals all over the world and explicit knowledge from the companies knowledge base (including the knowledge base of Accenture).

1.3 Problem Statement

In the twenty-first century, organizations are dealing with a fast changing environment. The changing environment is mainly caused by globalization, where geography and time boundaries no longer limit organizational processes. To react on this globalization, technology vendors develop more and more advanced technologies to accommodate organizations and give them competitive opportunities in this changing environment (Kakihara & Sorenson, 2002). One essential component in organizational success and increasing their competitive position is Business Intelligence (BI) (Wixom, Watson, 2010). But also with BI, globalization is changing the way organizations use BI. Researchers at Gartner (Gartner, 2011), but also in scientific literature like Chandler (2011), predict a shift in the way organizations use BI and it's driven by the upcoming Software as a Service (SaaS) solutions. This shift results in BIaaS solutions, Business Intelligence offered as a service on the internet (e.g. BI that is geographically and time independent). Gartner's Hype Cycle also makes specifically clear, that the BIaaS topic is a trending phenomenon and is still not reaching its peak (Figure 1). A "Hype Cycle" is a way to represent the emergence, adoption, maturity and impact on applications of specific technologies.

More and more BI vendors are exploring the BIaaS solution hype by improving their BI products with SaaS abilities and thereby developing BIaaS-like solutions. BI (in some form) exists for decades and extensive knowledge is available from different studies, and also SaaS as we know it now, exist for some years and researchers have followed this trend by doing research on SaaS and cloud computing. However, in contrast to BI and SaaS, combining these two together and offering BI as a Service (BIaaS) on the Internet is a very young field of research, and therefore there is not a specific BIaaS model available that elaborates on what BIaaS solutions should entail or a best practice for BIaaS solution development. Taking this lack of clarity about BIaaS and the missing guideline for BIaaS solution development into account, this master thesis research focusses on defining BIaaS and its contents and the creation of a roadmap for BIaaS solution development.

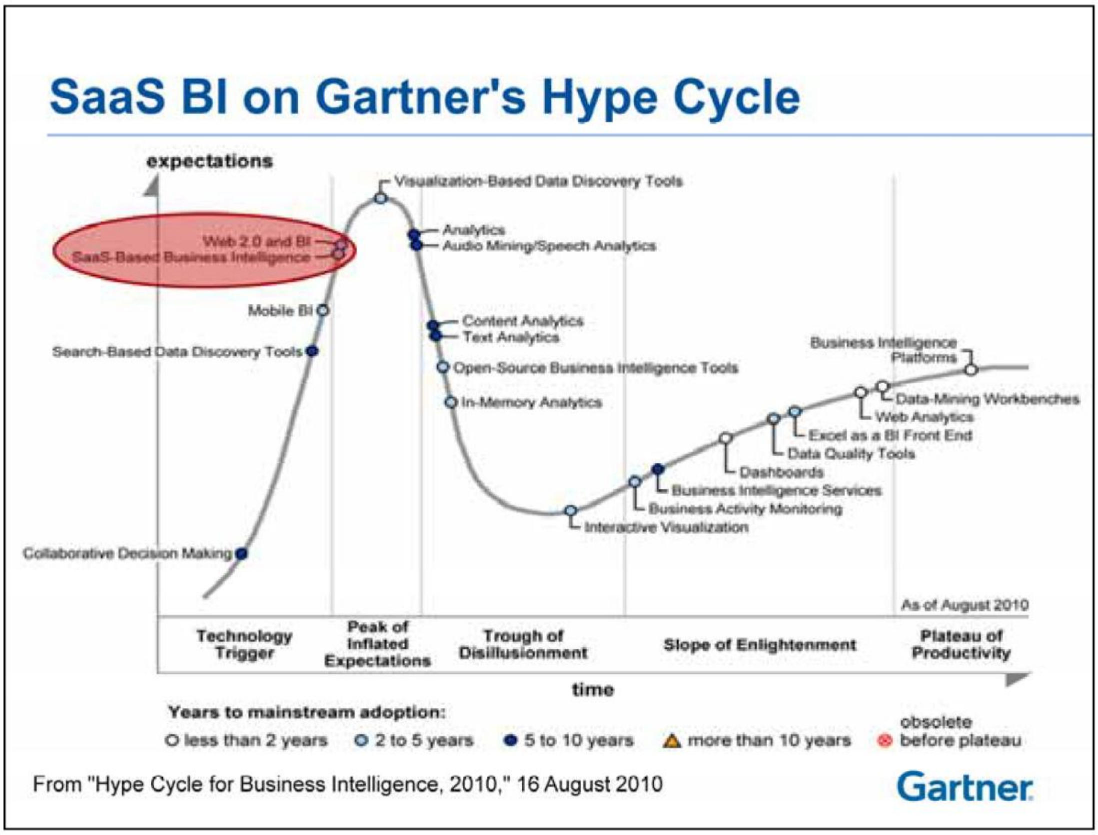


Figure 1: Gartner's Hype Cycle for Business Intelligence (2010)

1.4 Background

Understanding Business Intelligence and Cloud Computing is the basis for combining it into BIaaS. This paragraph provides the background of BIaaS and explains why BI and Cloud Computing are brought together to form business intelligence as a service.

1.4.1 Business Intelligence

The term Business Intelligence (BI) can best be explained by the definition given by Negash (2004): "BI systems combine data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to planners and decision makers". Implicit in this definition is the idea (perhaps the ideal) that business intelligence systems provide actionable information delivered at the right time, at the right location, and in the right form to assist decision makers. The objective is to improve the timeliness and quality of inputs to the decision process, hence facilitating managerial work (Negash, 2004).

While the term BI is relatively new, computer-based business intelligence systems appeared, in one guise or other, close to forty years ago (Power, 2004). In the 1980's, finance and telecommunication companies pioneered BI to support financial and market analysis of the large volumes of data that they had begun to accumulate electronically. The need for BI capabilities grew in the 80's and 90's in other industries as companies began capturing data electronically across the full range of their business activities. This need was further compounded by the growing interest in real time data access which required effective tools to mine and analyze, dramatically increased data volumes (Shobrys, 2003). To support this growing need, large software and services providers like IBM and Oracle launched major initiatives to bring data warehousing capabilities to the marketplace. The data warehouse grew to be the central component of BI systems and integrates data from transactional IS for analytical tasks (Inmon, Strauss, Neushloss, 2008; Kimball, Ross, Thornthwaite, Mundy, Becker, 2008). Enterprise Resource Planning (ERP) systems have also been used to capture data and enforce consistency, but although ERP systems, like data warehouses, use structured data from databases, they tend to be too inflexible to support ad hoc exploration of data (Negash, 2004; Blumberg & Atre, 2003a; Shobrys, 2003).

The emergence of the data warehouse as a repository, the advances in data cleansing that lead to a single truth, the greater capabilities of hardware and software, and the boom of Internet technologies that provided the prevalent user interface all combine to create a richer business intelligence environment than was available previously (Negash, 2004). Also, with each product update, BI capabilities increased as enterprises grew ever-more sophisticated in their computational and analytical needs (Negash, 2004). The richer environments usually start with flexible query and reporting capabilities that are combined with some mix of online analytical processing (OLAP), statistical analysis, forecasting and data mining techniques (Shobrys, 2003). BI was initially coined as a collective term for data analysis tools (Anandarajan, Anandarajan, Srinivasan, 2004), meanwhile, the understanding broadened towards BI as an encompassment of all components of an integrated decision support infrastructure (Baars & Kemper, 2008).

The past decennia the demand for Business Intelligence applications continues to grow even at a time when demand for most information technology (IT) products is soft (Soejarto, 2003; Whiting, 2003). One reason for this continuous growth is that BI is currently the top-most priority of many chief information officers (Watson & Wixom, 2007). BI has become an essential component of the information supply infrastructure and a contributor (and prerequisite) to the overall organizational success (Wixom, Watson, 2010). Therefore more and more enterprise solutions and platforms for Business Intelligence have been developed. These vendors and platforms include the commercial solutions such as IBM DB2 with Business Intelligence Tools (Gonzales, 2003), Microsoft SQL Server, NCR Teradata Warehouse, Hyperion/Brio, SAS, iData Analyzer, Oracle, Cognos, Business Objects, OLAP4All, and the open source project Pentaho (Thomsen & Pedersen, 2008), and MSMiner (Shi et al., 2004).

While relational database applications are query-driven, event-driven BI applications have become increasingly important. Event-driven applications are characterized by high event data rates, continuous queries, and millisecond latency requirements. While company data is typically stored in large relational databases, the growing need for BI, dramatically increasing data volumes and event-driven BI applications, making the relational databases for processing impractical. These requirements are shared by vertical markets such as: financial services, health care, IT monitoring, manufacturing, oil and gas, transportation, utilities, and web analytics. Event-driven applications use complex event processing (CEP) technology with the goal of identifying meaningful patterns, relationships and data abstractions from among seemingly unrelated events and trigger immediate response actions. A solution for this event-driven BI in combination with relational database is the current trend, in-memory data for fast data exploration. This is a relatively new technique that is also used in for instance the Twitter environment to insert and query large amount of data in a few milliseconds (de Boer, Lossek, Janssen, Neppelenbroek, 2011).

1.4.2 Cloud Computing

With the advancement of our modern human society, basic and essential services are delivered almost to everyone in a completely transparent manner. Utility services such as water, gas, and electricity have become fundamental for carrying out our daily life and are exploited on a pay per use basis. The existing infrastructures allow delivering such services almost anywhere and anytime so that we can simply switch on the light, open the tap, and use the stove. The usage of these utilities is then charged, according to different policies, to the end user. Recently, the same idea of utility has been applied to computing and a consistent shift towards this approach has been done with the spread of Cloud Computing (Vecchiola, Chu, Buyya, 2009). Therefore Cloud Computing can be defined as a recent technology trend whose aim is to deliver on demand IT resources on a pay per use basis (Vecchiola, Chu, Buyya, 2009).

Cloud Computing aims to be global and to provide computing services to the masses, ranging from the end user that hosts its personal documents on the Internet, to enterprises outsourcing their entire IT infrastructure to external data centers (Vecchiola, Chu, Buyya, 2009). Vecchiola, Chu and Buyya (2009) define three pillars on top of which Cloud computing solutions are delivered to the end users. These pillars are: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). The three defined pillars offer significant benefit to IT companies by freeing them from the low level tasks of setting up basic hardware (i.e. IaaS) and software infrastructures (i.e. PaaS and SaaS) and thus enabling them to focus on innovation and creating business value for their services (Buyya, Pandey, Vecchiola, 2009). Never before an approach to make IT a real utility has been so global and complete: not only computing and storage resources are delivered on demand but the entire stack of computing can be leveraged on the Cloud (Vecchiola, Chu, Buyya, 2009).

Software is a general term used to describe the computer programs, procedures, rules and the associated documentation, in relation to the operation of a computer system, which are stored in a read/write memory unit as part of the digital system (Langholz, et al., 1998; Wordreference.com, 2011). “The last few years there is a trend towards Software as a service (SaaS), which has been predicted by Rust and Kennan (2003) as early as 2003 when they still called it 'e-service'” (Abdat, Spruit, Bos, 2011). In SaaS, customers have the freedom to use the software as they require. Hence, SaaS is often referred to as 'on-demand' software. SaaS applications are being installed in data centers and no longer delivered as a product (physical object), but as a service (Figure 2). This fact has become the reason why the term 'product' does not fit any longer in the world of SaaS. The term 'solution' is being used instead (Abdat, Spruit, Bos, 2011) .

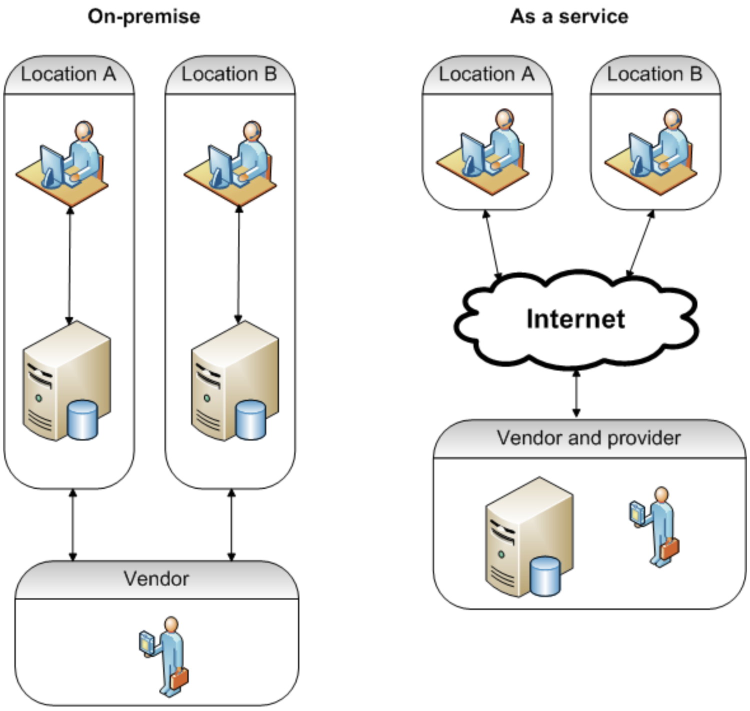


Figure 2: On-premise vs. as a service

1.5 Related Literature

Thomson and van der Walt (2010) outlined the essence for companies to invest in BI in the cloud. Their research conclusion stated that BIaaS solutions (i.e. BI solutions offered as services on the internet) will allow companies to reduce cost of having a BI solution and also having access to the latest software which will give the business an edge on their competition. This opposed edge on the competition by having access to the latest software is long supported by earlier research done for instance by Porter and Millar (1980), and somewhat more recently by Clemmons and Row (1991).

De Marco et al (2010) support the conclusion of Thomson and van der Walt (2010) on BIaaS solutions that reduce company costs, but they also seek to establish a conceptualization of the enabling factors in Business Intelligence as a Service solution adoption. Their main research objective was to propose a model containing enabling factors for the adoption of BIaaS solutions. They seek to expand on the Benlian, Hess and Buxmann (2009) model which is based on a theoretical framework including axioms from Transaction Cost Theory, Resource Based View and Theory of Planned Behavior.

Outlined the benefits for companies to have BIaaS solutions, it is clear why analysts at Gartner Inc. foresee a growth of investments companies will do in BIaaS solutions and therefore the importance for BI solution vendors to offer BIaaS solutions for their customers.

The subject of this thesis project is BI offered as a service in the cloud (BIaaS). A clear definition of cloud computing is therefore a necessity. For years many definitions of cloud computing have been made by a great variety of researchers, but they all seem to focus on certain aspects of technology (Bragg, 2009; Buyya, Yeo, Venugopal, 2008; Geelan, 2008; Gruman and Knorr, 2008; Hwang, 2008; McFedries, 2008; Milojicic, 2008). Recently Ambrust et al (2009) tried to make a more comprehensive definition of cloud computing. They define Cloud Computing as both the applications delivered as services over the Internet and the hardware and systems software in datacenters that provide those services. The services themselves have long been referred to as Software as a Service (SaaS), therefore the term SaaS will be used for services. The datacenter hardware and software is called a Cloud. BI in the cloud (BIaaS) is therefore Business Intelligence software solutions offered as a service in the Cloud.

The research in this thesis studies BI solutions in the cloud and therefore the features concerning BIaaS has to be defined. Vaquero et al (2009) already underlined a downside in their research; a *complete* overview of all cloud computing features is not available in literature yet. Nevertheless they have made a good start in defining those cloud computing features that are available in literature. BI on the other side, exist for some time and therefore more information about BI features is available. Langseth and Vivatrat (2003) did research on pro-active BI and summed up all the features of pro-active BI. Negash (2004) used the work of Langseth and Vivatrat (2003) to identify the features for BI in general. Although a lot of the BI features are identified in earlier research, more research has to be done to identify all BI features used nowadays.

The research and outcomes of this thesis complements the above, but it differs from them in that it finally will approach BIaaS from a vendor perspective. At the end of the research a usable model is delivered for BIaaS solution vendors who want to develop BIaaS solutions.

1.6 Research goal

The objective of this thesis research is to enrich the research field of BIaaS and furthermore assist BIaaS vendors in their start-up phase of developing BIaaS solutions. This thesis document will give an in-depth scientifically based overview of what BIaaS entails, but also a guideline for BIaaS vendors how they can improve their existing BIaaS (or BI) solutions to make them fully BIaaS mature. Vendors can use the introduced Capability Maturity Model and description as a guideline in their solution improvement process.

1.7 Research Questions

This research is driven by two main research questions. Due to the relatively new research field of BIaaS, the first research question is to define the research area of BIaaS and how this differ from conventional BI.

"*What are business intelligence as a service capabilities and how do they differ from conventional business intelligence capabilities?*"

Developing a capability model for BIaaS will partly answer this research question and give a clear conceptual model of the BIaaS research area.

The major goal of this thesis research is to assist BIaaS researchers and vendors in researching and developing BIaaS. Therefore a second research question is formulated to find a roadmap for the development process of BIaaS solutions.

"*How can business intelligence as a service capabilities be used to create a product portfolio roadmap for business intelligence as a service solution vendors?*"

By developing a Capability Maturity Model for BIaaS solutions accompanied with detailed descriptions and improvement steps, it aims to be of assistance for BIaaS vendors in their product management process.

1.8 Research Approach

The first research question is defined as "*What are business intelligence as a service capabilities and how do they differ from conventional business intelligence capabilities?*". The approach for answering this first research question is to do a literature study. The literature study is conducted using the structured literature method proposed by Webster and Watson (2002). This literature review form the basis of this research identifying the fundamentals of BI and Cloud computing and summarizing their main features. Although the term "*as a Service*", which is included in BIaaS, suggest only the SaaS aspects of cloud computing need to be taken into account, however, for this research all aspects of cloud computing are reviewed. SaaS features only cover a part of the most important features of cloud computing, therefore, to ensure a fully cloud enabled BIaaS solution, all aspects need to be reviewed.

The literature review is followed-up by a research conducting a BI and cloud computing product review. This research uses the capabilities from the literature review to identify BI and cloud computing capabilities currently available in vendors' solutions. A quantitative analysis of this research will spread light on the currently top most used capabilities (key capabilities), and therefore identifies the most common capabilities from practice. The aim for doing a literature review together with the product review is to identify the BI and Cloud Computing features from both theory and practice and give weight to the different capabilities.

After this first two steps in this research, the identified features from BI and cloud computing will be integrated into a BIaaS capability model, which is the real first deliverable. The capability model is formed by proposing the key capabilities to multiple experts on BI and cloud computing using the qualitative interviewing method of Kvale (1996). Feedback from the experts are used to modify the capability model into a final model. By delivering this capability model and analyze the differences between BIaaS capabilities and those of conventional BI capabilities, the first research question is answered.

The second research question is "*How can business intelligence as a service capabilities be used to create a product portfolio roadmap for business intelligence as a service solution vendors?*". The goal of the second research question is to find a solution to use the BIaaS capability model delivered from the first research question to create a product portfolio roadmap for BIaaS solution vendors in their process of developing BIaaS solutions. The second research question is answered by developing a maturity model for BIaaS solutions which can be used by vendors to assess BIaaS solutions (i.e. vendors own solutions). The assessment is used to develop their own roadmap using the descriptions for each maturity level.

To visualize the research process, a process delivery diagram (PDD) introduced by Weerd and Brinkkemper (2008) is used (Figure 3). The main activities of this research are shown on the left hand side of the diagram, which are connected via the dotted lines to the deliverables shown in the concepts on the right hand side of the diagram. The activities and corresponding deliverables are explained in more detail, in Table 1 and Table 2 respectively.

Table 1: Activities and sub-activities from PDD

Activity	Sub-Activity	Description
Perform literature study	Study related literature	Read scientific literature about BI and Cloud Computing and deliver the RELATED LITERATURE sub-chapter in the thesis document
	Identify BI capabilities	Identify BI CAPABILITY's and store them in the BI CAPABILITY DATABASE together with the reference to the scientific articles
	Identify Cloud capabilities	Identify CLOUD CAPABILITY's and store them in the CLOUD CAPABILITY DATABASE together with the reference to the scientific articles
Develop capability model	Extract major capabilities	Define thresholds for the number of scientific references per capability and extract MAJOR BI and CLOUD CAPABILITY's from the BI and CLOUD CAPABILITY DATABASE
	Identify focus areas	Identify strongly coherent groups of MAJOR BI and CLOUD CAPABILITY's
	Integrate focus areas	Integrate the BI and CLOUD FOCUS AREAS into one BIAAS CAPABILITY MODEL containing all MAJOR BI and CLOUD CAPABILITY's grouped into focus areas
Develop maturity model	Conduct product review	Review BI and Cloud Computing products and identify which MAJOR BI and CLOUD CAPABILITY's are implemented in the products
	Analyze product review	Analyze the outcome of the BI and CLOUD PRODUCT REVIEW's using quantitative analysis
	Develop maturity matrix	Develop the BIAAS MATURITY MATRIX using the analysis from the BI and CLOUD PRODUCT REVIEW as input to identify the levels of maturity per focus area
	Develop maturity model	Integrate the BIAAS MATURITY MATRIX into the BIAAS MATURITY MODEL making it useful for assessments of BIaaS solutions

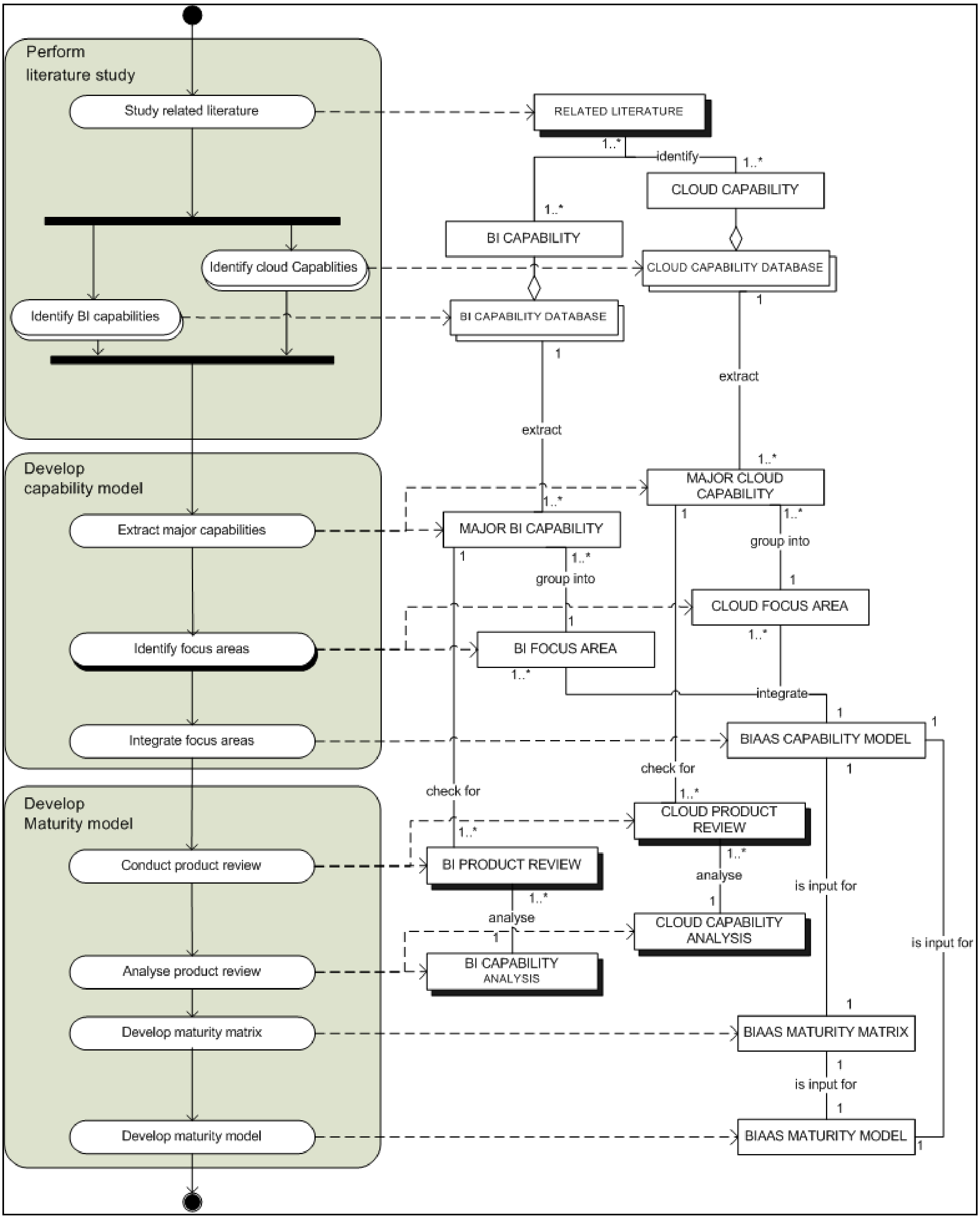


Figure 3: Process Delivery Diagram (PDD) with main research activities and deliverables

Table 2: Concepts (deliverables) from PDD

Concept	Description
RELATED LITERATURE	Related literature is subtracted from the performed literature study following the structured literature review method proposed by Webster and Watson (2002) (Figure 4). This related literature is summarized in the related literature chapter of the thesis document. References to the obtained literature are stored in a literature database.
BI CAPABILITY	A BI CAPABILITY is the ability to perform a set of co-ordinated tasks, utilizing technological resources, for the purposes of achieving a particular business intelligence result.
CLOUD CAPABILITY	A CLOUD CAPABILITY is the ability to perform a set of co-ordinated tasks, utilizing technological resources, for the purposes of achieving a particular Cloud Computing result.
BI CAPABILITY DATABASE	The BI CAPABILITY DATABASE is a database storing BI CAPABILITY's, holding references to scientific literature about the capability.
CLOUD CAPABILITY DATABASE	The CLOUD CAPABILITY DATABASE is a database storing CLOUD CAPABILITY's, holding references to scientific literature about the capability.
MAJOR BI CAPABILITY	A MAJOR BI CAPABILITY is a capability that is a component in BI literature from RELATED LITERATURE in more scientific articles than the set threshold.
MAJOR CLOUD CAPABILITY	A MAJOR CLOUD CAPABILITY is a capability that is a component in Cloud Computing literature from RELATED LITERATURE in more scientific articles than the set threshold.
BI FOCUS AREA	A BI FOCUS AREA is a group of MAJOR BI CAPABILITY's that are strongly coherent.
CLOUD FOCUS AREA	A CLOUD FOCUS AREA is a group of MAJOR CLOUD CAPABILITY's that are strongly coherent.
BIAAS CAPABILITY MODEL	The BIAAS CAPABILITY MODEL is an integrated overview of all MAJOR BI and CLOUD CAPABILITY's, grouped into FOCUS AREAS.
BI PRODUCT REVIEW	A BI PRODUCT REVIEW is an overview of all major BI capabilities available in a BI product investigated in the Product review phase of the thesis research.
CLOUD PRODUCT REVIEW	A CLOUD PRODUCT REVIEW is an overview of all major Cloud Computing capabilities available in a Cloud Computing product investigated in the product review phase of the thesis research.

Table 2 (part 2): Concepts (deliverables) from PDD

Concept	Description
BI CAPABILITY ANALYSIS	The BI CAPABILITY ANALYSIS consist of a quantitative analysis of all BI products that were reviewed in the product review phase.
CLOUD CAPABILITY ANALYSIS	The CLOUD CAPABILITY ANALYSIS consist of a quantitative analysis of all Cloud Computing products that were reviewed in the product review phase.
BIAAS MATURITY MATRIX	The BIAAS MATURITY MATRIX consist of ordered levels of maturity per focus area from the BIAAS CAPABILITY MATRIX, indicating a best practice order in which capabilities to be implemented.
BIAAS MATURITY MODEL	The BIAAS MATURITY MODEL uses the BIAAS MATURITY MATRIX and the BIAAS CAPABILITY MODEL to assess BIaaS solutions. The BIAAS MATURITY MODEL delivers a roadmap for BIaaS solution developers.

1.9 Research methods

Different research methods are used in this thesis research to come to all used conclusions. This sub-chapter will introduce all research methods used.

1.9.1 Structured literature review

The first phase of this research is the literature study. The literature study form the foundation of the whole thesis research, therefore a firm and valid method is used for this phase. The method used for the literature study is the *structured literature review* approach proposed by Webster and Watson (2002). This method propose a structured process to find and use the right scientific literature for a valid scientific foundation.

The process starts by finding related knowledge from major contributions in the research domain using the most common scientific journal and article databases and search engines (i.e. Google Scholar, Citeseer). The whole literature study is concept-centric, meaning concepts determine the organizing framework of the review. In this case the capabilities are the concepts. Concept matrixes (Webster, Watson, 2002) are compiled during the study of each article. The concept matrixes contain concept-names and hold references to the articles in which the concept is handled.

The second step in the process continues by finding more literature by *backward* and *forward reviewing* (Webster, Watson, 2002). Both techniques uses the major contributors from the first step to find more related literature. *Backward reviewing* is determining prior articles that could be of interest by using the citations from the articles from the first step. *Forward reviewing* on the other hand, uses the article itself to determine other articles that cited the article from the first step. This second step in the process continues until (almost) no new concepts are found anymore. During the second step, the concept matrixes are further completed.

During the structured literature study a variety of capabilities are found and inserted into the concept-matrix, resulting into a matrix with capabilities and numerous references. The key capabilities (i.e. the capabilities holding the highest amount of references) are extracted from the matrix by using a threshold. The steps are elaborated in more detail in chapter 2.1.1.

1.9.2 Quantitative analysis

The *literature review analysis* and also the *product review analysis* that are introduces in this thesis research, uses quantitative analysis to find useful knowledge in the findings. For valid analysis, both cases uses the statistical analysis methods proposed in Field (2009), with using cumulative thresholding on the compiled concept matrixes from the *literature review* and more advanced quantitative analysis for the results of the *product review.*

The cumulative thresholding method is often used for literature reviews where concepts are centric (Webster, Watson, 2002). By using a cumulative threshold, only the major concepts of the literature review are provided, ignoring the less significant concepts. In this case the major capabilities of BI and cloud computing were extracted. The approach is elaborated in more detail in chapter 2.1.2.

The quantitative analysis of the product review is performed using statistics. Analysis of the product review first categorize the major capabilities into groups of the same cumulative appearance, then use standard statistical methods for identifying possible patterns and finally use statistics formulas to prove (with a certain probability) that the patterns exists in the results. The analysis is used as input for the development of the BIaaS capability model. This method is elaborated in more detail in chapter 3.6.

1.9.3 Expert group sessions

A frequently used technique to evaluate models or to obtain useful business knowledge is the usage of (semi-)structured expert interviews as available in for instance Smith (1995). Although many researchers have proven that (semi-)structured interviewing also contain weaknesses, it is still a respected form for knowledge gathering in scientific research (Agarwal, Tanniru, 1990; Holtzblatt, Beyer, 1995). Group sessions, where a shared consensus is obtained between a group of experts, can have substantial value over individual semi-structured interviews, because the experts themselves analyze the findings instead of one (non-expert) analyst (Byrd, Cossick, Zmud, 1992; Agarwal, Tanniru, 1990). Because of the latter, this research used expert group sessions to obtain business knowledge which could not be obtained by the literature review. The experts sessions are setup by first let the experts perform a task individually. Afterwards the experts discuss the outcome of the task with the whole group and together form a consensus. The formed consensus is then used as a result for further research.

1.9.4 Maturity modeling

The most significant contribution of this thesis research is the development of the BIaaS capability maturity model. For the valid creation of the BIaaS CMM, the maturity model development framework is used proposed by deBruin, Rosemann, Freeze and Kulkarni (2005).

The development framework contains six phases; *scope, design, populate, test, deploy and maintain*. Only the first three phases of the framework are used in this thesis research, the latter three are related to the actual usage of the developed maturity model, which is out of scope of this thesis research and is left for further research.

The first phase in the development of a maturity model is to determine the *scope* of the desired model (e.g. domain specific, general, particular stakeholder, etc.). Scoping is important to distinguish the proposed model from other existing models and will also determine te specificity and extensibility of the model. For this research the focus is aimed on the BIaaS domain and uses the developed BIaaS capability model as input.

The second phase is the *design* phase, where a design or architecture for the model is determined, which forms the basis for further development and application. In particular, the design of the model should incorporates the needs of the intended audience and how these needs will be met. The needs are reflected in *why* they seek to apply the model, *how* the model can be applied, *who* needs to be involved in applying the model and *what* can be achieved through application of the model. Another important part of the *design* phase is the introduction of maturity stages. The number of stages may vary, however, the stage names and definitions should be clear. The maturity stages can be developed using either a *top-down* or a *bottom-up* approach. The *top-down* approach first introduces the definitions and then measures are developed to fit the definitions. With the *bottom-up* approach the requirements measurements are determined first and then definitions are written to reflect these. The *bottom-up* approach is used in this thesis research.

The third and, last phase for this thesis research, is the *populate* phase. This phase populate the model by focusing on the *scoped* domain and using the *design* from the previous phases. In this phase it is necessary to identify *what* needs to be measured in the maturity assessment and *how* this can be measured. The *what* is answered by the capabilities from the BIaaS capability model, which is developed prior to the development of the maturity model. The *how* is calculated by the product review analysis which is elaborated in more detail together with development of the maturity model in the third chapter.

1.10 Scope

This research introduce a model of what BIaaS entails and a guideline for BIaaS solution vendors to assess their own BIaaS solutions. Important is to understand BI and Cloud Computing before going into depth with combining the two concepts into a new model for BIaaS. Therefore the research scope can be distinguished into two main research goals. The *first* is to define both the concepts of Business Intelligence and Cloud Computing. This is done by defining the main capabilities of those two concepts from literature and practice. The *second* goal is about defining and using BIaaS. For this second goal of this research a BIaaS Capability model will be introduced accompanied by a capability maturity model (CMM).

The latter research goal focusses on a more specific goal, which is to introduce a new capability model for BIaaS to define the conceptual model for BIaaS and a BIaaS capability maturity model to assist BIaaS solution vendors. The CMM is seen as the most significant deliverable of this thesis and is intended to be used by vendors in the future. The CMM can be used by vendors for software process assessment and software capability evaluation.

1.11 Motivation

BI is currently the top-most priority of many chief information officers, because BI has become an essential component of the information supply infrastructure and a contributor (and prerequisite) to the overall organizational success (Wixom, Watson, 2010). Therefore many organizations want to invest into new BI capabilities to distinguish themselves and give them an edge on their competition in the changing and competitive market nowadays (Porter & Millar, 1980; Clemmons & Row, 1991). Combining BI with Cloud Computing capabilities and introducing BIaaS solutions provides these new BI capabilities, but also will reduce organizational costs (De Marco et al., 2010; Thomson & van der Walt, 2010). This provides BIaaS solution vendors with a rich sales market to sell their BIaaS solutions.

The primary goal of this research is to provide knowledge about BIaaS. From a scientific point of view, the produced capability model and CMM for BIaaS solutions certainly add value to existing BI, cloud computing and BIaaS scientific literature, as this research deliveres the first BIaaS capability and CMM available. The CMM also contribute from a practical business perspective, where vendors can use the CMM for software process assessment and software capability evaluation. The CMM provides detailed information for vendors to improve their existing solutions to make them more BIaaS mature. Additional, the CMM can be used as a roadmap for BIaaS solution development and improvement, which is also likely to be quite useful for BIaaS vendors. This thesis research does not include a field research, however, the developed BIaaS Capability Model and BIaaS CMM contain substantial information for further BIaaS research and significantly aid BIaaS vendors in the development process.

1.12 Thesis Outline

The outline of this thesis document is structured as follows:

Chapter 1 provides the reader with background information about the thesis subject. Furthermore the reason for doing this thesis project is explained, the main goal, scope and approach to perform the research is also included.

Chapter 2 elaborates on capabilities, why capabilities are used and how current BI and cloud computing capabilities are merged into one capability model for BIaaS. All capabilities are described in depth to make a conceptual model of BIaaS.

Chapter 3 introduces a capability maturity model for BIaaS and explains how this maturity model can be used to asses BIaaS solutions. An extensive elaboration into the different maturity levels is proposed to explain and provide a vendor with the knowledge to increase the maturity level of their BIaaS solutions. This chapter also elaborates on the product review that is conducted and gives a full analysis.

Chapter 4 discusses the conclusion of the thesis research and answers the research questions. Also possible future research is proposed.

2 THE BIAAS CAPABILITY MODEL

Buyya, Pandey and Vecchiola (2009) stated that "Cloud computing offers significant benefit to IT companies by freeing them from the low level tasks of setting up basic hardware (servers) and software infrastructures and thus enabling them to focus on innovation and creating business value for their services". Knowing organizations can benefit significantly from using Cloud computing and the research of Watson and Wixom (2007) which concluded "BI is currently the top-most priority of many chief information officers", combining BI with Cloud computing capabilities has a significant business and social relevance. This can be further increased because most BI solutions take a long time to install, build and deploy. The average implementation time for the larger BI solutions is about three to six months or longer (Zeng et al., 2006). A standard BIaaS solution can decrease this implementation time because installation, build and deployment of the software aren't necessary anymore, because BIaaS solutions can be deployed on demand. The only implementation time necessary for a BIaaS solution is the data conversion from original data into the pre-installed data warehouse of the BIaaS solution, which is mostly done through predefined ETL packages.

In SaaS, applications are stored in data centers and customers can use the software as they require via the internet. Because of this reason, SaaS is often referred to as 'on-demand' software (Abdat, Spruit, Bos, 2011). The software is no longer delivered as a product (physical object), but is made available as a service. Therefore the term 'product' does no longer fit the world of SaaS and the term 'solution' is being used instead. BIaaS is the merger of BI with SaaS, or better said, BI offered as a service on the internet. Therefore the term BIaaS 'solution' is used in this document.

This thesis research introduces a new term, Business Intelligence as a Service or BIaaS. To describe and explain what BIaaS entails, a conceptual model of BIaaS is developed. This chapter will introduce the conceptual model of BIaaS in the form of a BIaaS capability model. This capability model contains the most essential building blocks of every BIaaS solution. By developing the BIaaS capability model the first research question "*What are BIaaS capabilities and how do they differ from conventional BI capabilities?*" will be answered.

To describe the competences of software, often the term *features* are used. A feature describes one specific technological task the software product or solutions can handle. In this thesis the term capability is used instead, to describe the competences of BIaaS. A capability is a higher level construct of measurement than features, defined as 'a set of features'. More specifically, a capability is the ability of a solution to "*perform a set of co-ordinated tasks, utilizing technological resources, for the purposes of achieving a particular end result*" (Helfat & Peteraf, 2003, p. 1000).
BIaaS is the merger of BI and SaaS competences, therefore to develop a BIaaS model you have to understand BI and SaaS individually. Because the objective is to develop a capability model of BIaaS containing the most important capabilities of BIaaS, the first step in the development process is find the most important BI and SaaS capabilities, which is done through an extensive literature study.

2.1 Literature study

The approach used in this thesis research to find BI and cloud computing capabilities is through a literature study. The reason to use this approach for finding capabilities is because of BI and SaaS are both research domains that are well documented in today's scientific literature. Reviewing the scientific knowledge available from published journals and scientific articles provides a scientifically sound view of BI and SaaS until present time.

2.1.1 *Literature review approach*

The literature review approach is depicted in the PDD shown in Figure 4. Table 3 en Table 4 describe the activities and deliverables respectively. The literature study is conducted using a structured approach also proposed by Webster and Watson (2002). The process starts by finding new capabilities from the major contributions in the BI and SaaS research domain using the most common scientific journal and article databases and search engines (i.e. Google Scholar, Citeseer). The capabilities found from the reviewed literature are inserted into a database. After insertion, references to the corresponding scientific journals and articles are connected to the capabilities in the database. More literature is found by backward and forward reviewing. Backward reviewing is determining prior articles that could be of interest by using the citations from the articles found in the first step. Forward reviewing uses the article itself to determine other articles that cited the articles from the first step (Webster, Watson, 2002). This process of finding capabilities continues until (almost) no new capabilities are found anymore.

Table 3: Activities and sub-activities from PDD of literature review

Activity	Sub-Activity	Description
Find capabilities	Review major contributions	Review published journals and articles that made a major contributions to the BI and SaaS (cloud computing) research domain.
	Review backwards	Review backwards by reviewing prior articles that could be of interest by using the citations from the articles found in the major contributors.
	Review forwards	Review forwards by finding articles that cited the major contributors.
Analyze findings	Extract important capabilities	Extract from the founded capabilities those capabilities that are mentioned the most often (above threshold). Those are the most important capabilities.
Describe capabilities		Define the most important capabilities using the reviewed literature.

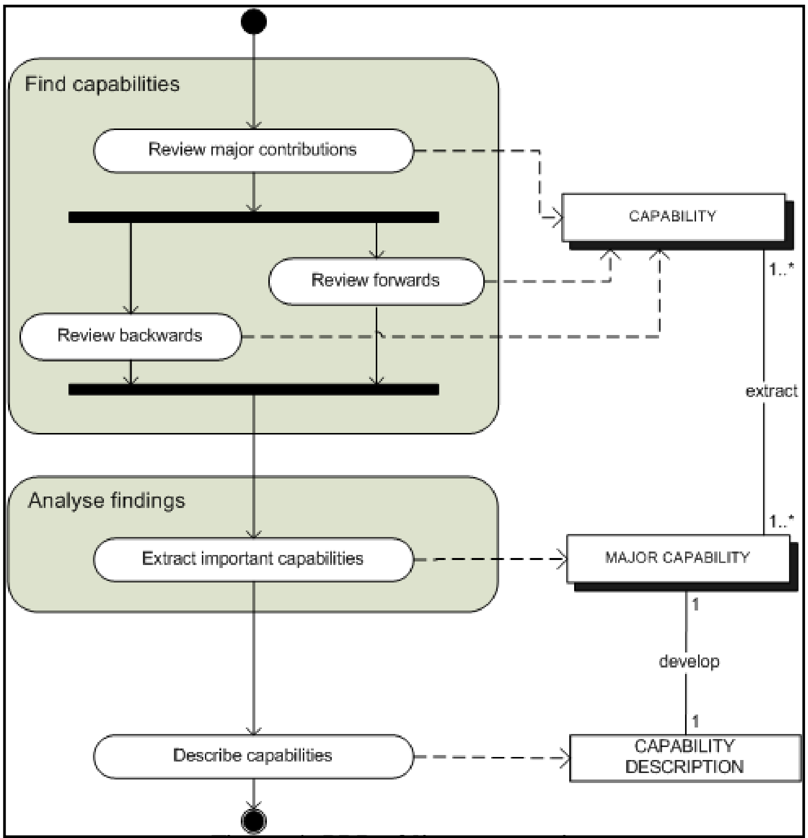


Figure 4: PDD of literature review

Table 4: Concepts (deliverables) from PDD of literature review

Concept	Description
CAPABILITY	Capabilities are defined by the reviewed literature. All capabilities found in literature are stored in the capability database. References to the corresponding scientific journals and articles are connected to the capabilities in the database.
MAJOR CAPABILITY	The most important capabilities are extracted from the found capabilities in literature by defining a threshold.
CAPABILITY DESCRIPTION	The capabilities are defined using the reviewed literature.

2.1.2 Literature review analysis

During the literature review a variety of BI and cloud computing capabilities are extracted. However not every capability is equally important for the BI or cloud computing research field. Based on the capability database storing all capabilities found from literature with corresponding references, supporting literature on BI and cloud computing and the input from experts (Table 5 and Table 11), a threshold is set for the minimum number of references for each capability. Using the threshold (≥ 5 for BI capabilities and ≥ 3 for cloud computing capabilities) the most important capabilities (e.g. key capabilities) are extracted from the database. Figure 5 illustrates with a Venn diagram how the BI and cloud computing capabilities stands towards each other.

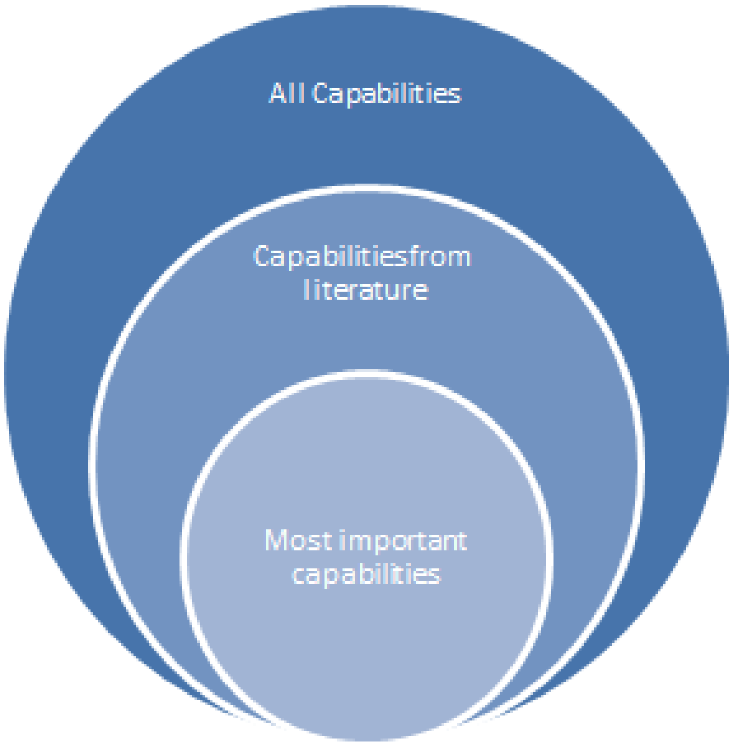


Figure 5: Venn diagram of capabilities

The key capabilities of each research domain are extracted from the database using a query output which provided all the capabilities above the set threshold. The following two sub-chapters elaborates on the key capabilities in BI and cloud computing that are found from the literature review. The sub-chapters show their title, main functionality, goal, understanding and references to related literature. These two sub-chapters also provide the prerequisites of each capability indicated with the letters A up to F, which will be further elaborated on in the following chapter introducing the capability maturity model (CMM).

2.1 BI focus areas

Best described by Watson and Wixom (2007) in their paper about the current state of BI, is that the basic idea of BI is "getting data in, and getting data out" (Watson & Wixom, 2007, pp. 96). Around this basic idea, lots of capabilities are developed to support BI. From the literature review and analysis, twenty-seven key BI capabilities are extracted from a variety of BI literature. The strongly coherent capabilities are grouped into *focus areas* containing capabilities supporting a specific area of the field of BI.

Table 5: BI experts participated in expert group sessions

Expert name	Title at Avanade	Experience in BI (years)
Erwin Haasnoot	Solution manager	10
Joris Valkonet	Project manager	7
Fabian Gutierres Ardts	Consultant	4
Wan Chi	Senior Consultant	5
Jeroen Schalken	Senior Consultant	5

The BI focus areas are formed after a structured expert group session. A group of five experts on BI (Table 5) were all individually asked to form coherent groups of capabilities with the capabilities from the literature research. After individually forming the groups, a groups discussion was setup to discuss each other's outcome. The experts then together form a consensus and develop the focus areas with corresponding names. Figure 6 graphically depict the focus areas formed and named by the experts, which support the idea of "getting data in, and getting data out"-process (Watson & Wixom, 2007) and are labeled *data gathering, data management, data processing, data analysis*, *consumerization* and *alerting*. These formed strongly coherent capability groups are also partly supported by the main BI areas opposed by Elena (2011) in her article about business intelligence.

The following sub-chapters elaborate on every major BI capability belonging to their corresponding BI focus area. Every focus area contain a number of capabilities, which are given a letter A until F. The letter given for each capability in a focus area is of importance for the BIaaS Capability Maturity Model where the letters are used as a reference to each capability. The use of the letters is further explained in detail in chapter 3.

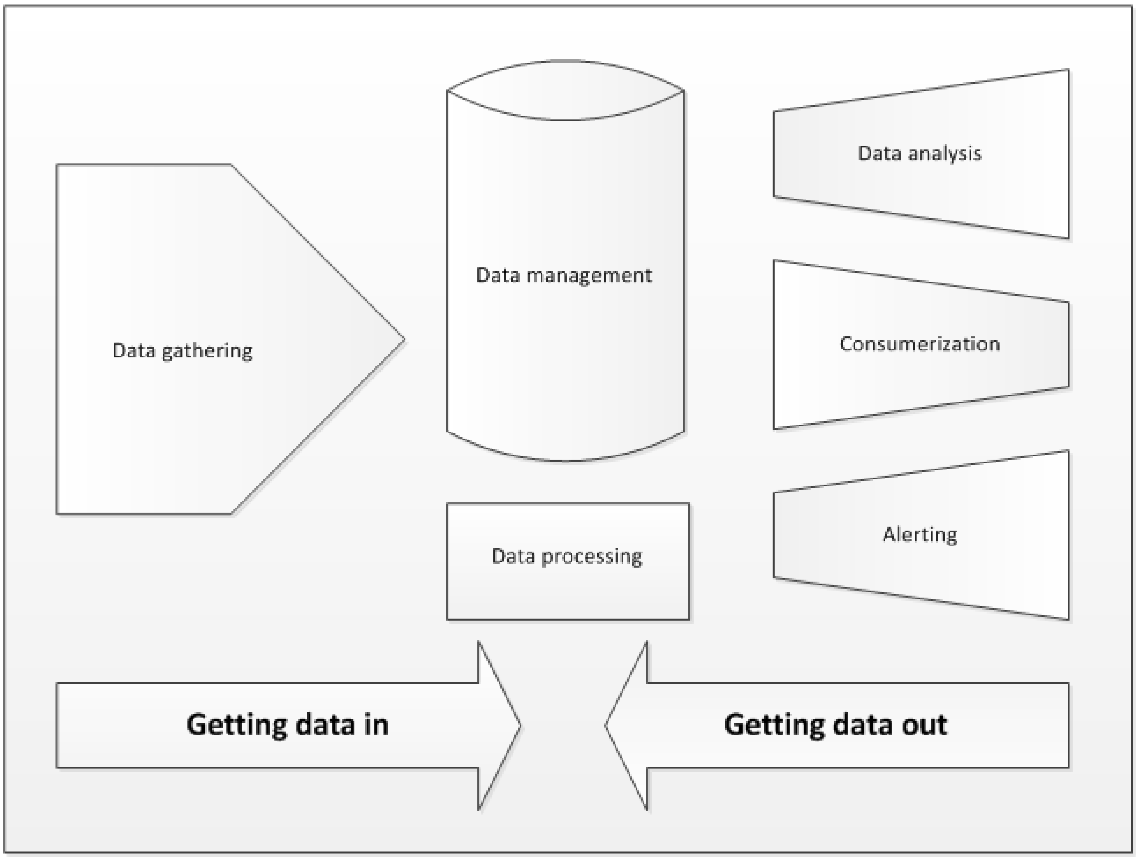


Figure 6: Capability units supporting the basics of BI: "Getting data in, and getting data out".

2.2.1 *Data gathering*

The data gathering focus area is about getting data in. Getting data in, or traditionally referred to as data warehousing, involves moving data from a set of source systems (e.g. an ERP system) into an integrated data warehouse. The source systems typically represent heterogeneous technical platforms and data structures (Watson & Wixom, 2007). The following capabilities are concerned with getting data into the data warehouse.

A. *Extract Transform and Load (ETL):* Before data can be analyzed for business purposes, the data should be inserted into a data warehouse. The ETL capability is concerned with automated extraction of data from a source system(s), transform this data into useful information (i.e. following a predefined format) for the target data warehouse and load this transformed data into the data warehouse (Thomsen & Pedersen, 2008; Golfarelli, Rizzi, Cella, 2004; Zeng et al., 2006).

B. *Data independency*: The most source systems, for instance ERP systems, contain structured data. This means the data is already structured in some way, making it easier to extract, transform and load into a data warehouse. However there also exist a lot of unstructured and semi-structured data stored in for instance text, images, business processes, web pages, news items, tweets, emails etc.. The evolved BI products nowadays can handle structured, but also semi-structured and unstructured data by automated software (Negash, 2004).

C. *Data follow-through workflow*: The traditional BI products depend on data analysts and consultants spending many hours to get the right data from a source system into the data warehouse, sometimes making use of additional data integration and migration software or ETL. The modern BI solutions contain features to automate this process and seamless getting data from the source system(s) into the data warehouse making use of an easy-to-use workflow, letting the user fully control the data transportation process (Negash,2004; Langseth & Vivatrat, 2003).

2.2.2 Data management

The data from the source systems as well as the data extracted from those systems should be managed properly to provide and analyze the best business information in BI products. This focus area contains the abilities to manage all the data to, in and from the data warehouse.

A. *Data Warehousing:* A data warehouse is a database that contains consolidated and transformed data (e.g. from *data input*) that is subject-oriented, integrated, time-variant end non-volatile (Inmon, 1992; Watson & Wixom, 2007; Chaudhuri & Dayal, 1997). The *data warehouse* capability is responsible for storing the right data in the right format for later BI processing. The data warehouse is the heart of your BI solution where all the information extracted from. The data warehouse is a necessary capability in a BI solutions (Langseth & Vivatrat, 2003; Negash, 2004; Zeng et al., 2006; Watson & Wixom, 2007; Inmon, Strauss, Neushloss, 2008; Kimball et al., 2008; Golfarelli, Rizzi, Cella, 2004; Chaudhuri & Dayal, 1997). Data warehouses contain consolidated data from several operational source databases, over potentially long periods of time; therefore they tend to be orders of magnitude larger than operational databases and projected to be hundreds of terabytes in size. The workloads of data warehouses are query intensive with mostly ad hoc and complex queries. The queries in a data warehouse can access millions of records and perform a lot of scans, joins an aggregates. Query throughput and response times are therefore more important than transaction throughput (Chaudhuri & Dayal, 1997). The data warehouse can be drilled-down into a data mart (i.e. smaller data warehouse), which contains specific data for specific BI output (Negash, 2004).

B. *Secure data delivery:* Organizations own a lot of data, to name a few, sales, employee, customer, business, marketing, etc.. Most of this data is for the organizations's eyes only, some are legally fixed like the private data of its employees, others contain sensitive business information that can influence the organizations's market position. Because the data is used in Business Intelligence, the security of the data needs to be fixed. This means assuring the privacy and integrity of private information, accuracy of public information and avoiding unintentionally revealing information that ought to be private (Boncella, 2003; Negash, 2004). These concerns can be managed through standard network security protocols and method (Boncella, 2000; Boncella, 2002), but also requires more sophisticated internet security methods against web defacing, web page hijacking, cognitive hacking and negative information (Boncella, 2003; Cybenko, Giani, Thompson, 2002; Hulme, 2003; Krasnow, 2000).

C. *Data Quality:* People do not rely on data that they don't trust; therefore it's important to maintain high-quality data (Watson & Wixom, 2007). High-quality data is reached by having correct, valid, integrated and in-time data (Zeng et al., 2006; Negash, 2004). These suggested terms for data quality can be assured in multiple ways. In either case it is not only important to have a well-designed connection with the source system (e.g. which is setup at *data input*) and import the correct data into the data warehouse, but also have a well-designed data model to extract the correct data and monitor this connection intensively to detect problems as early as possible. Integrated and in-time data can be for instance managed by setting up a (almost) real-time data warehouse with a direct connection with the source system, this ensures users from having the most up-to-date data at all times. These terms suggest that the data quality capability is not only about the technical specifications of the BI application, but also involves people and processes to maintain and guarantee the quality at all time. (Golfarelli, Rizzi, Cella, 2004; Negash, 2004; Watson & Wixom, 2007).

D. *Master Data Management:* Poor data quality in the source systems, results in poor data quality in the data warehouse, this can have multiple causes like wrong or insufficient data input, politics around data ownership or legacy technology (Watson & Wixom, 2007). *Master data management* is concerned around managing the master data (i.e. the data from (multiple) source systems) and keep the data quality in the source system(s) as high as possible. The same as with data quality for the data warehouse, managing master data can be done by automated technological means (i.e. using monitoring tools), but moreover includes setting-up processes, people and comities to ensure a continuous high value of master data quality (Watson & Wixom, 2007; Valkonet, 2011).

E. *Intelligent Warehousing:* Traditional relational servers (e.g. ERP systems) are not geared towards the intelligent use of indices and other requirements for supporting multidimensional views of data and extensive querying (Chaudhuri & Dayal, 1997). *Data warehousing* partly solves this problem by extracting useful data from the source systems and store the data into the data warehouse, which is specifically developed for data exploration by extensive querying. Many queries over data warehouses require summary data and therefore use aggregates. Hence, materializing summary data can help to accelerate many common queries (Chaudhuri & Dayal, 1997). For example, in a consultant organization calculating the employee productivity, the majority of queries may be based on the employee chargeability of the most recent month (or quarter) in the current fiscal year. Having summary data on these parameters can significantly speed up query processing. The technological features in the *Intelligent warehousing* capability can be summarized as an optimized data warehouse that deals with finding this common used data and summarize this in appropriate views, but also is concerned with optimizing queries and finding the most efficient way for getting the right data for the BI output (Zeng et al., 2006; Valkonet, 2011). Finding correct data in the most efficient way can be done for instance by implementing several techniques. For instance using the technique of minimal generators to narrow the search for the appropriate view from a set of candidate materialized views (Chaudhuri & Dayal, 1997).

F. *Information Management*: Information management is a continuous *process* of managing the information flow from the source system, through the data warehouse and processing, into understandable information output for organizations decision makers. The information management capability ensures that the right information data extracted from the source system, is processed into the right decision data and finally is translated into correct and clear information output that correspond to the business needs (Golfarelli, Rizzi, Cella, 2004; Negash, 2004). This process of information management continuously monitors and measures the business needs for BI and acts on these needs by adjust the data warehouse data input, change the processing of the source data or implement new information output (i.e. adjust or develop new KPI's or reports) (Golfarelli, Rizzi, Cella, 2004).

2.2.3 Data processing

A large part of Business Intelligence is about pre-processing data into actionable decision support data. The focus area of *data processing* is concerned with the abilities of processing data in the data warehouse and prepare the data for use in some form of BI output.

A. Data preprocessing: For high quality BI, data should be preprocessed into a particular form to be used in the BI product. Most data preprocessing is in the form of data cleaning, which involves dealing with missing information, and noisy data (Zeng et al., 2006). Noisy data includes incorrect attribute values, duplicate records and data smoothing. The most data preprocessing takes place before inserting into the data warehouse (e.g. the *data input* capabilities), but after insertion, data preprocessing is still an important capability to process data in a form for most efficient en high quality data which ensure higher quality and faster delivered data output.

B. Data transformation: The primary objective of Business Intelligence is delivering actionable decision information for decision makers (Zeng et al., 2006; Negash, 2004). Therefore transforming data into decision information is an important and vital capability of a BI solution. The *data transformation* can be seen as an extension of the *data preprocessing* capability. Where the *data preprocessing* capability delivers the data in the right form for best data usage, the *data transformation* is primary focussed on transforming the data in correct and valid decision information which can be directly used in *data output* capabilities.

C. Segmentation and clustering: Frawley, Piatetsky-Shapiro, Matheus (1992) estimated, as early as 1992 that the amount of information in the world doubles every twenty months. The growth in the size and the number of existing databases far exceeds human abilities to analyze such data, thus creating both a need and an opportunity for extracting implicit, previous unknown and potentially useful information (e.g. knowledge discovery) from these databases (Han, Cai, Cercone, 1992). The large amount of data intensifies the exploratory data analysis process, making it harder to find useful data, resulting for instance in performance decrease or missing potential useful information. One solution for this problem is *segmentation and clustering* (Negash, 2004; Watson & Wixom, 2007; Blumberg & Atre, 2003a; Zeng et al., 2006). *Segmentation and clustering* techniques use unsupervised classification of patterns in data (observations, data items, feature vectors) to segment the data into relevant data groups, called clusters (Zhang, Ramakrishnan, Livny, 1996). The clustering problem has been addressed by researchers in many disciplines and contexts, which reflects its broad appeal and usefulness as data processing capability to support the data exploration capabilities (Jain, Murty, Flynn, 1999; Zhang, Ramakrishnan, Livny, 1996; Chen et al., 2009; Porter, 1998). Segmentation and clustering capabilities are very popular for the use in *geographic information systems* (GIS) by linking data with electronic maps and thereby provide the ability to analyze spatial phenomena (Negash, 2004; Langseth & Vivatrat, 2003).

D. *Automated learning and refinement:* Segmentation and clustering techniques requires database analysts to implement and refine changing specifications. The *automated learning and refinement* capability will coop with changing specifications (e.g. changing fiscal year and thereby changing target clusters) and automatically change the segmentation and clustering specifications. Another example of a learning and refinement technique was introduced by Han, Cai and Cercone (1992), who introduced an automatic knowledge discovery learning algorithm. Automatic learning and refinement techniques are a useful capability in maintaining a high quality and pro-active data warehouse (Negash, 2004; Langseth & Vivatrat, 2003).

2.2.4 Data analysis

The early BI systems were just simple interactive *data analysis* products, which were designed to manually explore data to find meaningful and useful patterns to make decisions upon. The growth in the amount of data nowadays far exceeds human abilities to analyze the data without sophisticated tools. The data analysis focus area contains the abilities to aid decision makers in exploring data by for instance (predefined) querying and/or using data analysis interfaces and present complex and competitive information to make (strategic) decisions upon.

A. *Analytics:* Exploring data, extract useful information, and analyze the data for decision making is a hard process, especially for decision makers that often do not have expert knowledge about data exploration. For that purpose analytical capabilities are developed, which include predefined programs that build quantitative processes for a business to arrive at optimal decisions and to perform knowledge discovery (Shobrys, 2003; Negash, 2004; Watson & Wixom, 2007; Zeng et al., 2006; Elena, 2011). *Analytical* capabilities are often developed industry-specific for organizations to analyze (potential) customers in specific target areas (Elena, 2011).

B. *OLAP:* The data in the data warehouse is typically constructed to provide multidimensional and multi-level view, to facilitate complex analyses and visualization from different perspectives and with multiple granularities (Chen, Yan, Zhu, Han, Yu, 2008). A tool for fast and user friendly analysis of the multidimensional data in the data warehouse is Online Analytical Processing or shortened *OLAP* (Chaudhuri & Dayal, 1997; Chen et al., 2008). The key operations available in OLAP include rollup and drill-down along one or more dimension hierarchies, slice-and-dice, and pivot. Rollup will perform generalization to see a concise overview of the data and drill-down will decrease the level of aggregation to specialize on a particular part of the data and thereby increasing the level of detail. Slice-and-dice will focus on a particular aspect of the data by selection and projection. And the last, Pivoting, will be used for re-orienting the multidimensional view of data (Chaudhuri & Dayal, 1997; Chen et al., 2009). *OLAP,* together with *Data Warehousing* and *Data Mining* are seen as the most important BI capabilities (Zeng et al., 2006; Shobrys, 2003; Golfarelli, Rizzi, Cella, 2004; Watson & Wixom, 2007; Chaudhuri & Dayal, 1997; Gray et al., 1997).

C. *Data Mining*: Hand, Mannila and Smyth (2001, pp. 2) define the scientific field of *data mining* as: “The science of extracting useful information from large data sets or databases”.

This exactly depicts what the *data mining* capability encompasses. The *data mining* capability gives you the ability to explore, observe, and analyze large amounts of data in the data warehouse and extract useful information from it, which can be done by direct querying or using a data mining interface. Many researchers address *data mining* as one of the basic abilities of BI (Negash, 2004; Langseth & Vivatrat, 2003; Zeng et al., 2006; Shobrys, 2003; Golfarelli, Rizzi, Cella, 2004; Chen, Han, Yu, 1996; Hand, Mannila, Smyth, 2001; Elena, 2011).

D. Forecasting: For many organizations predicting the future is a powerful tool; "knowing" when and where to invest can gain higher market share and keep you one step ahead of your competitors. For instance in supply chain management forecasting can lower stock costs in predicting when stock needs to be supplemented and when you need to lower stock (Shobrys, 2003). *Forecasting* is the capability which extends *data mining* and *analytical* capabilities with prediction algorithms to predict the future (Hand, Mannila, Smyth, 2001). *Forecasting* uses given data sets and predefined algorithms (for instance Bayesian or other prediction algorithms and techniques) to predict the future using past data in the data sets.

E. Text mining: Text mining is an advanced data mining technique for semi-automated discovery of new knowledge from text (documents) (Hand, Mannila, Smyth, 2001; Swanson, 1987; Nasukawa, Nagano, 2001). This capability has gained more and more popularity (again) in the past few years. This is triggered by the the increasing usage of social media and other opinion-rich resources (Pang, Lee, 2008; Corley, Cook, Mikler, Singh, 2010). Analyzing these resources using text mining can give organizations insight into their market position (sentiment analysis) and can be a powerful marketing tool for positioning future advertisement campaigns (Corley, Cook, Mikler, Singh, 2010).

F. *Data Modeling*: Understanding relationships between data (sets) and how they affect each other, is a hard task for humans to perform. *Data modeling* is a powerful capability for visualization of the relationships and dependencies between data (Zeng et al., 2006). An interesting issue where often *data modeling* can give good insight, is in understanding relationships and dependencies which are related to the need for defining a consistent set of indicators, which can be used for instance in KPI's (Golfarelli, Rizzi, Cella, 2004).

2.2.5 Consumerization

The capabilities in this focus area are responsible for giving information to the user about analyzed data. The most capabilities are not interactive, but are just about providing the right information in the best possible way.

A. Reporting: Reporting capabilities like KPI's, dashboards, scorecards, story boards etc. are seen as the fundamental capabilities of a BI solution (Shobrys, 2003; Zeng et al., 2006; Watson & Wixom, 2007; Valkonet, 2011). KPI's, scorecards and dashboards for instance, visually summarize large amounts of data related to organizational performance; therefore many organizations implement these reporting types as key components of BPM initiatives (Watson & Wixom, 2007; Zeng et al., 2006; Golfarelli, Rizzi, Cella, 2004). Users can quickly see (in a single or few screens) how actual performance compares to goals, benchmarks and previous performance. Ideally *reporting* capabilities are (near) real-time to provide users with the newest information possible.

B. Data visualization: Because BI is mainly giving useful information about collected data, followed by the interpretation and evaluation of this data to use it for taking decisions; *data visualization* is an important and useful tool in this process (Negash,2004; Langseth & Vivatrat, 2003; Zeng et al., 2006; Chen et al., 2009; Theus, 2002). *Data visualization* is the set of techniques used to turn a set of data into visual insight. It aims to give data a meaningful representation by exploiting the powerful discerning capabilities of the human eye, therefore making it possible to analyze data easier and faster (Fisher, 2011). Interactive statistical data visualization reaches beyond the limits of the former static visualizations, making it increasingly popular (Theus, 2002). Interactive exploratory data visualizations is therefore currently a popular research field which is addressed by many researchers (Theus, 2002; Roberts, 2007).

C. Collaboration: Decision making in all levels of an organization is not a one person task, the multiple levels in an organization have often different managers and are interested in different sorts of BI or have different locations. To work together as one organization collaboration between the different dimensions in an organization is important, therefore *collaboration* capabilities should be in place to share the organizations information about BI (Cody, Kreulen, Krishna, Spangler, 2002). Collaboration capabilities can be “simple” like publishing or exporting reports, but can also be more advanced like inter-software collaboration by for instance using open standards (Thomsen & Pedersen, 2008).

D. Self-Service BI: Serve the broad population on multiple levels within an organization with BI is often referred to as “BI for the masses” (Negash, 2004). The new class of analytical capabilities that can serve “BI for the masses” is often called *self-service BI* where users create their own BI reports (e.g. creation of BI by non-specialists) (Negash, 2004). *Self-service BI* is the ability where users can create their own reports with information which is relevant for their own purposes, by selecting necessary datasets with corresponding visualizations using specialized self-service tooling.

E. Portability: In Gartner’s webinar by Richardson (2011), they foresee that by 2013, 33% of BI functionality will be consumed via handheld devices. That same webinar also shows that five of the total nine emergent technologies impacting BI the most is consumer-behavior-driven. This depicts the importance to increase the usability to optimize BI success (Negash, 2004; Diallo, Badardt, Hubert, Daniel, 2011). Increasing usability can be done on multiple levels like for instance the integration into popular business software like Microsoft Office, the ability for online usage or making it available for mobile devices. But also less tangible factors are important for user experience like performance (does the application provide the information in considerable time) and relevance (does the application give me the information in the form that is relevant for that particular user).

F. *Business Process Embedding*: Formally BI was mainly available for people in the strategic level of an organization, but research showed that facilitating BI tools in multiple organizational levels makes organizations more likely to be successful (Watson & Wixom, 2007). Make BI available to a whole organization on multiple granularities supporting their particular BI needs, ultimately makes the organization more agile and more competitive by empowering all employees to make faster and better decisions based on common understanding of markets, customers, and the data that shapes the business (Zeng et al., 2006). Spreading BI and making a larger use base by providing users information they need to perform their jobs better, implies the need for BI tools to be specified for certain business processes. This capability specifies the BI information for particular business processes to aid those users performing their jobs better (Watson & Wixom, 2007; Negash, 2004). The main challenges of accomplishing BI provided at all levels of the organization (i.e. non-BI-specialists) are easy creation and consumption of reports, secure delivery of information and a friendly user interface (Negash, 2004; McKnight, 2003).

2.2.6 Alerting

Acting on important events is key while dealing with business intelligence. Alerting capabilities can increase the awareness for important business events. The capabilities in the alerting focus area deal with alerting users in-time on an appropriate manner to take action upon.

A. BI alerting: BI alerting capability consists of features where the BI solution automatically compares operational BI metrics to user-defines business thresholds and rules. Based on this comparison, the application can send an alert to a business user to warn them of a potential business problem, issue or opportunity that requires action (White, 2005). *BI alerting* reduces the need for users to constantly monitor business information and is seen as an important capability for proactive-BI (Negash, 2004; Langseth & Vivatrat, 2003).

B. *Automated exception detection:* Connections with multiple systems (source systems connected with a data warehouse), data dependencies, inter table references, system errors and many other possible causes can lead to system exceptions. The *automated exception detection* capability is the ability which contain the feature who automatically detect exceptions and inform the appropriate persons to fix the possible problems (Negash, 2004; Langseth & Vivatrat, 2003). Complex Event Processing (CEP) is a currently popular technology and is used in event-driven BI applications with goal of identifying meaningful patterns, relationships and data abstractions from among seemingly unrelated events and trigger immediate response actions (Wu, Diao, Rizvi, 2006; Robins, 2010).

2.3 Cloud Computing capabilities

'Cloud computing' is a general term which is often used as umbrella for solutions containing a form of Software as a Service (SaaS), Infrastructure as a Service (IaaS) and/or Platform as a Service (PaaS) (Vecchiola, Chu, Buyya, 2009). SaaS, IaaS and PaaS are the three pillars on top of which cloud computing solutions are delivered and thus cloud computing encompasses a bigger field of research. This thesis focusses mainly on the SaaS part of Cloud computing, although IaaS is also taken into account.

The main idea of SaaS is defined by Abdat, Spruit and Bos (2011, pp. 156) as a "delivery model that supports multi-tenancy in which the vendors host and operate their software on a data center (either independently or through third-party) and provide it to their customers over the Internet and typically on a subscription basis and/or pay-per-use basis". They constructed this definition by extracting the key elements of SaaS from twelve different published articles of scientific and business studies. This subchapter elaborates on the key capabilities from the field SaaS, gathered from the literature review. As with the BI capabilities, the SaaS capabilities are gathered in focus areas containing strongly coherent groups of SaaS capabilities.

The focus areas for the SaaS capabilities are formed using a different approach then with the BI focus areas. The same as for the former software that is hosted on premise of the companies or individuals, also Software as a Service is a software model (Mietzner, Leymann, 2008). Therefore focus areas already exist in current software development literature, which can be enriched with a service component which was not included in the former software models. A major contributor to the system analysis and design research field and worldwide appreciated, is the book of Dennis, Wixom and Tegarden (2002). Which is also republished in many later editions. They conclude that every application system (i.e. a SaaS solution) can be divided into four general functions: data storage, access logic, application logic and presentation logic. Renaming this four functions taking the definitions into account and adding the service component of SaaS, provides the five SaaS focus areas data storage, accessibility, application logic, usability and service. This sub-chapter elaborates on each of the SaaS focus areas and the corresponding group of capabilities found in the literature review. The same as for the BI capabilities a letter is provided for each capability in every focus area, which is used in the BIaaS CMM and is further explained in chapter 3.

2.3.1 Data storage

SaaS solutions require data to be stored and retrieved, the data storage focus area is responsible for the abilities facilitate hardware for data storage and actions to retrieve it.

A. *Data center:* Vaquero et al. (2009) conclude that the data center is the basic unit of cloud computing. They use the definition given in the article of McFedries (2008) for the data center center, which is conceived as a huge collection of servers and clusters, offering huge amounts of computing power and storage by using spare resources. With other words, a data center contains the physical resources which constitute the foundation of Cloud computing (Buyya, Pandey, Vecchiola, 2009). This statement is strengthened by Abdat, Spruit and Bos (2011) who also performed a literature review and made a base-definition for SaaS from the most important elements in SaaS literature (Hoch, Kerr, Griffith, 2001; Blokdijk, 2008; Greschler, Mangan, 2002; Sääksjärvi, Lassila, Nordstrom, 2005). A SaaS solution can be hosted from multiple data centers making use of *resource allocation* and *virtualization* capabilities (Buyya, Pandey, Vecchiola, 2009; Geelan, 2008; Vaquero et al., 2009; Buyya, Yeo, Venugopal, 2008).

B. *Multi-tenancy:* The capability of *multi-tenancy* is the significant key capability which makes SaaS different from other online applications (e.g. ASP) (Abdat, Spruit, Bos, 2011). *Multi-tenancy* is an architecture principle in the realm of the SaaS business model in which multiple customers ("tenants") share the same application and database instance (Bezemer & Zaidman, 2010; Mietzner et al., 2009; Abdat, Spruit, Bos, 2011; Turner, Budgen, Brereton, 2003; Guo, Sun, Huang, Wang, Gao, 2007). This architecture principle allows making full use of the economy of scale. *Multi-tenancy* major benefits are increased utilization of hardware resources and improved ease of maintenance, in particular on the deployment side (Bezemer & Zaidman, 2010).

C. *Resource allocation:* SaaS solution's data is retrieved from data center(s), accessible via the internet. Therefore the geographic location of the data center or that of the user is not relevant anymore for accessibility. However it can be important for performance to access data that is geographically the closest or has for instance a faster connection available. Also quality of service can be a factor to have your resources at a particular (or multiple) location(s). *Resource allocation* abilities manage the physical infrastructure at multiple geographical locations to serve the SaaS solution (Buyya, Pandey, Vecchiola, 2009; Vaquero et al. 2009; Geelan, 2008).

2.3.2 *Accessibility*

Using a solution requires access to the application and its data. This focus area contains the capabilties responsible for accessing the solution and the data used by the solution.

A. *Internet centric:* This capability contains the ability of accessing the SaaS solution remotely by an Internet connection. The design principle of cloud computing state that the solution is a hosted service, thus making an internet connection a mandatory requirement to be able to access the cloud computing solution (Abdat, Spruit, Bos, 2011; Vaquero, Rodero-Marino, Caceres, Lindner, 2009). The solution access therefore must be *internet centric* to access remotely over the Internet (Geelan, 2008; Buyya, Yeo, Venugopal, 2008; Abdat, Spruit, Bos, 2011; Vaquero et al., 2009).

B. *Hosted Service:* Before the concept of cloud computing, typically software vendors develop software which is delivered to customers for implementation on the customers' computers or servers (i.e. on-premise software). Cloud computing solutions are a *hosted service* provided at a site the vendor has chosen, which differentiates cloud computing from on-premise software by the means that the software is installed in data centers and is provided to the customer over the Internet.. Abdat, Spruit and Bos (2011) extracted important elements of SaaS definitions from different published articles of scientific and business studies. The *hosted service* was found as one of the key elements for the SaaS definition, proving it is a vital (mandatory) capability for a SaaS solution. The research of Vaquero et al. (2009) conclude the same by doing similar research on twenty published scientific articles.

C. *Data protection:* A major question organizations have with respect to cloud computing is: "is my data secure?". This is one of the biggest argument organizations have when they think of moving to cloud computing. Especially for cloud computing where your solution is most likely to run on a multi-tenant environment, a security breach can result in the exposure of data to other, possibly competitive, tenants. This makes security issues such as *data protection* very important (Bezemer, Zaidman, 2010). Data protection abilities support better isolation among tenants in many aspects but maintain the multi-tenant requirement of high share efficientcy (Guo et al., 2007).

D. *Resource optimization:* Resources used for the cloud solution can be optimized (maximized) by resource optimization abilities. Resources are automatically optimized to provide an appropriate runtime environment for the SaaS solution satisfying the customers QoS and utilize the physical resources at best (Buyya, Pandey, Vecchiola, 2009; Vaquero et al. 2009; Geelan, 2008; Vaquero, Rodero-Marino, Caceres, Lindner, 2009).

2.3.3 *Application logic*

The application logic focus area is responsible for the behavioral capabilities of a cloud computing solution.

A. *Scalability:* The cloud solution is typically *scalable*, meaning they can be easily scaled up or down depending on the demand from the customer (Vaquero, Rodero-Marino, Caceres, Lindner, 2009; Geelan, 2008; Buyya, Yeo, Venugopal, 2008; Abdat, Spruit, Bos, 2011). With scalability customers can change usage resources within a certain period of time (e.g. changing the amount of users). This means that customers are able to subscribe to and unsubscribe from an application at any time, which requires that the environment the application runs in can be scaled up and down automatically when a new customer subscribes or unsubscribes from the application. The compute model from a customer's point of view is referred to as on demand computing (Mietzner, Leymann, 2008; Rust, Kannan, 2003; Buyya, Pandey, Vecchiola, 2009).

B. *Vendor controlled:* The SaaS principle focuses on separating the possession and ownership of software from its use. Delivering functionality as a set of distributed services can overcome many current limitations constraining software use, deployment, and evolution (Turner, Budgen, Brereton, 2003). The SaaS design approach where the solution is designed by vendors lets the set of services a vendor provides evolve without any user intervention, as the user requirements and its context change. (Chou & Chou, 2008; Abdat, Spruit, Bos, 2011; Turner, Budgen, Brereton, 2003). After development the solutions needs to be managed (updates, bug removals, etc.) which in Cloud computing is typically done by vendors; the vendor owns the solution (Chou & Chou, 2008). This eradicates the need for customers to download patches and upgrades, do bug fixing, and plan releases; rather than with on-premise hardware and software where this is the customers responsibility (Abdat, Spruit, Bos, 2011). The management of the solution is a continuous process by the vendor where the solution is monitored, problems are solved (within SLA) and update are done in rapid deployment cycles (Dym, 2009; Abdat, Spruit, Bos, 2011).

C. *Virtualization:* Vaquero et al. (2009) depicted in their research by extracting a minimum definition from twenty published scientific articles on cloud computing, that virtualization is one of the four characteristic capabilities. Virtualization hides the heterogeneity of the underlying resources, therefore provides the ability to virtualize the sum of parts into a singular wide-area resource pool (Vaquero et al., 2009; Geelan, 2008; Buyya, Yeo, Venugopal, 2008). The virtualization capability covers the whole solution, meaning data (files, databases etc.), the computing resources and also the hardware resources are virtualized (Bégin, 2008). Virtualization technologies provide features such as application isolation, quality of service, and sandboxing. Among the different solutions for virtualization, the most popular are hardware level virtualization and programming language level virtualization (Buyya et al., 2009). Hardware level virtualization guarantees complete isolation of applications and a fine partitioning of the physical resources, such as memory and CPU, by means of virtual machines. Programming level virtualization provides sandboxing and managed executions for applications developed with a specific technology or programming language (i.e. Java and .NET). The next generation data centers are developed as a network of virtual services (hardware, database, user-interface, application logic), making users able to access and deploy applications from anywhere in the world on demand.

Virtualization technologies therefore help in creating an environment where different services are integrated (Buyya, Pandey, Vecchiola, 2009).

D. *Automatic adaption:* Immediate scalability and resources usage optimization are key elements for cloud computing (Vaquero, Rodero-Marino, Caceres, Lindner, 2009; Geelan, 2008). Immediate response on important changes in usage without user intervention is provided by automation of resources management (i.e. using policies and real-time triggers) and is referred to as automatic adaption (Gruman, Knorr, 2008; Geelan, 2008).

2.3.4 *Usability*

Usability is concerned with the presentation of the solution to the user and the acceptance of user input.

A. *Monitoring:* Scalability and resource management are key elements of cloud computing. These capabilities are partly regulated by the user intervention. *Monitoring* capabilities provide real-time information about the current status of the users domain (e.g. it's application pool or network usage), which can be used to react upon (Gruman, Knorr, 2008; Geelan, 2008; Vaquero, Rodero-Marino, Caceres, Lindner, 2009).

B. *User friendliness:* The goal of cloud computing is to develop a solution for multiple customers. When usage is for the masses (multiple customers), everyone should be able to use it without learning (ease-of-use). Thus the solution should be *user friendly* (Vaquero et al., 2009; Geelan, 2008; Buyya, Yeo, Venugopal, 2008). Making a solution accessible for masses, includes hiding as much technical details as possible (i.e. deployment details) and making the use as easy as possible (Vaquero et al., 2009).

C. *Configurability:* Cloud computing solutions typically are designed for multiple customers rather than only one. The automatic provisioning of applications is an important task for the success of SaaS providers (Mietzner, Leymann, 2008). To attract a significant number of customers, cloud computing solutions have to be *configurable* to fulfill the varying functional and quality requirements of individual customers (Mietzner, Metzger, Leymann, Pohl, 2009; Abdat, Spruit, Bos, 2011).

2.3.5 *Service*

An important part of cloud computing are the service focus area capabilities. This focus area contains the service abilities of cloud computing which are very different to the on-premise software delivery. The capabilities in this focus area are not particularly concerned with tangible artifacts or the generating user output as with the capabilities from the other focus areas, however, it facilitates the business aspects and qualities of cloud computing.

A. Time & location independent: Cloud computing solutions are hosted services, making them accessible anytime and from any place using an Internet connection. Vendors should be aware of this *time and location independency* and provide accessibility independent of time and location, including customer support (Abdat, Spruit, Bos, 2011; Buyya, Yeo, Venugopal, 2008).

B. *Payment model:* Vendors are the owners of the cloud computing solution. Users can 'rent' the service of the application, therefore users do not need to buy licensing or have to install the software on-premises (Abdat, Spruit, Bos, 2011; Chou & Chou, 2008). The users are charged using a payment model, which can be pay-per-use based on usage metrics, or using a subscription model (fixed fee) (Vaquero, Rodero-Marino, Caceres, Lindner, 2009; Geelan, 2008; Buyya, Yeo, Venugopal, 2008; Abdat, Spruit, Bos, 2011; Vecchiola, Chu, Buyya, 2009; Buyya, Pandey, Vecchiola, 2009; Chou & Chou, 2008). Pay-per-use pricing is based on usage metrics, where the users are charged for the actual usage of computing capabilities. With a subscription model, users pay a fixed fee on a monthly, quarterly, or annually basis. These payment models are the most commonly used, but there are always possibilities that vendors come up with other alternative charging models.

C. *Quality of service (QoS):* With cloud computing the vendor is the owner of the solution and the customer is using the service on-demand. On the contrary to on-premise hardware and software, the customer does not have influence on quality of the service provided by the vendor. The *quality of service (QoS)* capability refers to measurable qualities of service (i.e. infrastructure uptime) that are predefined by the vendor (intrinsic QoS), ensuring those *qualities of service* are provided (Buyya, Pandey, Vecchiola, 2009; Buyya, Yeo, Venugopal, Broberg, Brandic, 2009; Vaquero et al., 2009).

D. *Service level agreements (SLA):* The level of service that is provided by the cloud computing provider is specified in a contract which is called *Service Level Agreement (SLA)* (Vaquero, Rodero-Marino, Caceres, Lindner, 2009; Geelan, 2008; Buyya, Yeo, Venugopal, 2008; Buyya, Pandey, Vecchiola, 2009; Hoch, Kerr, Griffith, 2001; Abdat, Spruit, Bos, 2011). SLAs can specify a variety of specifications like bandwidth availability, response times for routine and ad hoc queries and response time for problem resolution (network down, machine failure, etc.). They can be very general or extremely detailed, including the steps taken in the event of failure (Hoch, Kerr, Griffith, 2001). SLA can also be categorized into different domains, like for instance application SLAs and infrastructure SLAs, where the different SLAs specify the level of service for the different parts of the provided services (Vaquero, Rodero-Marino, Caceres, Lindner, 2009; Geelan, 2008; Buyya, Yeo, Venugopal, 2008).

E. *Competitive costs:* SaaS providers deliver software on-demand. A mature SaaS solution is virtualized and is able to deploy applications on-demand and from anywhere in the world. This easy on-demand creation and deployment of applications is the reason why SaaS solutions are provided for competitive costs (Abdat, Spruit, Bos, 2011; Buyya, Pandey, Vecchiola, 2009; Buyya, Yeo, Venugopal, Broberg, Brandic, 2009). Evolved SaaS solutions are offered to a wide range and high number of customers by using a single unified solution. Unification removes the complexity of dealing with multiple customer requirements which lower costs significantly (Buyya, Pandey, Vecchiola, 2009).

2.4 Combining BI with SaaS

In the previous sub-chapters the key capabilities of BI and Cloud Computing are introduced, explained and ordered in focus areas. This sub-chapter combines the capabilities into one capability model, that will form the conceptual model of BIaaS. Unfortunately there is not one technique available from capability modeling literature to combine this two fields. However, analyzing the meaning (definition), contents and focus areas of BI and SaaS provide a fortunate outcome.

BI refers to computer-based techniques providing historical, current and predictive views of business operations (Elena, 2011), SaaS on the other hand is a software model (Abdat, Spruit, Bos, 2011) containing different capabilities and different focus areas than BI. Conclusion, BI and SaaS capabilities and focus areas are so different that it is not necessary to integrate them, but can be used next to each other instead, to form the BIaaS capability model. The result is shown in Figure 7 and is referred to as the BIaaS conceptual model. Figure 7 include all BI and cloud computing focus areas and capabilities, where “BI” and “CC” provide the origin of each focus areas. A matrix form is chosen to visualize al focus areas and capabilities in one compact overview.

Recall the first research question: “*What are business intelligence as a service capabilities and how do they differ from conventional business intelligence capabilities?*”. This chapter explicitly provide the answer to the first part of the research question by providing the BIaaS capability model which include the key capabilities of BIaaS. The definitions of the individual capabilities are explained in the BI and SaaS sub-chapters. The second part of the research question is depicted by the BIaaS capability model, where the key BI capabilities are enriched with the key SaaS capabilities. Thus, they differ from the conventional BI capabilities in the fact that they also include SaaS capabilities.

Design		
Data Storage (CC)	**Accessibility (CC)**	
Data center	Hosted service	
Multi-tenancy	Internet centric	
Resource allocation	Data protection	
	Resource optimization	
Data		
Data Gathering (BI)	**Data Management (BI)**	**Data Processing (BI)**
ETL	Data warehousing	Data preprocessing
Data follow-through workflow	Data quality	Data transformation
Data independency	Information management	Segmentation and clustering
	Intelligent warehousing	Automated learning and refinement
	Master data management	
	Data security	
Data Analysis (BI)		
Analytics		
Data mining		
Data modeling		
Forecasting		
Text mining		
OLAP		
Usage		
Application Logic (CC)	**Usability (CC)**	**Consumerization (BI)**
Virtualization	Configurability	Collaboration
Vendor controlled	Monitoring	Business process embedding
Scalability	User friendliness	Data visualization
Automatic adoption		Reporting
		Self-service BI
		Portability
Support		
Alerting (BI)	**Service (CC)**	
Automated exception detection	Time & location independent	
BI alerting	Quality of service	
	Payment model	
	Competitive costs	
	SLA	

Solution unit **Focus area** **Capability**

Figure 7: The BIaaS capability model.

3 THE BIAAS MATURITY MODEL

Paulk et al. (1993, pp. 1) stated about process improvement in software organizations, "Setting sensible goals for process improvement requires an understanding of the difference between immature and mature software organizations". The same is true for developing new software products and the maturity of the developed products. Maturity describes a "state of being complete, perfect or ready" (Simpson & Weiner, 2011). To reach a desired state of maturity, an evolutionary transformation path from an initial to a target stage needs to be progressed (Fraser, Moultrie, Gregory, 2002). Maturity Models (MM's) are used to guide this transformation process.

Initially proposed in the 1970's (Gibson & Nolan, 1974), over hundred MM's have been published in the field of Information Science (Becker, Knackstedt, Pöppelbuß, 2009; Mettler & Rohner, 2009; Bekkers, Weerd, Spruit, Brinkkemper, 2010). Important characteristics of MM's are the maturity concept, the dimensions, the levels, the maturity principle, and the assessment approach, which are explained by the overview introduced by Lahrmann and Marx (2010) in Table 6. This thesis research most significant contribution is the BIaaS Capability Maturity Model (CMM) for BIaaS solution development. This chapter will introduce the BIaaS CMM.

Table 6: Properties of maturity models (Lahrmann & Marx, 2010)

Property	Description
Maturity concept	Three different maturity concepts (or understandings of maturity) can be distinguished (Mettler & Rohner, 2009). People (or workforce) capability defines "the level of knowledge, skills, and process abilities available for the performing an irganisation's business activities" (Curtis, Hefley, Miller, 2010). Process maturity defines "the extent to which a specific process is explicitly defined, managed, measured, controlled, and effective" (Paulk et al., 1993). Object (or technology) maturity defines the respective level of development of a design object (Gericke, Rohner, Winter, 2006)
Dimension	Dimensions are specific capability areas, process areas, or design objects structuring the field of interest. They should be exhaustive and distinct (de Bruin et al., 2005; Mettler & Rohner, 2009). Each dimension is further specified by measures (practices, objects, or activities) at each level (de Bruin et al., 2005; Fraser, Moultrie, Gregory, 2002).
Level	Levels are archetypal states of maturity of a certain dimension or domain. Each level has a distinguishing descriptor providing the level's intent and a detailed description (Lahrmann & Marx, 2010).
Maturity principle	MMs can be continuous or staged. Continuous MMs allow a scoring of activities at different levels. Therefore, the level can be either the (weighted) sum of the individual scores or the individual levels in different dimensions. Staged models require the compliance with all elements of one level (Fraser, Moultrie, Gregory, 2002). They specify a number of goals and key practices to reach a predefined level. Staged MMs reduce the levels to the defined stages, whereas continuous MMs open up the possibility of specifying situational levels.
Assessment	The assessment approach can be qualitative using descriptions or quantitative using e.g. Likert-like scales (Fraser, Moultrie, Gregory, 2002).

The following sub-chapters elaborate on the MM characteristics from Table 5 as reference for the development of the BIaaS CMM. The development process of the BIaaS CMM is setup using the CMM development framework proposed by de Bruin, Rosemann, Freeze and Kulkarni (2005). This chapter also elaborates on the product review research that is done to develop the maturity levels of the BIaaS CMM.

3.2 Overview of the BIaaS Capability Maturity Model

Although software engineers and managers often know the software development process in detail, they may disagree on which product capabilities are most important. Without an organized strategy for product development, it is difficult to achieve consensus on what capabilities to include and in which order. To develop the most profitable product, it is necessary to design an evolutionary path that increases the software product maturity in stages (Paulk et al., 1993).

The BIaaS capability maturity model is a framework representing a roadmap of recommendations for vendors developing BIaaS solutions. The BIaaS CMM can at least be used by the following type of users and activities:

- **BIaaS software developers** will use the BIaaS CMM as a roadmap to introduce new BIaaS solutions or improve existing .
- **Testing teams** will use the BIaaS CMM to assess developed BIaaS solutions and identify strengths and weaknesses of the solution. They also will use the BIaaS CMM to propose possible improvement points for the current solution. The SPM uses this information to make plans for direct (critical) or future (nice-to-haves) improvements.
- **Product managers of BIaaS solution vendors** will use the BIaaS CMM to understand the capabilities necessary to develop BIaaS solutions and produce development steps for new BIaaS solutions development process.
- **Researchers** will use the BIaaS CMM to understand and compare BIaaS capabilities for the use of future BIaaS research and to extend the proposed BIaaS CMM.

The above depict the variety of use and also scope the target audience of the BIaaS CMM, which is important for the next phases of the BIaaS CMM development.

3.3 Scoping the CMM

The first phase of the CMM development framework, suggesting a method to develop CMM's, is to determine the scope of the desired CMM (Bruin et al., 2005). The most significant decision that needs to be made in the scoping phase is the focus of the model. The focus refers to which domain the CMM would be target and applied, distinguishing the proposed model from other existing models.

The BIaaS CMM proposed in this thesis, focusses on the newly introduces BIaaS field. The BIaaS CMM tend to be used especially by BIaaS vendors using it to assess their own and competitor's BIaaS solutions, probably during the development phase of a newly introduced BIaaS solution or the improvement process of an existing BIaaS solution. The BIaaS CMM maturity concept is therefore technology oriented and depicts the respective maturity level of the BIaaS solution (Gericke, Rohner, Winter, 2006). The software product manager (SPM) of the BIaaS vendor is most likely to use the CMM, as he or she has the best overview of the development process (Ebert, 2007).

The BIaaS CMM is be used in three ways during the development process, each of which having different software product management purposes. The first usage is to assess the current existing (or just developed) BIaaS solution. Providing the SPM the maturity of its BIaaS solution, where from the market readiness can be derived. Secondly it is used to assess its own and competitor's BIaaS solutions, providing the SPM information about its own solution in contrast to its competitors. Product positioning can be derived from this information and can be used for marketing purposes. The third and maybe most important usage of the BIaaS CMM are the improvement steps that are provided by the model. The latter functionality of the BIaaS CMM provides the SPM an informed improvement approach of the assessed BIaaS solution. Each of these ways of using the BIaaS CMM can be an essential part of the development roadmap of the BIaaS solution.

3.4 Designing the CMM

The second phase of the CMM development framework is to determine a design or architecture for the model. The design phase provides *why* the model would be applied, *how* the model can be applied to varying solutions, *who* needs to apply the model and *what* can be achieved through applying the model (Bruin et al., 2005). This sub-chapter will provide the *why, how, who* and *what* of the designing phase, which is used later in this chapter with the actual development of the BIaaS CMM. Also the stage definitions are introduces that are used in the CMM.

This thesis introduce a new kind of BI and SaaS, namely BIaaS. Because this field is new, but vendors (like Avanade) see the potential of providing BIaaS solutions to their customers, they need an approach for the development of the new BIaaS solutions. The BIaaS CMM provides problem areas which aid SPM's at BIaaS vendors creating roadmap's for their BIaaS solutions. The BIaaS CMM also contributes scientifically to the BIaaS research field introducing the first comparison model for BIaaS capabilities, which aid researchers in further research on BIaaS.

The goal of this thesis is to introduce BIaaS for further scientific research on BIaaS, but also contribute on a practical basis by the introduction of BIaaS in the form of new BIaaS solutions. Both goals needs to extend the BI and SaaS scientific and development field respectively. The previous chapter introduced the BIaaS capability model which is a total merger of BI and SaaS key capabilities. The BIaaS CMM uses this capability model and therefore makes it easy to extend existing BI products with SaaS capabilities and vice versa. Current BIaaS solutions can be matured further by using the BIaaS CMM as roadmap for the development process. Using the CMM is therefore flexible and extends BI, SaaS and BIaaS by contributing both scientifically and practically.

Stage definitions needs to be developed to define the maturity levels of the CMM (Paulk et al., 1993). Defining maturity stages can be done either by using a top-down or bottom-up approach. With using the top-down approach, definitions are written first and then the measures are developed to fit the definitions. With the bottom-up approach the requirements and measures are determined first and then definitions are written to reflect these (Bruin et al., 2005). The previous chapter already introduces the capabilities and definitions, therefore the top-down approach is used. Looking at the maximum number of capabilities used per focus area in the

BIaaS capability model (hence maximum items per focus area in CMM), an eight-scale maturity level will best fit the CMM.

A maturity level is a well-defined evolutionary plateau toward achieving a mature software solution. Representation of maturity as a series of one-dimensional linear stages is widely-accepted and has formed the basis for assessment in many existing tools (Bruin et al., 2005). Each maturity level indicates a level of solution maturity. Historically labels are added to the levels of maturity used in CMM's (Paulk et al., 1993). However, IS literature on CMM refer to *process* maturity instead of *product* maturity, and software product maturity is often measured by code or module formula's that is not relevant for this current research. Therefore new maturity stages are introduced that refer to the maturity of the BIaaS solution taking market readiness into account. Table 7 provide the maturity stages that are used in the BIaaS CMM. Typically eleven levels of maturity are used in maturity modeling (zero up to maturity ten), however, the focus areas of the BIaaS Capability Model has an average population of five capabilities (smallest contain two and largest contain six). For the purpose of readability and usability of the BIaaS CMM, a nine scale maturity model is used (zero up to maturity eight).

Table 7: Maturity stages of the BIaaS CMM

Stages	Definitions
Level-0: Not implemented	The solution is not BIaaS, it does not contain all the vital capabilities for a BIaaS solution. The solution is incomplete.
Level-1: Initial	The initial capabilities of BIaaS are implemented. The current state is a foundation for every BIaaS solution, but need further improvement to be of any value.
Level-2: Basic	A basic level is obtained were the fundamental capabilities are in place. The solution is working as a basic BIaaS solution.
Level-3: Evolved	The solution is evolving and is becoming almost a complete BIaaS solution. More improvement is recommend before offering it for public use.
Level-4: Ready	The most common used capabilities currently available are implemented. The solution is ready to be offered as a BIaaS solution.
Level-5: Ready+	The most common used capabilities currently available are implemented.
Level-6: Maturing	The solution is maturing and includes additional competitive capabilities.
Level-7: Qualitative	The solution include all qualitative capabilities.
Level-8: Mature	The solution include all BIaaS capabilities currently available.

Maturity level zero indicates that the assessed solutions is incomplete and therefore is not a BIaaS solution. The solution has not all of the basic capabilities that a BIaaS solution must have. Maturity levels one up to three indicate a BIaaS solution that is becoming ready to be offered as a BIaaS solution but is not having all of the average capabilities current BI and SaaS solutions have. Maturity level four and five, *ready* and *ready+*, are seen as the average maturity of a BIaaS solution. These levels indicates that the assessed BIaaS solution is ready for the vendor to be offered as a complete BIaaS solution on the market, including the most popular BI and SaaS capabilities that are currently used in BI and SaaS solutions. Levels six, seven and eight are the evolving BIaaS maturity levels that indicate a more maturing BIaaS solutions that finally have the most advanced capabilities currently known on the market.

3.5 Capability comparison

The previous chapter introduced capabilities extracted from previous research obtained in the conducted literature review. These capabilities are used in the BIaaS capability model. This chapter will introduce the BIaaS CMM which is used to assess BIaaS solutions. The CMM uses maturity levels introduced in the preceding sub-chapter, which indicate the solution maturity by means of capability implementation. A requirement for using this method for measurement of solution maturity is that you can compare capabilities with each other.

When comparing entities with each other, some system should be available where comparison is possible. Common comparisons are <, > and =. However, to use such comparisons on capabilities, it must be clear which capability is more important than one other. A problem arises when you want to introduce a comparison for capabilities, because from literature it is not clear if one capability is more important than another one. So how do you find out? The answer is that there is not one particular way to state that one capability is more important than another capability, only if the first capability is necessary for the second capability to exist. When using logic representations; with current literature you can only state about two capabilities C_1 and C_2: $C_1 \rightarrow C_2$ (if C_1 then also C_2). This is not sufficient for the development of the BIaaS CMM, where it is necessary to introduce the comparisons > and <, so that can be stated $C_1 > C_2$ (C_1 is more important than C_2). Current literature does not give this information.

When taking the consumer market into account, market forces decide which capabilities are important for the consumer. This comes about the capabilities vendors put in their offered solutions, which implies that consumers want these capabilities in their purchased solutions. By using this market information, a comparison can be obtained using quantitative analysis. A new method for positioning capabilities is introduces, the capability maturity positioning method (i.e. CAMP). This method consist of a product review where the top most currently available BI and cloud computing solutions are reviewed and analyzed to develop a comparison technique for BIaaS capabilities. To increase the phenomenon of market forces, only the top most vendors of BI or cloud computing solutions, who develop solutions for commercial purposes, are taken into account for the product review. The following sub-chapter elaborates on CAMP introducing a comparison technique by using a product review.

3.6 Capability Maturity Positioning Method

In this sub-chapter literature and practice are put together by introducing the capability maturity positioning method (CAMP). The method contains of a product review and analysis, where current top-solutions are examined and tested on the availability of capabilities from the capability model (i.e. in this case the BIaaS capability model introduced in the previous chapter). For the product review twenty cloud computing and thirty-three BI solutions are examined. The product review is conducted to find the possibility to compare capabilities with each other. Comparing capabilities with each other is important for the development of the BIaaS capability maturity model.

The product review is conducted by examining cloud computing and BI solutions (by installing or use a trial account), consult product documentation and (where possible) interview solution experts. For every solution the capabilities that are implemented in or for the solution are checked and stored in a specially developed database which can be consulted for the quantitative analysis which is performed after the review.

3.6.1 BI Products

At this point a selection of BI solutions has to be made for the use in the product review. To increase the phenomenon of market forces, only the top most vendors of BI, who develop BI solutions for commercial purposes, are taken into account for the product review. To find these vendors it is useful to make advantage of the market research that is done by commercial researchers, like Gartner's research institute. Therefore, the selection of BI solutions chosen for the product review is made by taking the top companies designated by Gartner Research Institute in their Magic Quadrant of business intelligence platforms (Figure 8), which is published every year. This suggest the top companies producing BI solution platforms and are thereby the best choice to find the most evolved BI solutions currently on the commercial market.

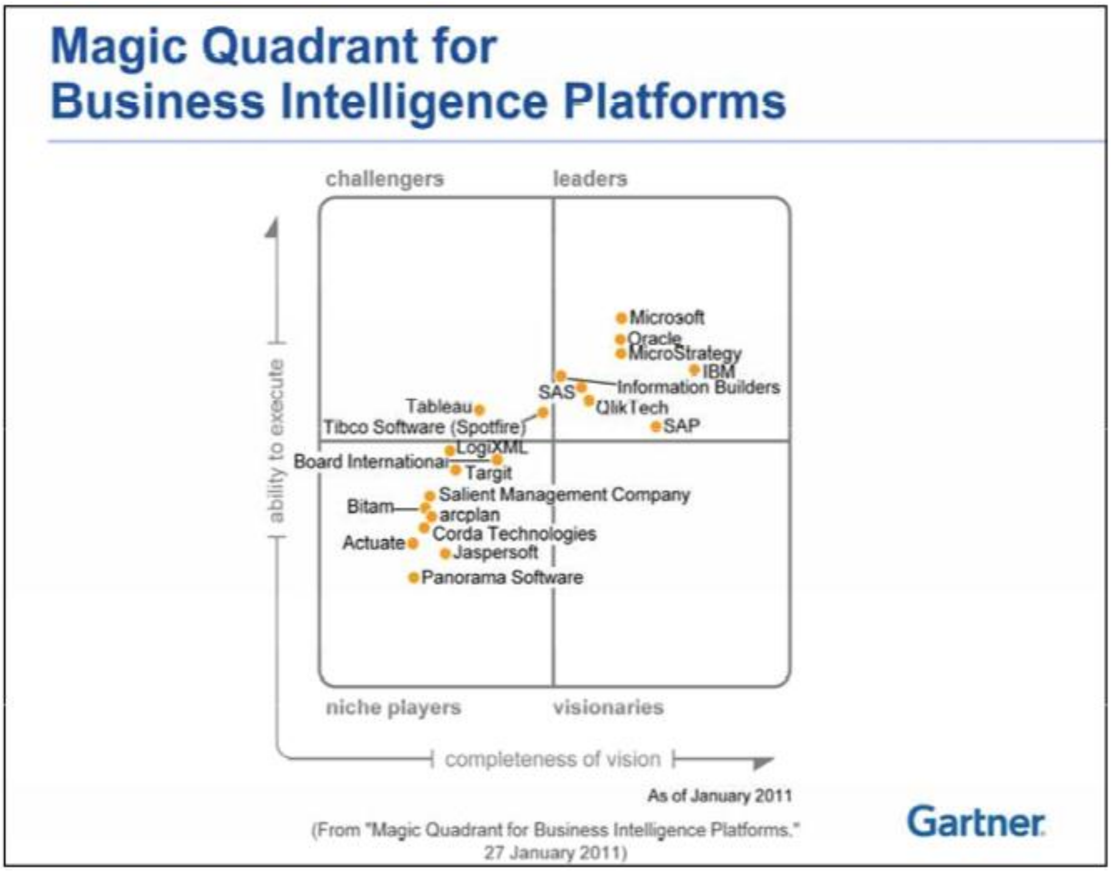


Figure 8 - Gartner's Magic Quadrant for BI platforms (Gartner, 2011b)

Using Gartner's Magic Quadrant for Business Intelligence Platforms in combination with the information that can be found on the BI solutions provided by these vendors, thirty-three BI solutions are reviewed from seventeen companies. The vendors and their solutions reviewed in this product review are summarized in Table 8. The solutions that are selected for the product review from each vendor are the newest and most complete BI solutions provided by those vendors.

The outcome of the product review per solution is summarized in Appendix 1. Because some companies (e.g. Microsoft) use a broader BI product portfolio where multiple products has to be used besides each other to function properly (e.g. SQL server 2008 to store the data, Report Builder or Excel to make reports and Sharepoint 2010 to publish the reports), also a summary is provided per vendor available in Appendix 2. This latter summary indicates which capabilities are implemented per vendor in at least one product (possible in multiple products).

Table 8 - BI vendors their solutions used in the product review

Vendor	Solution
Microsoft	SQL server 2008
	Sharepoint 2010
	Report Builder
	Excel (with PowerPivot)
IBM	Cognos
	SPSS
	Infosphere
Oracle	Business Intelligence Enterprice Edition 11g
	Database 11g
SAP	Business Objects
	Data Warehousing
	Netweaver
SAS	Enterprise BI server
	Visual BI
	Customer Intelligence
	Analytics
	Data Management
	Supply Chain Intelligence
Tableau Software	Tableau Server
Advizor Solutions Inc.	ADVIZOR Analyst
	ADVIZOR Analyst/X
TIBCO	Spotfire
	Data Quality
LogiXML	Logi Info
Arcplan	Arcplan Enterprice
MicroStrategy	MicroStrategy 9
BOARD	Management Intelligence
TARGIT	BI Suite
Predixion Software	Insight
Appistry	CloudIQ
Qliktech	Qlikview

3.6.2 *Quantitative analysis on BI solutions*

Every solution from Table 8 is carefully examined, the documentation is read and where possible also experts are questioned to find all the capabilities that are implemented in every individual solution. Appendix 1 and 2 provide the entire list of capabilities available in each of the examined thirty-three BI solutions.

While conducting the product review, a irregularity raised among the BI capabilities found per solution. For instance SAS scattered their BI capabilities over multiple solutions, each having their own specific abilities and a product can function as an addition to another product, resulting in each product to have different BI capabilities. However, TARGIT on the other hand offers their customers one BI solution including their whole variety of BI capabilities. The outcome of the analysis can be distort while conducting the quantitative analysis on all the solutions separately, knowing not all solutions are relatively the same. Therefore a different approach is used for the analysis of the BI solutions, namely a quantitative analysis taking the outcome per vendor (i.e. summary of capabilities available per vendor) into account. The analysis per vendor gives a better outcome of the analysis because the capabilities found per vendor are relatively the same, thereby suggest they are comparable. Appendix 2 contains the outcome of the product review per vendor, which is summarized in Table 9. Table 9 is the input for the quantitative analysis.

Table 9 - Summarized outcome of the product review per BI capability

Capability	#	Capability	#
Analytics	16	Forecasting	13
Automated exception detection	3	Information management	6
Automatic learning and refinement	1	Intelligent warehousing	4
Business process embeded	5	Master data management	8
Collaboration	14	OLAP	15
Data mining	14	BI alerting	7
Data modeling	6	Reporting	16
Data preprocessing	9	Secure data delivery	11
Data quality	11	Segmentation and clustering	7
Data transformation	7	Self-service BI	13
Data visualization	14	Data independency	4
Data warehousing	12	Text mining	9
Data follow-through workflow	2	Usability	12
ETL	13		

Recall from sub-chapter 3.4 that the product review is conducted to introduce a comparison between capabilities. The analysis on the sum of appearances introduced in Table 9 can give us such a comparison when we sort the capabilities on the total number of appearance in the reviewed solutions. However for the CMM it is equally important to know the levels of importance, referring to the maturity levels of the CMM. To find these levels, another analysis is performed using IBM's SPSS statistical analysis tool and the analysis methods proposed by Field (2009).

A frequency histogram is computed using the results from Table 9 as input (Figure 9). The calculations provided by SPSS provide *mean = 9,33* and *standard deviation = 4,498*. However, these calculation are made under the assumption of a normal distribution. Numerous techniques are available to show normality, like histograms (Figure 9), normal P-P plots (Appendix 3), normal Q-Q plots (Appendix 4) and boxplots. There are also methods to calculate and proof with a predefined probability that your results are normal. A method to proof the distribution is significantly (>95%) normal, is the Kolmogorov-Smirnov (K-S) test (Field, 2009).

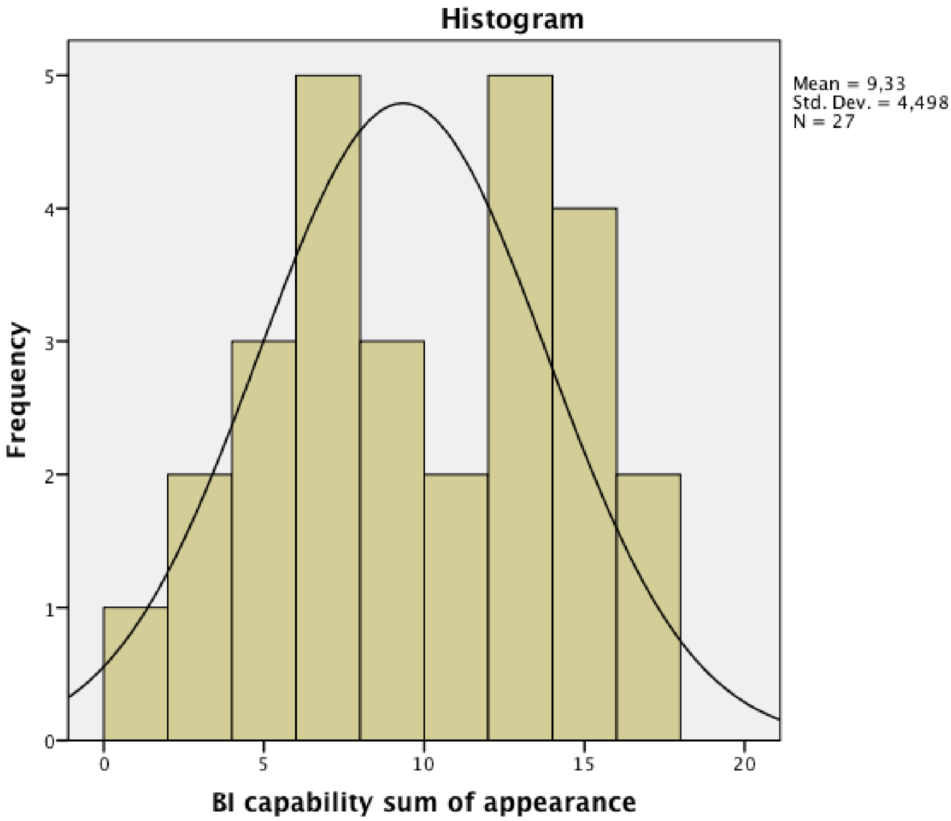


Figure 9 - Histogram for the sum of BI capabilities from product review

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
BI capability sum of appearance	,131	27	,200*	,947	27	,184

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure 10 - BI product review outcome test of normality

The SPSS analysis is used to compute a K-S test, the output is shown in Figure 10. The K-S test significance should be above ,05 to proof normality with a probability of 95% ($p < ,05$) (Field, 2009). The computed test calculated a K-S test significance of ,200, which is higher than the threshold of ,05. The results from Table 8 are therefore normal distributed with a probability of 95% ($p < ,05$).

The calculations of the mean and standard deviation from the analysis and the results from Table 8 can be used to group all BI key capabilities into levels of importance. The level of the BI key capability can calculated by taking the value of the result (x_i) from Table 8 and map it to the ranges provided in Table 9, where "μ" is the mean and "s" is the standard deviation. The levels from Table 9 are used for the calculation of the maturity levels in the BIaaS CMM by mapping all the results from Table 9 with the corresponding level from Table 10.

Table 10 - Mapping capability results to maturity levels

Value	Level
$x_i \geq (\mu+1,5s)$	Level-1
$(\mu+,5s) > x_i < (\mu+1,5s)$	Level-2 & level-3
$(\mu+,5s) \leq x_i \geq (\mu-1,5s)$	Level-4 & level-5
$(\mu-,5s) < x_i > (\mu-1,5s)$	Level-6 & level-7
$x_i \leq (\mu-1,5s)$	Level-8

Figure 11 illustrate the idea of normal distribution (curve) and the use of standard deviation (horizontal axis) to segment results (vertical lines and percentages). Due to this normally distribution and the use of a standard deviation, the most capabilities have a value near to the mean (indicated with green, yellow and orange) and only a small percentage are higher or lower than $\mu \pm 1,5*s$ (indicated with purple and blue respectively). Due to the latter, the mapping of the near-to-average values are mapped into two levels of maturity and the highest and lowest value into one.

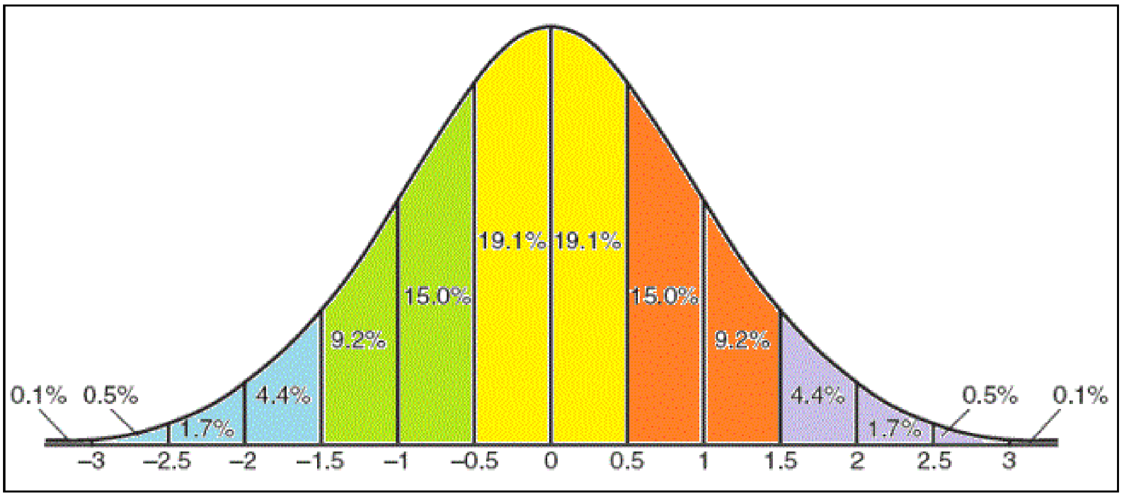


Figure 11 – Normal curve with standard deviation percentages

3.6.3 Cloud computing solutions

There are three pillars on top of which cloud computing solutions are delivered to end users. These are: software as a service (SaaS), platform as a service (PaaS), and infrastructure/hardware as a service (IaaS/HaaS) (Vecchiola, Chu, Buyya, 2009). For this research the SaaS part is of most interest, but it is not possible to provide SaaS without PaaS and IaaS. Therefore all aspects are taken into account making only looking at Gartner's Cloud Computing, PaaS, IaaS or SaaS Magic Quadrant not sufficient enough for the concept of BIaaS. Also the concepts of PaaS, IaaS and SaaS are different and therefore a lot of different companies and products are available.

Table 11: Cloud experts participated in expert brainstorm sessions

Expert name	Title	Experience in Cloud Computing (years)
Tijmen van de Kamp	Director at Avanade Netherlands	8
Kurt Claeys	Cloud strategy advisor at Microsoft Belgium	6

To make a selection of cloud computing solutions for this review a brainstorm session is setup with two cloud computing experts (Table 11). The companies and products from Gartner's four Magic Quadrants are taken into account during the brainstorm session and both experts are asked to provide those companies and solutions that are best fit for this BIaaS research. Finally the top twenty cloud computing solutions from both sessions are selected for the product review. Table 12 provide all cloud computing solutions that are reviewed in the product review.

The solutions are carefully examined (full or trial version), the documentation is read and where possible also experts are questioned to find all the capabilities that are implemented in every solution. Appendix 5 and Appendix 6 provide the outcome of the cloud computing review, which is summarized in Table 12.

Table 12 - Cloud computing vendors and their solutions used in the product review

Vendor	Solution
Microsoft	Azure
	Office 365
	CRM online
Google	Cloud connect
	Docs
	Sites
Amazon	Elastic compute cloud (Amazon EC2)
	CloudFront
	RDS
	S3
Salesforce.com	force.com
	Sales Cloud
Appistry	CloudIQ
AT&T	Hosting services
	Cloud computing services
	Application management services
Engine Yard	AppCloud
	xCloud
Enomaly	Elastic Compute Platform (ECP)
Flexiant	Flexiscale

3.6.4 *Quantitative analysis on Cloud computing solutions*

The same approach is used as for the BI product analysis to find a method for comparison between the capabilities. Table 13 provides the comparison information of total appearance of the capabilities in the cloud computing solutions. Again, this provides a comparison, but not directly the importance of each capability.

Table 13 - Summarized outcome of the product review per cloud computing capability

Capability	#	Capability	#
Data center	19	Multi-tenancy	10
Resource alocation	7	Hosted service	15
Internet centric	20	Data protection	13
Resource optimization	6	Virtualization	14
Vendor controlled	17	Scalability	18
Automatic adoption	4	Configurability	9
Monitoring	11	User friendliness	10
Time and location independent	20	QoS	15
Payment model	18	Competitive costs	2
SLA	10		

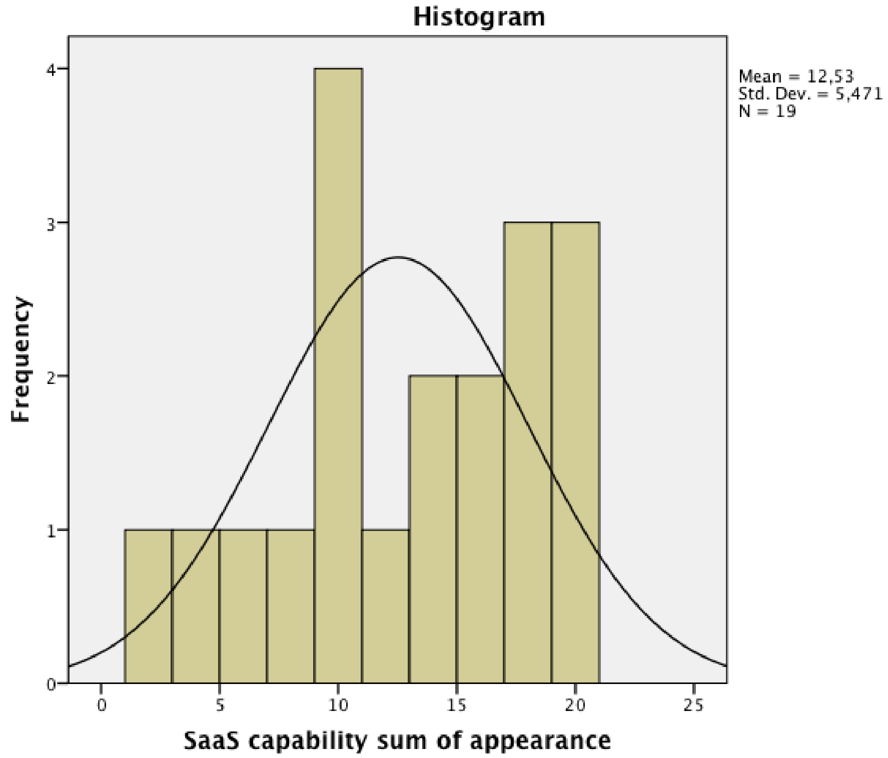


Figure 12 - Histogram for the sum of Cloud Computing capabilities from product review

The outcome from the cloud computing product review (Table 12) is analyzed using IBM's SPSS analysis tool. SPSS provides the frequency histogram and the calculation outcome for *mean = 12,53* and *standard deviation = 5,471* shown in Figure 12. These later calculations are made under the assumption of normal distribution. To show the results are significantly (95%) normal distributed, SPSS is used to calculate the Kolmogorov-Smirnov (K-S) test (Field, 2009).

Figure 13 provides the outcome of the K-S test, which is calculated by SPSS on a significance of ,200. An outcome of a K-S test significance above ,05 indicates a probability higher than 95% (p>,05) that the results are normal distributed (normal P-P plot and normal Q-Q plot provided in Appendix 7 and Appendix 8). Therefore, with a probability of a least 95%, can be stated that the results from the Cloud computing product review are normal distributed.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
SaaS capability sum of appearance	,109	19	,200*	,952	19	,430

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure 13 – Cloud Computing product review outcome test of normality

The calculation output of SPSS providing the mean and standard deviation can be used to group the capabilities into levels. The level for each SaaS key capability can be calculated by taking the result from Table 13 (x_i) and map the result to the calculation Table 10, where "µ" is the mean and "s" is the standard deviation. The levels provided by this mapping are the levels of maturity used in the BIaaS CMM.

3.7 Contents of the BIaaS CMM

Sub-chapter 3.5 provide a mapping method to calculate the levels of maturity for every BIaaS capability from the BIaaS capability model. Using the mapping from Table 9, the results from the product review (Table 9 and Table 13) and the BIaaS capability model, a BIaaS capability maturity matrix is developed which is shown in Table 13. The BIaaS capability maturity matrix consist of BIaaS focus areas, each with their own number of maturity levels. The focus areas are represented in the left-most column and the specific maturity levels are represented by the characters A-F in a maturity range from 1 to 8 (see top-most row). An extended version of the maturity matrix with all capabilities is presented in Appendix 9. The levels of maturity per focus area indicate a best practice order in which the capabilities are implemented from left to right.

The BIaaS capability maturity matrix is the foundation of the BIaaS capability maturity model (CMM). The BIaaS CMM is a physical assessment tool which uses the maturity matrix in a working model using Microsoft Excel® (deBoer, 2012). The model uses an assessment sheet as input in the form of yes/no answers to questions regarding specific capability functionality. The input is calculated by the calculation sheet and provides output in the BIaaS CMM sheet. The BIaaS CMM sheet is the calculated feedback for the assessment providing overall maturity, problem areas and ordered maturity per focus area. The following sub-chapters elaborated on the three available sheets in the BIaaS CMM.

Table 14 - The BIaaS capability maturity matrix

	0	1	2	3	4	5	6	7	8
Design									
Data Storage		A			B		C		
Accessibility		A	B		C				D
Data									
Data gathering		A						B	C
Data management		A	B	C	D		E		F
Data processing					A		B	C	D
Data analysis		A	B	C	D	E	F		
Usage									
Application logic		A	B		C				D
Usability					A	B	C		
Consumerization		A	B	C	D	E	F		
Support									
Alerting							A		B
Service		A	B	C	D				E

3.7.1 Assessment sheet

The BIaaS CMM is used to assess existing BI, SaaS or BIaaS solutions to provide a strategy for an evolutionary path to BIaaS solution maturity (development roadmap). This is done by providing the current state of solution maturity but also provide specific problem areas for product improvement. The assessment starts by providing yes/no answers to questions provided in the first sheet of the BIaaS CMM, the *assessment sheet*.

The assessment is preferably carried out by the software product manager or the project leader for the development team, because they can answer the question best about the different focus areas and implemented capabilities. All the questions are answered with *yes* or *no* by selecting the correct answer from the drop-down menu to the right of each question. The answer column (the right-most column) will change color from orange to white when an answer is provided. When all questions are answered, hence all boxes in the answer column are white, the assessment is done and the results can be analyzed in the CMM sheet.

3.7.2 Calculation Sheet

The calculation sheet is a hidden sheet where all calculations for the BIaaS CMM outcome is performed. The calculation sheet consist of three types, the *problem areas, CMM ordered* and *overall maturity*.

The *problem area* part is sub-categorized in three categories each represent all the focus areas in the first column with their associated capabilities in the corresponding row. The capabilities in each row are ordered to their importance (level-1 to level-8) from left to right.
The *first* (left-most) category of the problem areas checks if a capability is implemented. The outcome is a boolean value True or False, where True means the capability is implemented and False otherwise (not implemented or not answered).
The *second* category checks wether the questions that correspond to the capability is answered (regardless of the answer) or not and returns True or False respectively.
The *third* and last category of the problem area is the merger of the other two calculating for each capability if it is *implemented* (answer to question is "yes"), *missing* (answer is "no") or *unknown* (not answered yet).

The *CMM ordered* part of the calculation sheet contain the BIaaS focus areas and corresponding capabilities as shown in the maturity matrix and are also divided into three sub-categories.
The *first* table checks if the capabilities are implemented, making use of the first table from the problem area part. If the capability is implemented the value is set to zero (0), otherwise to one (1). The sum of each column is provided on the top of the column of each solution unit. The empty field from the maturity matrix are default set to zero (no value).
The *second* table does not contain the focus areas, only the solution units and calculates the maturity per unit. The maturity per unit is calculated first looking at the levels individually (starting at level one, the column with the "1" in the header). For each level of the unit check the corresponding sum from the first table. If the sum is not zero (hence, there are unimplemented capabilities at this level), the maturity of that level is equal to the previous level (note that the left-most level is zero to start with). However, if the sum is zero (all capabilities of this level are implemented) than the maturity at this level becomes equal to the corresponding maturity level (equal to the header of the level), but only if the previous level is also equal to the previous header (the previous capabilities are also implemented). With other words, the maturity of a solution unit at particular level is calculated by checking if all capabilities at that level are implemented and also all previous capabilities from the lower levels are implemented. There is a special case in this calculation, when there are no capabilities at a particular level in a solution unit. In that case, the previous maturity is taken and is only updated when the next level is fully

implemented or the units at that same level are all fully implemented.
The *third* table is the merger of the third from the problem area part and the characters from the maturity matrix from Table 14. This table provides for each field of the maturity matrix if it is implemented, missing or unknown with the corresponding capability character as shown in Chapter 2.

The *third* part contains the BIaaS maturity calculation of the assessed solution. It contains the maximum maturity level for each solution unit calculated on each row in the second table of the *ordered CMM.* The overall solution maturity is calculated by taking the minimum of the maturity levels of each solution unit.

3.7.3 CMM sheet

The last sheet in the BIaaS CMM provides all the information about the assessment in one overview. The *Problem area* part provide information about which capabilities from each focus area are implemented. The *Ordered CMM* provides the implemented and missing capabilities per maturity level, which can be used as roadmap for maturity improvement. The *Solution maturity* provides the solution maturity per unit and the overall maturity.

3.8 BIaaS assessment using the BIaaS CMM

BIaaS assessment using the BIaaS CMM requires explicit knowledge about the technological design, data management, the usage and the supporting model of the solution. Therefore the BIaaS assessment is preferably performed by a BIaaS vendor's software product manager or the project leader for the development team.
The assessment starts with the assessment sheet which contains forty-seven questions. All questions need to be answered with "yes" or "no" using the drop-down functionality available at the end of each question. When all questions are answered (there are no orange fields left in the answer column), the outcome is provided in the CMM sheet. This sheet provides the overall *maturity* level of the BIaaS solution, which is composed from the maturity levels of the solution units which are also provided. The *problem areas* pinpoints those capabilities that are missing from the assessed solution, and thus need attention from the developer of the solution. Additional information about the missing capabilities can be found in Chapter 2 which also contains references to detailed scientific research for each capability. The *CMM ordered* part of the CMM sheet can be used as a roadmap for solution improvement by providing an implementation order for each capability. The improvement path starts at the left-most maturity column and by implement each capability from that column working to the right level to level.

Using the BIaaS CMM as an assessment tool which provides a roadmap for BIaaS solution improvement answers the final research question "*How can BIaaS capabilities be used to create a product portfolio roadmap for BIaaS solution vendors?*". The BIaaS CMM is developed using the capability model developed in Chapter 2 which contain all BIaaS key capabilities, and the BIaaS CMM can be used to create a roadmap for solution improvement to be used by BIaaS solution vendors.

4 DISCUSSION AND CONCLUSION

4.1 Discussion

This thesis research major contributions and deliverables are the BIaaS capability model, which conceptually model BIaaS, and the BIaaS capability maturity model, which introduces a assessment model for BIaaS solution and aim to be used as a roadmap for BIaaS solution vendors. There are however some limitations to the introduced models.

The BIaaS capability model is constructed using a literature review and extract the key capabilities from existing literature. Although this method, where literature is used to position concepts, is often used in scientific research, it has its limitations. Particular the information technology branch has fast changing capabilities and introduces newly development technologies relatively often. Therefore the constructed model should be updated probably at least ones in the five to ten years, to exclude absolute capabilities and introduce new key capabilities (if available).

The BIaaS CMM is developed using the introduction of the capability maturity positioning method (CAMP) and is dependent on the products currently available on the market. CAMP positions capabilities from a capability model by examine the current top most solutions on the market. Therefore the BIaaS capability maturity model is dependent on the solutions available on the market at the time of the performed CAMP (i.e. snapshot). The outcome of an assessment using the BIaaS CMM is therefore always bound to a particular period in time. To keep the BIaaS CMM up-to-date, preferably the CAMP method should be performed once in the two years, as this is the average cycle for major IT vendors for releasing new versions of their solutions.

Different methods are used to develop the BIaaS CMM. First the BIaaS capability model is introduced by using a structured literature review introduced by Webster and Watson (2002) and expert input. The BIaaS capability model is used for the development of the BIaaS capability maturity model, which is constructed by the CAMP method including a product review and statistical analysis. Although strictly following the methods used for the construction of the model can conclude valid research is performed, the model is not been evaluated in practice yet. To strengthen the model, an evaluation should be performed to correct possible unforeseen flaws in the model or the development process.

4.2 Conclusion

This thesis research first aim is to conceptual model the integration of two research domains, business intelligence and cloud computing, into the new concept *business intelligence as a service* (BIaaS), to extend research on BIaaS and to aid BIaaS solution vendors in the development (or improvement) of new BIaaS solutions. This aim is been met by answering the first research question proposed in the first chapter:

> "*What are business intelligence as a service capabilities and how do they differ from conventional business intelligence capabilities?*"

As mentioned earlier in this thesis document, BI and cloud computing are two research domains that exist for some time. This thesis research extracted the main capabilities from BI and cloud computing research by conducting a literature review. The literature review spread light on the available capabilities in each domain, and an analysis of the review extracted key BI and cloud computing capabilities. Strongly coherent groups of key capabilities are formed using literature and expert group sessions which form the BIaaS focus areas. The key capabilities, grouped into focus areas, result into the business intelligence as a service capability model. The BIaaS capability model, together with the explanations of the capabilities, conceptually model the new concept of business intelligence as a service and answering the first research question.

For a BIaaS solution vendor to successful develop BIaaS solutions, there are a number of aspects which needs to be considered, including which capabilities should be implemented and in which order. The second aim of this thesis research is to use business intelligence as a service capabilities for the creation of a roadmap for BIaaS solution development. This aim is met by answering the second research question:

> "*How can business intelligence as a service capabilities be used to create a product portfolio roadmap for business intelligence as a service solution vendors?*"

This thesis research introduces the capability maturity positioning (CAMP) method, which calculate the position of capabilities in the current commercial market. Analysis provide an ordered value for each capability in their focus area, resulting in a BIaaS maturity matrix. The BIaaS maturity matrix is input for the developed BIaaS capability maturity model which can be used as an assessment tool for BIaaS solutions. The second research question is answered by the development of the BIaaS capability maturity model (CMM), which is constructed and elaborated in detail in the third chapter of this thesis document. The BIaaS CMM and the knowledge provided by the BIaaS capability model and additional background knowledge, will assist BIaaS solution vendors in the development of BIaaS solutions. An assessment using the BIaaS CMM provide an ordered development process, depict problem areas for solution improvement and calculate a level of maturity for solution positioning. The BIaaS CMM aid BIaaS solution vendors by creating a roadmap for their development process.

4.3 Future research

The biggest limitation of this research is a missing evaluation of the BIaaS CMM. An interesting follow-up study could therefore evaluate the model by case studies where BIaaS (or maybe BI) solutions are assessed. The use of experts can be of importance for the case studies, whereas the experts can compare the assessment output with their own expected outcome. Analysis of the similarities and differences can provide possible improvements for the model or the methods used to develop the model.

Another recommendation for further research is to use the capability maturity positioning (CAMP) method on available open source products. The comparison can provide a better understanding between open source and commercial product. Moreover the comparison of the resulting maturity matrix can be compared with the suggested maturity matrix in this thesis research, where analysis can perhaps strengthen the current model or provide suggestions for improvement.

REFERENCE

Abdat, N., Spruit, M., Bos, M. (2011). Software as a Service and the Pricing Strategy for Vendors. Business science reference, pp. 154-192. IGI Global, Hershey, New York.

Agarwal, R., Tanniru, M. R. (1990). Knowledge acquisition using structured interviewing: an empirical investigation. *Journal of Management Information Systems*, 7(1), 123–140.

Anandarajan, M., Anandarajan, A., Srinivasan, C.A. (2004). *Business Intelligence Techniques - A Perspective from Accounting and Finance*. Springer, Berlin 2004.

Armbrust, M., Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M. (2009). *Above the Clouds: A Berkeley View of Cloud computing*. Technical Report No. UCB/EECS-2009-28, University of California at Berkley, USA, Feb. 10, 2009.

Avanade (2011). Avanade's company website, available at http://www.avanade.com.

Baars, H., Kemper, H.G. (2008). *Management Support with Structured and Unstructured Data – An Integrated Business Intelligence Framework*. Information Systems Management, vol. 25, no. 2, pp. 132-148, 2008.

Becker, J., Knackstedt, R., Pöppelbuß, J. (2009). *Developing Maturity Models for IT Management - A Procedure Model and its Application*. Business & Information Systems Engineering, vol. 1, no. 3, pp. 213-222.

Bégin, M. (2008). *An EGEE Comparative Study: Grids and Clouds - Evolution or Revolution?*. 23rd Open Grid Forum (OGF23), June 2008.

Bekkers, W., Weerd, I. van de, Spruit, M., & Brinkkemper, S. (2010). A framework for process improvement in software product management. *Proceedings of the 17th European Conference on Systems, Software and Services Process Improvement (EUROSPI),* Grenoble, France, 1-12.

Benlian, A., Hess T. and Buxmann, P. (2009). Drivers of SaaS-Adoption- An Empirical Study of Different Application Types. Business & Information Systems Engineering, 5: pp. 357-368, 2009.

Beyer, K.S., Ramakrishnan, R. (1999). Bottom-up computation of sparse and iceberg cubes. *In: SIGMOD Conference 1999*, pp 359–370.

Bezemer, C.P., Zaidman, A. (2010). Multi-tenant saas applications: Maintenance dream or nightmare?. *Proceedings of the 4th International Joint ERCIM/IWPSE Symposium on Software Evolution (IWPSE-EVOL)*. ACM, 2010, p. To appear.

Blokdijk, G. (2008). *SaaS 100 succes secrets - How companies successfully buy, manage, host and deliver software as a service (SaaS)*. Brisbane, Australia: Emereo Pty Ltd.

Blumberg, R. and Atre, S. (2003a). The Problem with Unstructured Data. DM Review 2003, September 12, 2003.

Blumberg, R. and Atre, S. (2003b). Automatic Classification: Moving to the Mainstream. DM Review, http://dmreview.com/master.cfm? NavID=55 &EdID=6501.

Boncella, R. J. (2000). Web Security For E-Commerce. Communications of the Association for Information Systems, Volume (4)11, November 2000.

Boncella, R. J. (2002). Wireless Security: An Overview. Communications of the Association for Information Systems, Volume (9)14, October 2002.

Boncella, R.J. (2003). Competitive intelligence and the web. Communications of the Association for Information Systems, Volume (12) 21, September 2003.

Bragg, R (2009). Cloud computing: When computers really rule. Tech News World, July 2008. Electronic Magazine.

Buyya, R., Pandey, S., Vecchiola, C. (2009). Cloudbus toolkit for market-oriented cloud computing. In CloudCom '09: Proceedings of the 1st International Conference on Cloud Computing, volume 5931 of LNCS, pages 24–44. Springer, Germany, December 2009.

Buyya, R., Yeo, C.S., Venugopal, S. (2008). Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. CoRR, (abs/0808.3558), 2008.

Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J., Brandic, I. (2009). Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility. Future Generation Computer Systems, 25(6): 599-616, Elsevier Science, Amsterdam, The Netherlands, June 2009.

Byrd, T. A., Cossick, K. L., & Zmud, R. W. (1992). *A synthesis of research on requirements analysis and knowledge acquisition techniques*. MIS Quarterly, 16(1), 117–138.

Cates, J.E., Gill, S.S., Zeituny, N. (2005). *The Ladder of Business Intelligence (LOBI): a framework for enterprise IT planning and architecture*. International Journal of Business Information Systems, vol. 1, no. 1-2, pp. 220-238.

Chandler, N (2011). Fast Forward to 2014 — Gartner Predicts the Future of Business Intelligence and Data Analytics.TIBCO Spotfire's business intelligence blog, available at http://spotfireblog.tibco.com/?p=4779#more-4779.

Chaudhuri, S., Dayal, U. (1997). *An Overview of Data Warehousing and OLAP Technology*. ACM SIGMOD Record 26(1), 1997.

Chen M.S., Han, J., Yu, P.S. (1996). Data mining: An overview from a database perspective. IEE Transactions on Knowledge and Data Engineering, 8(6):866883, 1996.

Chen, C., Yan, X., Zhu, F., Han, J., Yu, P.S. (2009). *Graph OLAP: Towards online analytical processing on graphs*. Knowledge Information Systems 21, pp. 41–63, 2009.

Chou, D.C. & Chou, A.Y. (2008). Software as a Service (SaaS) as an Outsourcing Model: An Economic Analysis. *Proceedings of the 39th Southwest Decision Science Institute Conference*, Houston, Texas, 2008.

Clemmons, E.K. , Reddi, S.P. and Row, M.C. (1991). The Impact of Information Technology on the Organization of Economic Activity: The 'Move to the Middle' Hypothesis. Journal of Management Information Systems, 10(2): pp. 9-35, 1991.

Cody, W.F., Kreulen, J.T., Krishna, V., Spangler, W.S. (2002). *The Integration of Business Intelligence and Knowledge Management*. IBM Systems Journal, (41)4, pp. 697-713.

Corley, C.D., Cook, D.J., Mikler, A.R., Singh, K.P. (2010). Text and structural data mining of influenza mentions in web and social media. *Public Health Informatics special issue in International Journal of Environmental Research and Public Health*, vol. 7, 2010.

Curtis, B., Hefley, W.E., Miller, S.A. (2010). *The People Capability Maturity Model – Guidelines for Improving the Workforce*, 2 ed.. SEI Series in Software Engineering, Boston, MA: Addison-Wesley.

Cybenko, G., Giani, A., Thompson, P.. (2002). Cognitive Hacking: A Battle for the Mind. IEEE Computer (35)8, August 2002, pp. 50–56.

deBoer, T. (2012). The business intelligence as a service capability maturity model. Master thesis research deliverable for graduation at Utrecht University. Source: http://www.timdeboer.eu/thesis_mbi/biaas_cmm_v20120826.xlsx

deBoer, T., Lossek, M., Janssen, R., Neppelenbroek, M. (2011). Twitter: An architectural review. Scholars paper on Software Architecture at Utrecht University. Source: http://www.timdeboer.eu/paper_publishing/Twitter_An_Architectural_Review.pdf

deBruin, T., Rosemann, M., Freeze, R., Kulkarni, U. (2005). *Understanding the Main Phases of Developing a Maturity Assessment Model*. In Proceedings of ACIS 2005.

De Marco, M., Rossignoli, C., Ferrari, A., Mola, L., Zardini, A. (2010). BI as a service: an attempt to understand the leading adoption factors. The 2010 International Conference on E-Business Intelligence (pp. 120-127). Atlantis Press, 2010.

Dennis, A., Wixom, B., Tegarden, D. (2002). *Systems Analysis and Design: An object-oriented approach with UML*. New York: John Wiley & Sons Inc.

Diallo, B.A.A., Badardt, T., Hubert, F., Daniel, S. (2011). *Towards context-awareness mobile geospatial BI (GEOBI) applications*. University Laval, Quebec, Canada. Retrieved from icaci.org.

Dym, R. (2009). *Why Software as a service? Helping our customers reduce costs and increase revenue*. OpSource - The Business of Web Operations database.

Ebert, C. (2007). The impacts of software product management. *Journal of Systems and Software 80*(6), 850 - 861.

Elena, C. (2011). *Business intelligence*. Journal of Knowledge Management, Economics and Information Technology, ScientificPapers.org, vol. 1(2), pages 10, February 2011.

Field, A.P. (2009). *Discovering statistics with SPSS* (3rd ed.). London: Sage.

Fisher, R. (2011). *Edinburgh Online Graphics Dictionary*. University of Edinburgh, December 2, 2011, http://homepages.inf.ed.ac.uk/rbf/GRDICT/grdict.htm.

Fraser, P., Moultrie, J., Gregory, M. (2002). *The Use of Maturity Models/Grids as a Tool in Assessing Product Development Capability*. In Proceedings of IEMC 2002, Cambridge, UK, pp. 244-249.

Frawley, W.J., Piatetsky-Shapiro G., Matheus, C.J. (1992). *Knowledge discovery in databases: an overview*. AAAI MIT Press, IA Magazine Vol. 13, Nr. 3, pp. 57-70.

Friedman, T. (2002). *How Secure Is Your Business Intelligence Environment*. Gartner Research Inc., August 2002, pp. 1-6.

Gartner (2011a). Gartner's company website, available at http://www.gartner.com.

Gartner (2011b). Slide deck from Gartner's International Business Intelligence summit 2011 in London.

Geelan, J. (2008). Twenty one experts define cloud computing. Virtualization, August 2008. Electronic Magazine.

Gericke, A., Rohner, P., Winter, R. (2006). *Networkability in the Health Care Sector - Necessity, Measurement and Systematic Development as the Prerequisites for Increasing the Operational Efficiency of Administrative Processes*. In Proceedings of ACIS 2006, Adelaide, Australia.

Gibson, C.F., Nolan, R.L. (1974). *Managing the four stages of EDP growth*. Harvard Business Review, vol. 52, no. 1, pp. 76-88.

Golfarelli, M., Rizzi, S., Cella, I. (2004). Beyond data warehousing: What's next in business intelligence? In Proc. DOLAP, pages 1–6, 2004.

Gonzales, M.L. (2003). IBM Data Warehousing with IBM Business Intelligence Tools. Wiley, 2003.

Gray, J., Chaudhuri, S., Bosworth, A., Layman, A., Reichart, D.,Venkatrao, M., Pellow, F., Pirahesh, H. (1997). Data cube: a relational aggregation operator generalizing group-by, cross-tab, and sub totals. Data Mining Knowledge Discovery 1(1):29–53.

Greschler, D., Mangan, T. (2002). Networking lessons in delivering 'software as a service' - Part II. *International Journal of Networdk Management*, 12(5), 317-321. doi:10.1002/nem.446

Gruman, G. & Knorr, E. (2008). What cloud computing really means. InfoWorld, April 2008. Electronic Magazine.

Guo, C. J., Sun, W., Huang, Y., Wang, Z. H., Gao, B. (2007). A Framework for Native Multi-Tenancy Application Development and Management. *In Proceedings of CEC/EEE 2007.*

Han, J., Cai, Y., Cercone, N. (1992). Knowledge discovery in databases: an attribute-oriented approach. In *Proceedings of the 18th International Conference on Very Lange Data Bases*. pages 547-559, Vancouver, British Columbia, Canada, 23-27, August 1992.

Hand, D.J., Mannila, H., and Smyth, P. (2001). Principles of Data Mining. MIT Press.

Helfat, C.E. and Peteraf, M.A. (2003). The dynamic resource-based view: capability lifecycles. Strategic Management Journal, 24 (10), 997 – 1010.

Hoch, F., Kerr, M., Griffith, A. (2001). *Software as a service: strategic backgrounder*. Retrieved November 25, 2011, from http://www.siia.net/estore/ssb-01.pdf.

Holtzblatt, K., Beyer, H. R. (1995). Requirements gathering: the human factor. *Communications of ACM*, 38(5), 30–32.

Hulme, G. W. (2003). *Hack in Progress*. Information Week, http://www.informationweek.com/news/14400070 (current September 8, 2003).

Hwang, K. (2008). Keynote. massively distributed systems: From grids and p2p to clouds. In The 3rd International Conference on Grid and Pervasive Computing-gpc-workshops, pp. 22. 2008.

Inmon, W.H. (1992). Building the Data Warehouse. QED Information Sciences, Wellesley, MA, 1992.

Inmon, W.H., Strauss, D., Neushloss, G. (2008). DW 2.0: The Architecture for the Next Generation of Data Warehousing. Amsterdam: Elsevier Science.

Jain, A. K., Murty, M. N., Flynn, P. J. (1999). *Data clustering: A review*. ACM Computing Surveys, Vol. 31, 264–323.

Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., Becker, B. (2008). The Data Warehouse Lifecycle Toolkit, 2 ed.. New York, John Wiley & Sons.

Kakihara, M., Sorenson, C. (2002). *Mobility: an extended perspective*. Proceedings of the 35th Annual Hawaii International Conference on System Sciences, pp 1756-1766. Hawaii.

Krasnow, J. D. (2000). The Competitive Intelligence and National Security Threat from Website Job Listings. http://csrc.nist.gov/nissc/2000/proceedings/papers/600.pdf. (current September 1, 2003).

Kvale, S. (1996). InterViews: an introduction to qualitative research interviewing. 1996. Sage Publications: Thou-sand Oaks, CA.

Lablong, M.E. (2011). *Data Governance POV with an overview of Data Quality and MDM*. Accenture research on Data Governance, version 1.0, 17 Dec. 2011. Retrieved 20 Dec. 2011 from Accenture's internal knowledge exchange base.

Lahrmann, G., Marx, F. (2010). *Systematization of Maturity Model Extensions.* In Proceedings of DESRIST 2010, St. Gallen, Springer, pp. 522-525.

Langseth, J. and Vivatrat, N. (2003). Why Proactive Business Intelligence is a Hallmark of the Real-Time Enterprise: Outward Bound. Intelligent Enterprise, (5)18, pp. 34-41.

Mahowald, R. P. (2009). SaaS, PaaS, and cloud: choices for success. Proceedings of the IDC SaaS Summit Spring 2009, Document # 217935. Retrieved April 5, 2009, from IDC database.

Paul McFedries (2008). The cloud is the computer. *IEEE Spectrum Online*, August 2008. Electronic Magazine, available at http://spectrum.ieee.org/computing/hardware/the-cloud-is-the-computer.

McKnight, W. (2003). Bringing data mining to the front line, part 2. DM Review, 13(1), pp. 50, 2003.

Mettler, T., Rohner, P. (2009). Situational Maturity Models as Instrumental Artifacts for Organizational Design. *In Proceedings of DESRIST 2009*, New York.

Mietzner, R., Leymann, F. (2008). Towards Provisioning the Cloud: On the Usage of Multi-Granularity Flows and Services to Realize a Unified Provisioning Infrastructure for SaaS Applications. *In Proceedings of IEEE Congress on Services - Part I*, pages 3–10, Los Alamitos, CA, USA, 2008. IEEE Computer Society.

Mietzner, R., Metzger, A., Leymann, F., Pohl, K. (2009). Variability modeling to support customization and deployment of multi-tenant-aware software as a service applications. *Proceedings of the ICSE Workshop on Principles of Eng. Service Oriented Systems (PESOS)*, pages 18–25. IEEE, 2009.

Milojicic, D. (2008). Cloud computing: Interview with russ daniels and franco travostino. IEEE Internet Computing, (5):7–9, Sept/Oct 2008.

Nasukawa T. and Nagano T. (2001). Text Analysis and Knowledge Mining System. *IBM Systems Journal 40*, No. 4, 967–984, 2001.

Negash, S. (2004). Business Intelligence. Communications of the Association for Information Systems. Volume. 13, pp. 177-195.

Pang, B., Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, vol. 2, nos. 1–2, pp. 1–135, 2008.

Paulk, M., Curtis, B., Chrissis, M., Weber, C. (1993). *Capability Maturity Model for Software, Version 1.1*. CMU-SEI-93-TR-24, Software Engineering Institute, Pittsburgh PA 15213.

Porter, M.E. (1998). Clusters and the new economics of competition. Harvard Business Review November-December 1998: pp. 77-90, 1998.

Porter, M.E. & Millar, V.E. (1980). How information gives you competitive advantage. Harvard Business Review, 63(4): pp. 149-160, 1980.

Power, D. (2004). A brief history of decision support systems. DSSResources.COM, World Wide Web, http://DSSResources.COM/history/dsshistory.html, version 4.0, March 10, 2007.

Richardson, J. (2011). The Consumerization of BI: Emerging Technologies to Watch in 2011. Gartner webinar on BI consumerization from http://gartner.com/webinar.

Roberts, J. (2007). *State of the Art: Coordinated and Multiple Views in Exploratory Visualization*. Coordinated and Multiple Views in Exploratory Visualization, ETH, Switzerland, IEEE Press, 61-71.

Robins, D.B. (2010). Complex Event Processing. In *CSEP 504,* 2010.

Rust, R.T., Kannan, P.K. (2003). E-service: A new paradigm for the business in the electronic environment. Communications of the ACM, 46(6), 36-42. doi:10.1145/777313.777336.

Sääksjärvi, M., Lassila, A., Nordstrom, H. (2005). Evaluating the software as a service business model: From CPU tie-sharing to online innovation sharing. Proceedings of the IADIS International Conference e-Society 2005, Qawra, Malta.

Shi, Z.Z., Huang, Y.P., He, Q., Xu, L.D., Liu, S.H., Qin, L.X. (2004). MSMiner - a developing platform for OLAP. Decision Support Systems, 2004.

Shobrys, D. (2003). Business intelligence and supply chain management. Peer Publishing, 2003. http://supplychain.ittoolbox.com/documents/peer -publishing/business-intelligence-and-supplychain- management-1948.

Simpson, J.A., Weiner, E.S.C. (2011). The online Oxford English Dictionary. Oxford, UK: Oxford University.

Smith, J.A. (1995). Semi-structured interviewing and qualitative analysis. In: J. A. Smith, R. Harré and L. Van Langenhove (eds), *Rethinking Methods in Psychology*. London: Sage

Soejarto, A. (2003). Tough Times Call for Business Intelligence Services, an Indisputable Area of Growth. http://vb.channelsupersearch.com/news/var/40682.asp or www.VarBusiness.com (May 2003).

Swanson, D.R. (1987). Two medical literatures that are logically but not bibliographically connected. Journal of the American Society for Information Retrieval, 38(4), pp. 228–233.

Theus, M. (2002). Interactive data visualization using mondrian. *Journal of Statistical Software*, 7(11), 2002.

Thompson, W.J.J. & Van der Walt, J.S. (2010). Business intelligence in the cloud. SA Journal of information Management 12(1), Art. #445, 5 pages. DOI: 10.4102/sajim.v12i1.445.

Thomsen, C., Pedersen, T.B. (2008). *A Survey of Open Source Tools for Busienss Intelligence*. A DB Technical Report, TR-23, September 2008, pp. 1-25.

Thomsen, E. (2003). BI's Promised Land. Intelligent Enterprise, (6)4, pp. 21-25.

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

Vaquero,L. Rodero-Marino, L., Caceres, J., Lindner, M. (2009). A break in the clouds: towards a cloud definition. SIGCOMM Computer Communication Review, 39, 137–150.

Valkonet, J. (2011). Interview with drs. Joris Valkonet on BI capabilities. May 20, 2011. Senior Consultant BI at Avanade.

Vecchiola, C., Chu, X., and Buyya, R. (2009). Aneka: A Software Platform for .NET-based Cloud Computing. In Gentzsch, W., Grandinetti, L., and Joubert, G. (eds), High Performance & Large Scale Computing, Advances in Parallel Computing, IOS Press, 2009.

Watson H.J., Volonino L. (2002). *Customer Relationship Management at Harrah's Entertainment*. Decision Making Support Systems: Achievements, Trends and Challenges for the Decade, M. Mora, G.A. Forgionne, and J.N.D. Gupta, eds., Idea Group Publishing, 2002, pp. 157-172.

Watson, H.J., Wixom, B.H. (2007). The current state of business intelligence. Computer 40(9), 2007, pp.96–99.

Webster, J., Watson, R.T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly (26:2) 2002, pp xiii-xxiii.

Weerd, I. van de, Brinkkemper, S. (2008). Meta-modeling for situational analysis and design methods. In M.R. Syed and S.N. Syed (Eds.), Handbook of Research on Modern Systems Analysis and Design Technologies and Applications (pp. 38-58). Hershey: Idea Group Publishing.

White, C. (2005). The Next Generation of Business Intelligence: Operational BI. *DM Review*, May 2005,13–14.

Whiting, R. (2003). Look Within—Business-Intelligence Tools have a New Mission: Evaluating All Aspects of a Company's Business. InformationWeek Januari 2003, pp. 32.

Willen, C. (2002). Airborne Opportunities. Intelligent Enterprise, (5)2, pp. 11-12. Januari 14, 2002.

Wixom, B.H., Watson, H.J. (2010). The BI-Based Organization. International Journal of Business Intelligence Research, vol. 1, no. 1, pp. 13-28.

Wu, E., Diao, Y., Rizvi, S. (2006). High-performance complex event processing over streams. In *Proceedings of the 2006 ACM SIGMOD international conference on Management of data.* ACM, 2006, pp. 407–418..

Zeng, l., Xu, L.D., Shi, Z.Z., Wang, M., Wu, W. (2006). Techniques, Process, and Enterprise Solutions of Business Intelligence. In Proceedings of the 2006 IEEE International Conference on Systems, Management and Cybernetics. Taipei, 2006.

Zhang, T., Ramakrishnan, R., Livny, M. (1996). BIRCH: An efficient data clustering method for very large databases. In *Proceedings of the 1996 ACM SIGMOD International Conference on Management Data*. Montreal, Canada, June 1996.

APPENDICES

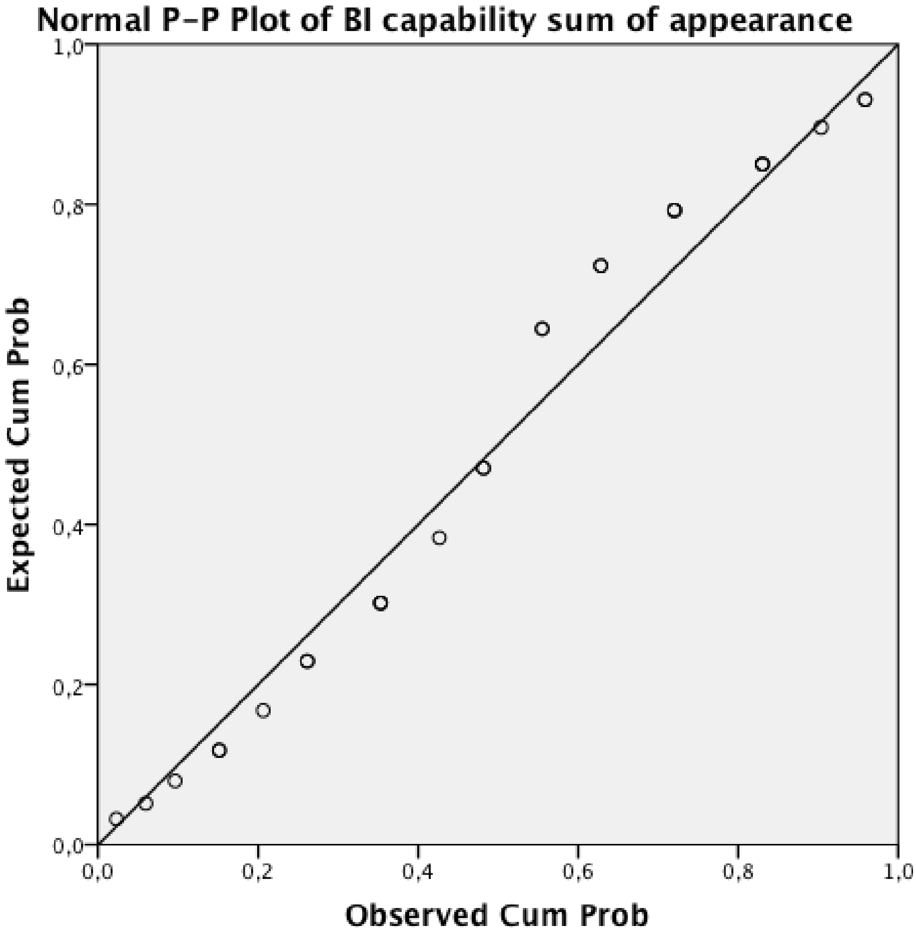
Appendix 1 - BI capability product comparison per product

	Microsoft				IBM			Oracle		SAP			SAS						Tableau software	ADVIZOR Solutions, Inc.		TIBCO		LogiXML	Arcplan	Strateg y	BOARD	TARGIT	Predixion Software	Appistry	Qliktech	Information Builder		
	SQL server	Sharepoint 2010	Report Builder	Excel (with PowerPivot)	Cognos	SPSS	InfoSphere	BI Enterprise Edition 11g	Database 11g	Business Objects	Data Warehousing	Netweaver	Enterprise BI Server	Visual BI	Customer Intelligence	Analytics	Data Management	Supply Chain Intelligence	Tableau Server	ADVIZOR Analyst	ADVIZOR Analyst/X	Spotfire	Data Quality	Logi Info	Arcplan Enterprise	MicroStrategy 9	Management Intelligence	BI suite	Insight	CloudIQ	Qlikview	webFOCUS	iWay	SUM
Data Warehousing	X						X		X		X		X				X		X	X	X				X		X			X	X		X	14
Data Mining	X		X	X	X	X			X	X			X			X			X		X	X		X	X	X		X	X		X			18
Automated exception detection									X													X										X	X	4
BI alerting								X	X	X					X							X			X			X				X	X	9
Data follow-through workflow									X																	X								2
Automatic learning and refinement									X																									1
Data visualization	X	X	X	X	X	X		X		X			X	X		X			X	X	X	X		X		X	X		X		X	X	X	22
Secure data delivery	X						X	X	X	X							X		X						X	X			X		X		X	12
Analytics	X			X	X	X	X	X	X	X				X	X	X		X	X	X	X	X			X	X	X	X	X	X	X	X		24
OLAP	X				X		X	X	X	X			X						X			X		X	X	X	X	X		X	X	X		17
Text mining	X				X				X	X					X	X										X		X		X	X			10
Reporting	X	X	X	X	X	X		X		X			X	X	X	X		X	X	X	X	X		X	X	X	X	X	X		X	X		25
Usability			X		X			X		X			X						X			X			X	X		X			X	X		12
Forecasting	X				X			X	X	X				X	X	X		X			X	X			X	X		X	X	X		X		17
Segmentation and clustering						X			X						X			X								X		X			X	X	X	9
Information management	X			X	X												X									X			X			X	X	8
Master data management	X						X		X			X					X						X							X			X	8
Data quality	X				X		X		X	X	X	X					X						X			X	X			X	X		X	14
Data independency	X																X			X	X												X	5
Collaboration		X		X	X			X		X					X			X	X					X	X	X	X	X	X		X	X		16
Business process embedded					X					X					X			X							X			X						6
Data transformation	X						X	X	X	X	X		X		X		X			X											X			11
Data modeling	X					X	X	X		X		X	X				X															X	X	10
Data preprocessing	X				X		X		X	X	X		X	X	X	X	X						X		X						X		X	15
Intelligent Warehousing							X		X		X																						X	4
Self service BI		X	X	X	X	X		X		X			X	X					X	X	X	X				X		X		X	X	X		18
ETL	X				X		X	X	X	X	X		X				X		X			X	X	X		X	X		X		X		X	18
TOTAL:	17	4	5	7	15	7	11	13	18	18	6	3	11	6	10	7	10	6	11	7	8	11	4	6	12	16	8	12	9	8	16	13	14	329

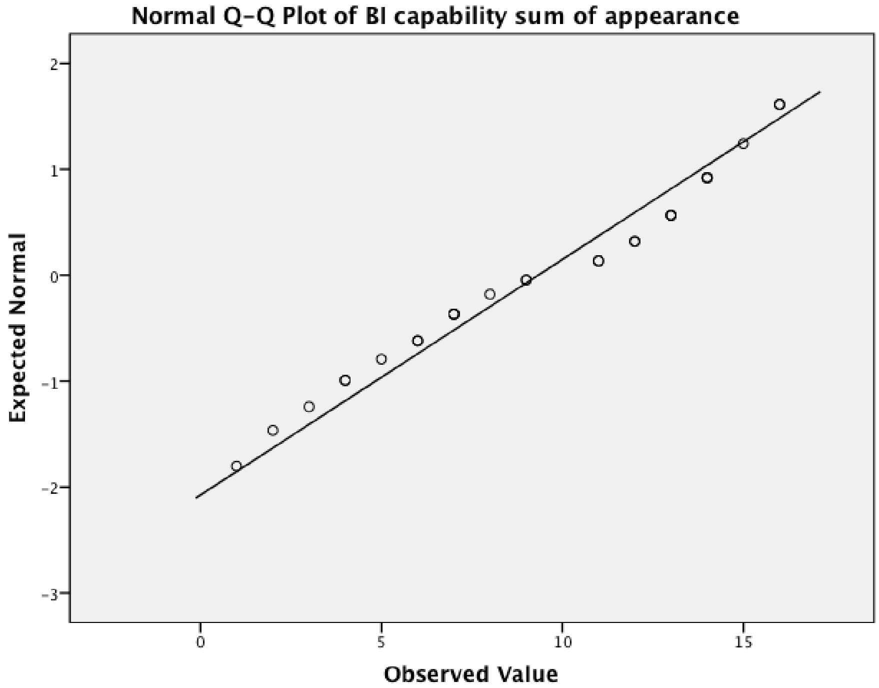
Appendix 2 - BI capability product comparison per vendor

	Microsoft	IBM	Oracle	SAP	SAS	Tableau software	ADVIZOR Solutions, Inc.	TIBCO	LogiXML	Arcplan	MicroStrategy	BOARD	TARGIT	Predixion Software	Appistry	Qliktech	Information Builder	SUM
Data Warehousing	X	X	X	X	X	X	X			X		X			X	X	X	12
Data Mining	X	X	X	X	X	X	X	X	X	X	X		X	X		X		14
Automated exception detection			X					X									X	3
BI alerting			X	X	X			X		X			X				X	7
Data follow-through workflow			X								X							2
Automatic learning and refinement			X															1
Data visualization	X	X	X	X	X	X	X	X	X		X	X		X		X	X	14
Secure data delivery	X	X	X	X	X	X				X	X			X		X	X	11
Analytics	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	16
OLAP	X	X	X	X	X	X		X	X	X	X	X	X		X	X	X	15
Text mining	X	X	X	X	X						X		X		X	X		9
Reporting	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	16
Usability	X	X	X	X	X	X		X		X	X		X			X	X	12
Forecasting	X	X	X	X	X		X	X		X	X		X	X	X		X	13
Segmentation and clustering		X	X		X						X		X			X	X	7
Information management	X	X			X						X			X			X	6
Master data management	X	X	X	X	X			X							X		X	8
Data quality	X	X	X	X	X			X			X	X			X	X	X	11
Data independency	X				X		X										X	4
Collaboration	X	X	X	X	X	X			X	X	X	X	X	X		X	X	14
Business process embedded		X		X	X					X			X					5
Data transformation	X	X	X	X	X		X									X		7
Data modeling	X	X	X	X	X												X	6
Data preprocessing	X	X	X	X	X			X		X						X	X	9
Intelligent Warehousing		X	X	X													X	4
Self service BI	X	X	X	X	X	X	X	X			X		X		X	X	X	13
ETL	X	X	X	X	X	X		X	X		X	X		X		X	X	13
TOTAL:	20	22	24	21	23	11	9	14	6	12	16	8	12	9	8	16	21	252

Appendix 3 - Normal P-P Plot of BI product review



Appendix 4 - Normal Q-Q Plot of BI product review



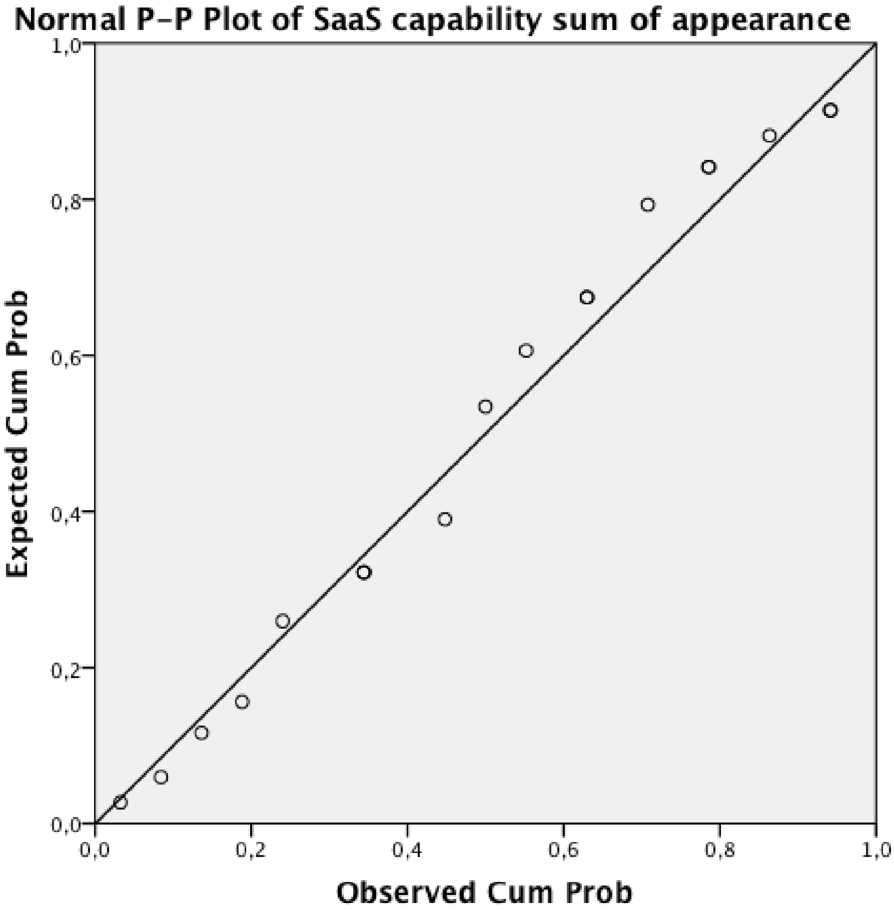
Appendix 5 - SaaS capability product comparison per product

	Microsoft			Google			Amazon				Sales-force.com		Appistry	AT&T			Engine Yard		Enomaly	Flexiant	
	Azure	Office 365	CRM online	Cloud Connect	Docs	Sites	Elastic Compute Cloud (Amazon EC2)	CloudFront	RDS	S3	force.com	Sales Cloud	CloudIQ	Hosting Services	Cloud Computing Services	Application Management Services	AppCloud	xCloud	Elastic Computing Platform (ECP)	Flexiscale	SUM
Automatic adaptation							X	X		X									X		4
Competitive costs				X		X															2
Configurability							X	X			X	X			X		X	X	X	X	9
Data center	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	19
Data protection		X	X			X	X		X	X		X		X	X	X	X		X	X	13
Vendor controlled	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X			X	X	17
Hosted service		X	X	X	X	X	X	X	X	X	X	X	X	X		X				X	15
SLA		X	X				X		X	X	X	X		X		X			X		10
Internet Centric	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	20
Monitoring		X	X				X		X					X	X	X	X	X	X	X	11
Multi-tenancy	X			X			X	X		X				X		X	X		X	X	10
Payment model	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		X	18
QoS	X	X	X			X	X		X	X	X	X		X	X	X		X	X	X	15
Resource allocation				X			X	X		X					X		X		X		7
Resource optimization		X					X	X		X							X		X		6
Scalability	X		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	18
Time & location independent	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	20
User friendliness				X		X		X	X	X				X		X	X		X	X	10
Virtualization	X	X					X	X	X	X	X	X		X	X		X	X	X	X	14
TOTAL:	9	12	11	11	5	11	17	13	13	16	14	12	7	14	11	13	13	9	16	14	238

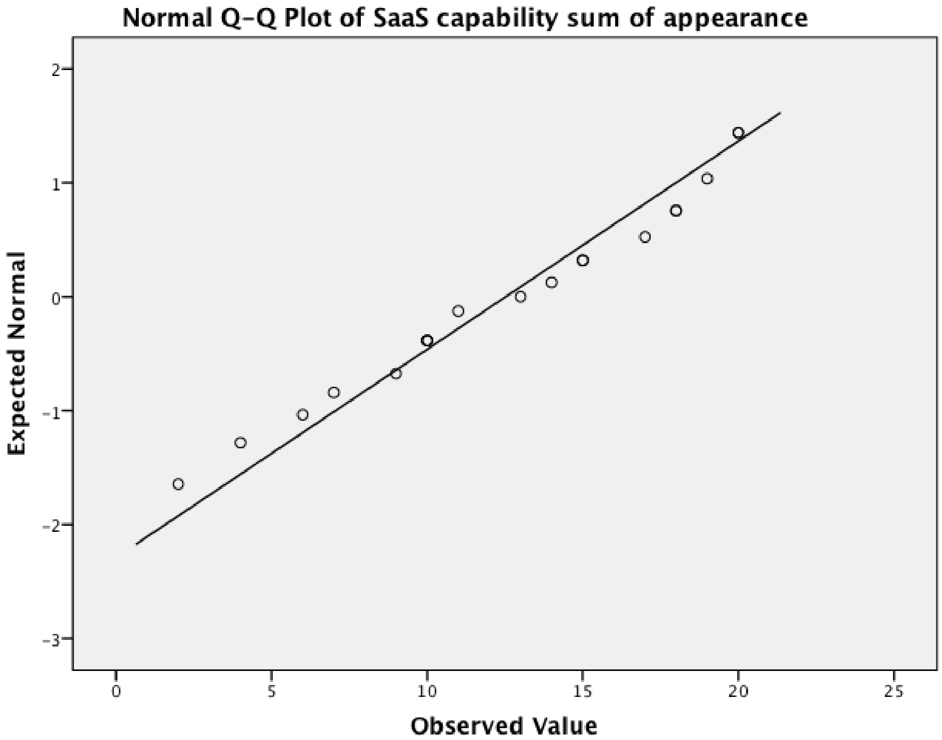
Appendix 6 - SaaS capability product comparison per vendor

	Microsoft	Google	Amazon	Salesforce.com	Appistry	AT&T	Engine Yard	Enomaly	Flexiant	SUM
Automatic adaptation			X					X		2
Competitive costs		X								1
Configurability			X	X		X	X	X	X	6
Data center	X	X	X	X	X	X	X	X	X	9
Data protection	X	X	X	X		X	X	X	X	8
Vendor controlled	X	X	X	X	X	X		X	X	8
Hosted service	X	X	X	X	X	X	X	X	X	9
SLA	X		X	X		X		X		5
Internet Centric	X	X	X	X	X	X	X	X	X	9
Monitoring	X		X			X	X	X	X	6
Multi-tenancy	X	X	X			X	X	X	X	7
Payment model	X	X	X	X	X	X	X		X	8
QoS	X	X	X	X		X	X	X	X	8
Resource allocation		X	X			X	X	X		5
Resource optimization	X		X				X	X		4
Scalability	X	X	X	X	X	X	X	X	X	9
Time & location independent	X	X	X	X	X	X	X	X	X	9
User Friendliness		X	X			X	X	X	X	6
Virtualization	X		X	X		X	X	X	X	7
TOTAL:	14	13	18	12	7	16	15	17	14	126

Appendix 7 - SaaS product review analysis normal P-P plot



Appendix 8 - SaaS product review analysis normal Q-Q plot



Appendix 9 – The BIaaS Maturity Matrix

	0	1	2	3	4	5	6	7	8
Design									
Data Storage		A: Data Center			B: Multi-tenancy		C: Recource Allocation		
Accessibility		A: Internet Centric	B: Hosted Service		C: Data Protection				D: Resource Optimization
Data									
Data gathering		A: ETL						B: Data Independency	C: Data follow-through workflow
Data management		A: Data Warehousing	B: Secure data delivery	C: Data Quality	D: Master Data Management		E: Intelligent Warehousing		F: Information management
Data processing					A: Data preprocessing		B: Data transformation	C: Segmentation and Clustering	D: Automated learning and refinement
Data analysis		A: Analytics	B: OLAP	C: Data Mining	D: Forecasting	E: Text Mining	F: Data Modeling		
Usage									
Application logic		A: Scalability	B: Vendor Controlled		C: Virtualization				D: Automatic adaption
Usability					A: Monitoring	B: User Friendliness	C: Configurability		
Consumerization		A: Reporting	B: Data visualisation	C: Collaboration	D: Self-service BI	E: Portability	F: Business Process Embeded		
Support									
Alerting							A: BI alerting		B: Automated exception detection
Service		A: Time & location independent	B: Payment Model	C: QoS	D: SLA				E: Competitive costs