

**A DATA MANAGEMENT AND ANALYTIC MODEL FOR BUSINESS
INTELLIGENCE APPLICATIONS**

by

MISHECK BANDA

Submitted in accordance with the requirements for the degree of

MASTER OF SCIENCE IN COMPUTING

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF E K NGASSAM

FEBRUARY 2017

DECLARATION

I declare that "A Data Management and Analytic Model for Business Intelligence Applications" is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

I further declare that I have not previously submitted this work, or part of it, for examination at UNISA for another qualification or at any other higher education institution.

A handwritten signature in black ink, appearing to read "Bawli", is positioned above a horizontal line.

SIGNATURE

22-05-2017

DATE

ACKNOWLEDGEMENT

First and foremost, I would like to give all the glory and thank the almighty God for making it possible for me to complete my dissertation.

Secondly, this dissertation would not have been written without the help of my supervisor, Professor Ernest Ketcha Ngassam. Thank you for your guidance, constructive feedback and for always encouraging me and keeping me on track.

Thirdly, my special thanks to my editor Dr. Jackie A. Fourie for proofreading and editing my dissertation as well as providing advice on formatting and referencing style.

Fourthly, thank you to the respondents from various organisations for finding time to complete the research survey questionnaires.

Fifthly, my special thanks to my family, and my wife Cynthia Banda for her support, understanding, encouragement, and love. Thanks also to my children, Chipwali Misheck Jr, Natalie, and Angelina for their patience and support. To my mom, Densina Banda for her support and encouragement, and my sister, Sheena Mundale, and her husband, Charles Mundale for their support.

Finally, thanks to all my friends and other family members for their support and encouragement, as well as my colleagues at Mindworx Consulting for their support and encouragement throughout my studies.

LIST OF PUBLICATIONS

Papers reviewed by the IST-Africa's International Programme Committee and accepted for publications:

- Banda, M & Ketcha Ngassam, E. (2017). A data management and analytic model for business intelligence applications. In *IEEE IST-Africa Conference Proceedings, 2017:1-16*.

ABSTRACT

Most organisations use several data management and business intelligence solutions which are on-premise and, or cloud-based to manage and analyse their constantly growing business data. Challenges faced by organisations nowadays include, but are not limited to growth limitations, big data, inadequate analytics, computing, and data storage capabilities. Although these organisations are able to generate reports and dashboards for decision-making in most cases, effective use of their business data and an appropriate business intelligence solution could achieve and retain informed decision-making and allow competitive reaction to the dynamic external environment.

A data management and analytic model has been proposed on which organisations could rely for decisive guidance when planning to procure and implement a unified business intelligence solution. To achieve a sound model, literature was reviewed by extensively studying business intelligence in general, and exploring and developing various deployment models and architectures consisting of naïve, on-premise, and cloud-based which revealed their benefits and challenges. The outcome of the literature review was the development of a hybrid business intelligence model and the accompanying architecture as the main contribution to the study.

In order to assess the state of business intelligence utilisation, and to validate and improve the proposed architecture, two case studies targeting users and experts were conducted using quantitative and qualitative approaches. The case studies found and established that a decision to procure and implement a successful business intelligence solution is based on a number of crucial elements, such as, applications, devices, tools, business intelligence services, data management and infrastructure. The findings further recognised that the proposed hybrid architecture is the solution for managing complex organisations with serious data challenges.

Key Words:

Analytics, Business Intelligence Applications, Big Data, Data Management, Data Warehouse, Cloud/Cloud Computing, Extract, Transform, and Load, On-line Analytical Processing, On-line Transaction Processing, Open Source Licensed BI Tools, BI Total Cost of Ownership

TABLE OF CONTENTS

DECLARATION	II
ACKNOWLEDGEMENT	III
LIST OF PUBLICATIONS.....	IV
ABSTRACT	V
LIST OF FIGURES.....	IX
LIST OF TABLES	X
DEFINITIONS OF KEY CONCEPTS	XI
CHAPTER 1 INTRODUCTION.....	1
1.1 Introduction	1
1.2 Background and motivation.....	3
1.3 Statement of the research problem.....	9
1.4 Purpose and objectives of the research	9
1.5 Research questions	11
1.6 Research outcomes	12
1.7 Research methodology	12
1.8 Preliminary literature review	15
1.9 Ethical considerations.....	19
1.10 Significance of the study.....	19
1.11 Limitations of the study.....	20
1.12 Dissertation layout.....	21
1.13 Conclusion.....	21
CHAPTER 2 BUSINESS INTELLIGENCE OVERVIEW	23
2.1 Introduction	23
2.2 Definition of business intelligence	23
2.3 Business intelligence value chain	26
2.4 Naïve business intelligence model.....	32
2.5 Business intelligence architecture	36
2.6 Application of on-line transaction processing	42
2.7 Application of on-line analytical processing	43
2.8 Total cost of ownership of BI deployment.....	44
2.9 Conclusion.....	47
CHAPTER 3 ON-PREMISE BUSINESS INTELLIGENCE.....	48
3.1 Introduction	48

3.2	On-premise business intelligence deployment	48
3.3	On-premise business intelligence model.....	54
3.4	Architecture for on-premise deployment of business intelligence	58
3.5	Advantages of on-premise business intelligence deployment	62
3.6	Limitations of on-premise business intelligence	65
3.7	Total cost of ownership of on-premise BI deployment.....	67
3.8	Conclusion.....	68
CHAPTER 4 CLOUD-BASED BUSINESS INTELLIGENCE		69
4.1	Introduction	69
4.2	Foundation of cloud-based business intelligence	69
4.3	Cloud-based business intelligence model.....	81
4.4	Cloud-based business intelligence architecture	86
4.5	Advantages of cloud-based business intelligence	94
4.6	Limitations of cloud-based business intelligence	99
4.7	Total cost of ownership of cloud-based BI deployment	102
4.8	Conclusion.....	104
CHAPTER 5 HYBRID DEPLOYMENT OF BUSINESS INTELLIGENCE		106
5.1	Introduction	106
5.2	Preliminaries.....	107
5.3	Overview of hybrid business intelligence.....	116
5.4	Hybrid business intelligence model	119
5.5	Architecture for a hybrid business intelligence solution	123
5.6	Advantages of hybrid business intelligence	128
5.7	Limitations of hybrid business intelligence.....	133
5.8	Total cost of ownership of cloud-based BI deployment.....	136
5.9	Conclusion.....	137
CHAPTER 6 DATA COLLECTION.....		139
6.1	Introduction	139
6.2	Case study	140
6.3	Data gathering	140
6.4	Research instrument	142
6.5	Objectives and categories of questions.....	145
6.6	Research software tools.....	147

6.7	Sample design.....	147
6.8	Data preparation and cleaning.....	149
6.9	Ethical clearance and considerations	150
6.10	Pilot study.....	150
6.11	Study limitations	151
6.12	Conclusion.....	152
CHAPTER 7 STATE OF BI DEPLOYMENT IN ORGANISATIONS		153
7.1	Introduction	153
7.2	Data analysis	153
7.3	Results of part 1: BI business users	156
7.4	Findings part 1: state of BI utilisation in organisations.....	164
7.5	Conclusion.....	167
CHAPTER 8 AN IMPROVED HYBRID BI ARCHITECTURE		169
8.1	Introduction	169
8.2	Data analysis	169
8.3	Results of part 2: BI experts responses.....	170
8.4	Research findings	184
8.5	Improved BI architecture	188
8.6	Practical demonstration of the proposed hybrid BI architecture	191
8.7	Conclusion.....	196
CHAPTER 9 SUMMARY AND CONCLUSION.....		198
9.1	Introduction	198
9.2	Research findings	198
9.3	Research summary	201
9.4	Future research.....	206
9.5	Conclusion.....	207
REFERENCES.....		209
APPENDIX 1: ETHICAL CLEARANCE DOCUMENTATION		221
APPENDIX 2: RESEARCH QUESTIONNAIRE – BI EXPERTS.....		222
APPENDIX 3: RESEARCH QUESTIONNAIRE – BI USERS		237
APPENDIX 4: LETTER FROM EDITOR.....		246
APPENDIX 5: TURNITIN SUMMARY REPORT.....		247

LIST OF FIGURES

Figure 2-1: Business Intelligence Value Chain	26
Figure 2-2: Naïve BI Model	33
Figure 2-3: BI Architecture	37
Figure 2-4: Corporate Sales Overview	38
Figure 3-1: Isolated BI Deployment	50
Figure 3-2: Overlapping BI Deployment	51
Figure 3-3: Centralised BI Deployment	53
Figure 3-4: On-Premise BI Model.....	55
Figure 3-5: On-Premise BI Architecture.....	59
Figure 4-1: Cloud Services Models	71
Figure 4-2: Public Cloud Deployment.....	75
Figure 4-3: Private Cloud Deployment	77
Figure 4-4: Cloud-based BI Model.....	82
Figure 4-5: Cloud-based BI Architecture	87
Figure 4-6: SaaS Architecture	92
Figure 4-7: SaaS and PaaS Architecture.....	93
Figure 4-8: SaaS and IaaS Architecture	94
Figure 5-1: Traditional Query vs In-memory Query Approach	108
Figure 5-2: Integrated Analytical Approach	110
Figure 5-3: Traditional Data Integration and Analytics	112
Figure 5-4: Conventional Data Processing vs Embedded Processing	115
Figure 5-5: Hybrid Business Intelligence Model.....	120
Figure 5-6: Hybrid BI Architecture	124
Figure 7-1: Years of experience	157
Figure 7-2: Organisation size	158
Figure 7-3: Industry of organisations.....	159
Figure 7-4: Users accessing BI services	159
Figure 7-5: Software capabilities supporting BI solution	161
Figure 7-6: Management of data source or generating systems.....	162
Figure 7-7: Location and management of BI solutions.....	163
Figure 7-8: Buying or licensing options	164
Figure 8-1: Years of Experience.....	171
Figure 8-2: Industry of Organisation	172
Figure 8-3: Level of Deployment of BI	173
Figure 8-4: Geographical Distribution of Organisations	173
Figure 8-5: Expertise and Knowledge of BI Deployment Options	175
Figure 8-6: Ideal Pricing Models for BI	176
Figure 8-7: Front-end Layer – Proposed Components	177
Figure 8-8: Front-end Layer: Additional Components.....	178
Figure 8-9: Service Layer – Proposed Components.....	179
Figure 8-10: Service Layer – Additional Components.....	180
Figure 8-11: Back-end Layer – Proposed Components.....	181
Figure 8-12: Back-end Layer – Additional Components	182

Figure 8-13: Infrastructure Layer – Proposed Components	183
Figure 8-14: Infrastructure Resources.....	184
Figure 8-15: Improved Hybrid BI Architecture	189

LIST OF TABLES

Table 0-1: Definition of Key Terms	XI
Table 2-1: Naïve TCO Cost Categories.....	45
Table 3-1: TCO On-premise Deployment.....	67
Table 4-1: TCO Cloud-based Deployment	103
Table 5-1: On-premise and Cloud-based Summary Challenges.....	119
Table 5-2: Hybrid BI Related Advantages	129
Table 5-3: Hybrid BI Related Disadvantages.....	134
Table 5-4: TCO Hybrid Deployment	136
Table 6-1: Data Preparation and Cleaning Process	149
Table 7-1: Data Preparation Process	154
Table 8-1: TCO Calculation for BI Deployment.....	195

DEFINITIONS OF KEY CONCEPTS

The business environment is ever changing and so is the introduction of new business terminology. In order to have a common understanding and eliminate ambiguity in respect of the key terms used in this research, the relevant definitions are provided below:

Table 0-1: Definition of Key Terms

Key term	Definition
<i>Analytics</i>	Kolb (2012:5) defines analytics as the “collection of technologies and processes that turn raw data into usable knowledge in order to inform decisions and drive action.” Analytics is the utilisation of technology that allows organisations to infer why certain business transactions happened, and could happen, in a certain manner by providing intelligence and insightful reports (Harriott 2013:36). It is critical that analytical tools are used in addressing the key challenges that an organisation faces as they have capabilities of predicting future scenarios based on current situations.
<i>Business Intelligence Applications</i>	<p>Gartner (2013:1) defines BI as an “umbrella term that includes the applications, infrastructure and tools, and best practices that enables access to, and analysis of information to improve and optimize decisions and performance”. BI is “a type of application software that is designed to report, analyse, and present data” (Manyika et al 2011:31). This is achieved through using BI tools that are often used to read data that has been previously stored in a data warehouse or data mart. <i>The Dictionary.com</i> (2013:1) defines BI as “the methods and technologies that gather, store, report, and analyse business data to help people make business decisions” which are often referred to as BI software; BI tools; BI applications; the list is not exhaustive.</p> <p>All these definitions have a common theme, “<i>manipulating stored data in order to present information so that business people can make informed decisions</i>”. This is usually achieved by using analytical tools to offer data as an asset as they are capable of finding patterns and data trends with data mining tools (Kolb 2012:4). BI in this context refers to BI applications with capabilities of performing management and analytical functions in respect of stored business data with the view to generating business value.</p>
<i>Big Data</i>	Big data refers to datasets the size of which is “beyond the ability of typical database software tools to capture, store, manage, and analyse” (Manyika et al 2011:1) and it is assumed that, as technology advances over time, the size of datasets that qualify as big data will also increase. Gartner (2017:1) defines big data as “high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation.” Big data are the massive or large volumes of business data that organisations generate, accumulate and experience problems in managing and analysing with their current BI applications.
<i>Data Management</i>	Data management involves the processing and delivery of collected and stored data whether old data, new big data, or both (Russom 2013:5). Data management is crucial for BI, analytics and business value, as without old and new data, organisations will not be able to make informed decisions.
<i>Data Warehouse</i>	Data warehousing is a process of gathering all organisational data from multiple data source systems into a structured central repository for the purpose of optimal utilisation in decision-making processes (Boateng, Singh & Singh 2012:2; Manyika et al 2011:32). For BI applications to work effectively, they need to be connected to a stable data store, such as, a data warehouse as it contains data fit for BI purpose.

Key term	Definition
<i>Cloud/Cloud Computing</i>	Armbrust, Fox, Griffith, Joseph, Katz, Konwinski, Lee, Patterson, Rabkin, Stoica, and Zaharia (2010:50) define cloud computing as “both the applications delivered as services over the Internet and the hardware and systems software in the data centres that provide those services”, such as, SaaS BI, etc. The hardware and software applications in the data centre are what are often referred to as the Cloud. There are various cloud offerings available, such as, the public cloud, and the private cloud which are made available to the general public or private organisations on a pay-as-you-go manner for using the necessary BI services.
<i>Extract, Transform, Load</i>	ETL tools use a complex process of extracting data from various heterogeneous sources, transforming it into correct and standardised data fit for BI purpose and loading quality data into a data repository, such as, a data warehouse, and data mart. (Ouf & Nasr 2012:655; Manyika et al 2011:32). This process ensures that data fit for BI purpose is stored in the right data repository.
<i>On-line Analytical Processing</i>	OLAP are analytical systems that are able to execute analytical functions on data stored in the data warehouse, data mart or operational data store for the purpose of BI (Williams & Williams 2007:201). OLAP systems are capable of being linked to multiple data sources and provide analytical functions which provide flexible reporting (Power 2008:150). The analytical functions in OLAP enable organisations to generate reports which assist in better decision-making.
<i>On-line Transaction Processing</i>	OLTP are operational systems with the capabilities to process big data emanating from transactions in real-time and to update data continuously (Ong, Siew & Wong 2011:4). OLTP systems and other sources use ETL tools to generate most of the data found in the data warehouse (Sen, Ramamurthy & Sinha 2012:336). The data that is generated from OLTP systems is crucial for BI applications for reporting purpose.
<i>Open Source Licensed BI Tools</i>	Open source licensed applications allow any user to freely make changes to the source code (Thomsen & Pedersen 2008:75). If the user makes changes and uses the source code privately, then the user does not need to release the code. However, if shared, then the code must be released under a general public license. There are a number of BI tools that use open source platforms, such as, DBMS, ETL, and OLAP servers which have been discussed in this study.
<i>BI Total Cost of Ownership</i>	TCO is a comprehensive assessment of IT-related costs or other costs across the organisation boundaries over a period of time (Gartner 2015b). The TCO analysis helps to obtain an accurate and transparent cost model for BI investments through direct costs, indirect costs and related costs (King 2007:8; Plattener & Zeier 2012:16).

CHAPTER 1 INTRODUCTION

1.1 Introduction

Nowadays, organisations are processing large numbers of transactions which result in the storage of large volumes of data that are difficult to manage and analyse in order to produce useful information for decision-making (Rahman, Aldhaban & Akhter 2013:543; Sun, Gao, Fan & An 2012:512). Acker, Grone, Blockus, and Bang (2011:133) emphasise that the volume of business data that is being digitally generated and managed is growing rapidly, and at the same time, the need to gain competitive advantage by studying the past and predicting future situations, has put the responsibility on organisations to manage and retain business data and use it for many for years.

Large volumes of data, also known as big data, have exceeded the capacity of existing data management technologies making it impossible to process, store and manage business data in-house (Mate, Llorens & De Gregorio 2012:111). Both small and large organisations are finding it difficult to put together huge amount of data so that business intelligence (BI) systems can be utilised to perform real-time analytics and other relevant functions. As a result, most organisations remain uncompetitive due to a lack of informative reports and quality reporting capabilities.

Big data and informative reports are associated when the organisation attempts to forecast its business operations, but organisations find it difficult to select the right strategies for implementing a BI solution with analytics capabilities, such as predictive, and descriptive analysis. (Barnes, Ring, Yamunan, Franks, Hiltbrand, Taylor, & Williams 2012:21; O'Donnell, Sipsma & Watt 2012:210). A further problem resulting from large volumes of data is the difficulty of managing a variety of mixed information coming from different sources that reside in local servers or in the cloud. Increasingly, most organisations have opted to invest in third party tools in order to manage large volumes of mixed data so that data can be delivered from various sources for reporting purposes.

In the past decades, some of the organisations that have used decision-making systems have experienced tremendous changes, such as been able to store, combine, and aggregate data and use the results for analysis, and have created vast

opportunities that are associated with the use of big data (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh & Byers 2011:1). However, it remains uncertain as to whether organisations are utilising the appropriate BI applications that are capable of capturing, storing, managing and analysing data that is rapidly increasing.

There has also been an increasing demand for the usage of smart mobile devices in the workplace, prompting the need to have sophisticated information generating applications running on the devices and providing information in real time (O'Donnell et al 2012:209). Most organisations promote, or accommodate the use of mobile devices into their operational activities which are capable of running a number of applications that could include BI applications, and data mining applications, but often face the challenge of configurations and integrations. The use of mobile devices is constantly increasing and, at the same time, it also encourages collaboration among organisations and business users from various locations.

Collaborative BI applications that enable organisations, their users and service providers to work together, irrespective of where they are located, are rarely found on various technological devices, such as, computers, laptops, and mobile devices (Poonen 2011:3). It is critical that organisations are able to use a range of collaborative BI applications that enable the sharing of critical information in real-time from various locations. Collaboration also comes with the need to have proper governance in place with respect to data management and reporting systems as well as communication among different technology platforms (Minoli 2008:442). Most organisations grant business users access to their BI systems without properly configuring the systems with regards to indicating which users are supposed to work on which specific reports, this often results in duplication and inconsistency in reporting.

Some organisations are still using old versions of BI applications because they are comfortable with the status quo as they are not able to easily find evidence of a return on their investment (O'Donnell et al 2012:209). This is causing organisations to miss out on the benefits that come with using state-of-art technology that currently supports BI applications. Organisations need to have quality information and guidelines that would encourage them to upgrade to any available latest BI application that is affordable and fit for purpose.

Use of old versions of applications often result in challenges in regard to application availability due to issues related to technological incompatibility (Andreadis, Fourtounis & Bouzakis 2015:67). New BI applications that are procured either through a once-off purchase or license per user option may fail to work on older workstations or mobile devices. This may be due to the fact that the operating system is not compatible with the new applications. This will mean that all business users who make use of state-of-the-art technology will not be able to access and utilise BI applications that make use of legacy systems. If so, organisations need to consider investing in software that will enable legacy systems to work with new BI solutions, which may be costly.

Having discussed some of the above-mentioned challenges and benefits, the remainder of this chapter is structured as follows: section 1.2 provides the background to the research, and the statement of the research problem is discussed in section 1.3. The objectives of the research are listed in section 1.4 followed by related research questions and research outcomes in section 1.5 and section 1.6 respectively. In section 1.7, a synopsis of the data collection and data analysis methodology used in the study is discussed. This is followed by ethical considerations in section 1.8. Section 1.9 and section 1.10 discuss the significance and limitations of the study, respectively. Definitions of key concepts for the study are provided in section 1.11 while the layout of the dissertation is provided in section 1.12. The chapter concludes with a summary discussion in section 1.13.

1.2 Background and motivation

In the past decades, data management and analytics systems have evolved tremendously in an effort to keep up with technology advancement and to meet business requirements (Gray 1996:39; Kempler, Lynnes, Vollmer, Alcott & Berrick 2009:21-24; Power 2008:150). They further declare that data management and analytics systems date back prior to the 1950s when everything was done manually and data was stored on punch cards.

The mid-1950s to 1970 saw the beginning of the software programming languages, such as, COBOL which was used to develop computer programs that could read,

analyse and report on records (Gray 1996:39). Programs, such as, batch transaction processing systems were used to process data that were captured and stored on cards or magnetic tapes which was collected in batches and stored in a database called, Masterfile. The type of computer programs or systems that were used during this period includes data processing, electronic data processing systems, management reporting, management information systems, and database management systems. and used magnetic tapes and/or hard-disk drives to store data (Zhang 2009:1; Trikha 2010:1449). During this era, a few drawbacks were experienced, such as, late detection of transaction errors as business users had to wait until the daily run for a Masterfile to be completed in order to attend to it. It was also difficult to know the current status of the Masterfile as changes were made only after the overnight batch run.

The mid-1960s to the 1980s, saw the birth of on-line network databases which shifted the focus from having to wait for batched processed data to data being available immediately on-line (Gray 1996:40). Tele-processing monitors and computers were the hardware used to interactively access data on-line by facilitating the exchange of requested messages from terminal to servers and back. The type of programs that were commonly used during this period includes management reporting, management information systems, database management systems, decision support, decision support systems, and group decision support systems. (Zhang 2009:1). Some systems continued using magnetic tapes and, or hard-disk drives to store data (Trikha 2010:1449). During this period data was organised in indexed sequential records or advanced set-oriented records which came with the problem of updating records as data was redundantly stored. Data independence was maintained and organised through the use of logical schemas that ensured logical relationships among records. Concurrent access to the Masterfile, which was not possible in the previous period, was resolved in this period by the introduction of a system that did not allow to make changes to a record that is opened or accessed, which is also known as transaction locking. Transaction locking ensured that systems could access certain records that were tracked through a log and be able to make changes to them.

Relational databases and client-server computing emerged between the 1980s and 1995 (Gray 1996:41). Relational databases presented a solution to challenges faced

in the previous period by offering uniform representation, unified language for data definition, data navigations and data manipulation and mostly used hard-disk drives for storing data (Tripathi 2010:1450). During this period, client-server computing offered the capturing of inputs and displaying of outputs to the client while the server managed the storing of data in the database, processing of requests and generating the desired outputs. Relational databases promoted parallel processing of data which reduced data mining tasks that were taking weeks to complete, to only a few hours. During this period, the use of relational databases and graphical user interfaces enabled multiple users to construct and execute complex queries at a given time. Furthermore, according to Gray (1996:41) SQL database language was also standardised in 1985, becoming the common model and language for data modelling and application programming across almost all database systems. One of the key limitations of this period was that domain specialists were unable to define and implement different data types in the database systems which made it difficult to search, compare and manipulate data.

This period saw the birth of intelligence decision support systems, executive information systems, on-line analytical processing (OLAP) systems, etc., which were also referred to as BI systems (Power 2008:150; Zhang 2009:1). They further asserted that these BI systems were capable of analysing large volumes of historical data and produced data-driven business reports that helped organisations make informed decisions. The success of data-driven business reporting relied mainly on easy and rapid access to large volumes of good quality business data residing in a data warehouse (Silva, Vollino, Becker & Galante 2012:42). Data warehousing was used to centralise and integrate historical and subject-oriented data which was targeted for decision-making. Another limitation of this period was that, organisations started creating an increasing amount of historical data which resulted in storage and retrieval problems (Power 2008:153).

From the mid-1990s to the mid-2000s, multimedia databases were introduced which were capable of storing different richer forms of information, such as, data, records, documents, images, maps, sound recordings, videos, and graphs (Gray 1996:43). This reflected the transition of traditional relational databases to object-oriented relational databases that are capable of storing complex data types. It was during this

period that unifying procedures and data replaced the client-server computing model with procedures and workflow. This meant that preconditions and scripted tasks were executed when a certain event occurred. The World Wide Web became very active, generated and allowed movement of huge volumes of different types of data and objects which brought about new data management challenges.

Some of the types of business information systems that emerged during this period were Enterprise Resource Planning, Enterprise Application Integration, data warehousing and data mining, BI, On-line Transaction Processing (OLTP), etc. (Zhang 2009:1). The use of Extract, Transform and Load (ETL) processes and Data Warehousing was crucial to ensuring data quality. Complex ETL processes were responsible for extracting data from various internal and external data sources, such as, ERP, and CRM and transforming it into a format that fits BI purpose and thereafter loading it into the data warehouse (Silva, Becker & Galante 2013:244). Data warehousing, on the other hand, was responsible for ensuring that correct, useful and granular data was available. Data warehousing was used to address some of the challenges that emerged in the previous period relating to absence of standards and unclear processes in the distributions of different data sources. Traditional BI tools, such as, pivot tables, and dashboards were used to explore the data warehouse in order to generate the required reports.

In the mid-2000s to the present, one of the projects that pushed data management systems to the limit was the implementation and evolution of the National Aeronautics and Space Administration's (NASA) project for the Earth Observing System Data and Information System (EOSDIS) (Kempler et al 2009:23). Some of the objectives of the EOSDIS were, "increasing end-to-end data system efficiency and autonomy while decreasing operations costs; increasing data interoperability and usability by the science research, application, and modelling communities; improving data access and processing; and ensuring continued safe stewardship". In order to achieve this, NASA had to evolve architectural components, that is, data management, data access, search services, data processing, discipline-specific archive, operational support and community interaction. This evolution resulted in an integrated data management system and sophisticated techniques to performing data analysis. However, the

common challenge of how to manage big data still persisted during this period which encouraged many organisations to seek for solutions.

Since business data was constantly growing, this situation brought a number of challenges. For instance, there were limited guidelines, a shortage of analytical and managerial talent needed to utilise large volumes of data, and storage capacity limitations which organisations needed to begin addressing immediately (Manyika et al 2011:3). The process of pulling out business knowledge from large volumes of information poses significant research and practical business challenges, especially considering the legal requirements for building a privacy-preserving and reliable information infrastructure (Chen, Chiang & Storey 2012:1171).

This period saw a wide adoption of various data storage media, such as, solid-state drives, in-memory technology, and cloud storage. In the early 2000s, due to technology advancement, the use of solid-state drives increased although its existence dates from the 1980s (Huang, Serres & Narayana 2010:42). The other storage medium which has gained popularity is the use of in-memory technology that has faster retrieval and response speed than any other storage medium or device (SAP 2013:1; Plattener & Zeier 2012:21). Although both these storage media have been in existence for some time, they have only been widely adopted recently due to the advancement of the underlying technology and available affordable options (SAP 2013:1). Cloud storage is another system that has been embraced widely due to the ease of access to unlimited data storage resources which reside in cloud data servers (Tripathi 2010:1451). A business user with a device, such as a laptop that connects to the Internet can easily access, store, manipulate and retrieve data using a web-based interface. However, many business users still remain concerned with the security and reliability of cloud storage services.

In order to address some of the above-mentioned challenges, this period saw the beginning of BI solutions that were provided through cloud computing resources, such as, Software-as-a-Service (SaaS) and supported through Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS) (Yang 2012:4, Garrison, Kim & Wakefield 2012:63). SaaS introduced web-based analytical tools that are accessible anywhere in the world as long as a user can connect to the Internet. However, clear guidelines or models for guiding organisations to implement a sound BI solution that

can leverage from available cloud computing resources remained limited. The problem also arose from the fact that little was known about what business value can be achieved from using BI applications in the cloud (Khan, Amin & Lambrou 2010:7). Over the past few years, the understanding has grown that decision-making systems have evolved and supported organisations to a certain extent.

The primary focus of traditional BI's has been to analyse historical data and to produce and meet certain business requirements (Barnes et al 2012:20). However, organisations realised that there is more that they can do with the ever growing business data in order to try and forecast business operations. In order to meet this need, BI solution service providers improved their analytical offering by introducing features, such as, predictive analytics with capabilities of predicting the future and prescriptive analytics which ensures that organisations know what to do to influence the predicted future situation (Hazen, Boone, Ezell & Jones-Farmer 2014:72; Barnes et al 2012:21). It has been understood that business data will continue to grow, and as it grows, the utilisation of appropriate BI applications will be crucial. The more accurate data an organisation has, the better it will be able to predict the direction of the business and make informed decisions. This will also provide a competitive advantage or better decision-making since not all big or small organisations are getting the full benefit of having a BI solution.

Apart from sales information emanating from business white papers and reports, some questions still remain unanswered academically, such as, what would be the preferable implementation type for data management systems, BI and analytic applications? Literature has shown that a number of BI solution vendors are moving their current locally based BI software solutions to the cloud in order to take full advantage of BI and analytical resources in the cloud. Cloud computing resources have been believed to allow most organisations to overcome the problem of processing power faced by most of the traditional BI solutions which were used to manage big data (Acker et al 2011:132). In a local server based environment, some BI functions usually take a lot of time to deliver expected results leading to delays in getting and providing the required information on time.

In the light of all the data management and BI-related challenges and limitations discussed above, a comprehensive investigation was proposed to be undertaken to

address them. A possible solution was to develop a data management and analytic model that could inform, guide and encourage most organisations to procure and use appropriate BI applications for managing large volumes of data in order to make better decisions and remain competitive.

1.3 Statement of the research problem

Organisations continue to be uncertain as to whether the utilisation of BI applications will enable them to adequately capture, store, manage and analyse rapidly increasing data (Manyika et al 2011:1; Khan et al 2010:7). This situation has continued because most organisations lack appropriate technology, data management systems, and BI solutions that can effectively assist them to make informed decisions (Russom 2013:5). The net result is that many organisations face growth limitations, inadequate BI computing and data management capabilities and often cannot realise the full potential of using their business data as the main information asset.

For organisations to keep up with this dynamic environment and, at the same time, remain competitive in an ecosystem of constant technological evolution, relevant interventions are required. One such intervention would be to methodically guide them throughout the implementation lifecycle of data management systems and BI applications. Such a contribution would go a long way in sustaining the competitiveness of the organisation in the sense that appropriate analytic solutions are available for real-time decision-making.

1.4 Purpose and objectives of the research

The purpose of this research study was to develop a data management and analytic model that can guide organisations in the implementation of BI applications. The model is needed to address current challenges faced by many organisations when managing large volumes of business data.

Based on the background to the research and the statement of the research problem, one main objective and four sub-objectives were derived to address the challenges which organisations face when procuring and implementing BI applications.

1.4.1 Research main objective

The main objective of this study was:

- ***To develop a data management and analytic model that can serve as a guide for BI applications implementation.***

1.4.2 Research Sub-objectives

The main research objective was achieved through the following sub-objectives:

- 1) ***Sub-Objective 1: To investigate and understand BI in general.*** For this sub-objective, an intensive literature review on BI was necessary. This was intended to glean information about the challenges and opportunities faced by organisations in deploying BI applications mostly from a business perspective. Increasingly, BI is becoming a driver aimed at enabling business to make real-time decisions so that they can remain competitive in the business environment. As such, achieving this sub-objective assisted in establishing a basis for exploring different technological possibilities in order to support organisations in their technological endeavour for effective and efficient deployment of BI applications.
- 2) ***Sub-Objective 2: To investigate and understand the various BI models and their associated advantages and limitations.*** A range of existing BI deployment models was investigated. In the process, limitations and drawbacks for each model under investigation were derived. The result of the research in regard to this sub-objective collectively contributed to the development of the proposed BI model aimed at limiting to some extent the challenges faced by organisations when relying on legacy deployment models.
- 3) ***Sub-Objective 3: To develop a BI model to overcome the existing BI deployment limitations.*** This sub-objective relied on the results of the

foregoing sub-objectives to contribute to the literature with a widely acceptable deployment model that could make up for existing shortcomings.

4) Sub-Objective 4: To assess the state of BI utilisations in organisations by business users and validate the proposed BI model by experts. The suggested model had to be tested theoretically in a real-life context at organisational level to ascertain its degree of acceptability and applicability. Furthermore, since the suggested model was conceptual in nature, for validation purpose, we relied on a survey targeting BI practitioners and experts so as to address shortcomings within our model and put forward an improved model that enjoys contributions from participating experts and in this way contribute to the body of knowledge on BI.

In order to achieve the main objective of this dissertation, an appropriate basic research question was derived from the main objective, of which a range of sub-research questions linked to each sub-objective were also derived as presented in the next sub-section.

1.5 Research questions

The following research questions were derived from the main research objective and sub-research objectives respectively.

1.5.1 Main research question

- ***What is an appropriate data management and analytic model that can serve as a blueprint for BI applications deployment for organisations?***

1.5.2 Sub-research questions

To answer the foregoing main research question, the following sub-research questions had to be answered:

- ***Sub-Research Question 1: What is the current state-of-the-art of BI usage in organisations?***
- ***Sub-Research Question 2: What are the most currently used models at organisation level and the limitations thereof?***

- ***Sub-Research Question 3: How can an improved model be developed taking into consideration existing limitations?***
- ***Sub-Research Question 4: How can the proposed model be evaluated and validated in a real-life context?***

1.6 Research outcomes

The main outcome of this research study is a data management and analytic model that organisations could use as a reference guide when implementing a BI solution. Its contribution consists of a hybrid BI deployment model and architecture. The model could be used as a guide when an organisation is planning to implement a robust BI solution and is uncertain as to where to start. Although the model is generic in nature, the architectural component is designed with key features that could be used to address key technological aspects needed for a successful BI solution implementation. The proposed architecture is independent of a technology platform giving organisations the opportunity to pilot several available proprietary and, or open sources technology options.

1.7 Research methodology

This section briefly explains the various data collection approaches and data analysis methods used to achieve the above-mentioned research objectives. Data collection approaches are explained in detail in chapter 6, and data analysis methods in chapters 7 and 8, respectively.

1.7.1 Data collection

Case studies which enabled the researcher to examine closely the data within specific industries were used in this research (Rugg & Petre 2007:64). The case study method ensured that the study focused on two specific types of respondents, namely, BI experts and BI users exposed to BI solutions. The first case study focused on BI users to assess the current state of BI utilisation in their organisations, and the second case

study focused on BI experts to evaluate and validate the proposed architectural model and any relevant improvements.

The research questionnaires with specific investigative questions were used to get sufficient information aimed at answering the research questions (Cooper & Schindler 2006:358). The questionnaires were the main research instrument used for data gathering in the survey conducted. The questionnaires were designed with close-ended questions with specific answers together with some questions requesting alternative answers in the respondent's own words (Kothari 2004:118). This helped to control the way in which the respondents provided answers. It also helped respondents to understand the meaning of each question to be answered. The use of questionnaires ensured that the administration costs were low, the resultant information was easy to quantify and maintain, and that there was a fair chance of obtaining large volumes of data (Cooper & Schindler 2006:254). Each questionnaire had specific questions that were linked to questionnaire objectives which they intended to achieve. The questionnaire's objectives were derived from the research sub-objectives.

Data gathering for this evidence-based research relied on measurable data collected through clearly defined survey questionnaires, that were easy to administer and easy to response; and with consideration of all crucial research requirements before data collection (Alasuutari 2010:149; Sibanda 2009:3; Kothari 2004:103). Information was collected in two parts. First, an extensive literature review was conducted on BI in general and on various deployment types and related Total Cost of Ownership (TCO), and then two case study surveys were administered to validate the findings from the literature review. To reach the relevant respondents residing at various geographic locations in South Africa, the questionnaires were administered through three options, namely, an email with a link to an online survey, an email with an attached soft copy (Word document) or hand delivery of a printed hard copy. Before engaging in the actual survey administration, a pre-test and pilot study was conducted to identify any flaws and weaknesses in the research questionnaire.

1.7.2 Validity and reliability

Validity and reliability concerns were addressed to ensure that the measuring instruments measure precisely what it is supposed to measure with accuracy and precision during the measurement procedure (Cooper & Schindler 2006:318).

Two major forms of validity were addressed, namely, internal and external validity (Cooper & Schindler 2006:318). Internal validity was addressed in terms of face validity, content validity, criterion validity and concurrent validity so that the research instrument measured what it is supposed to measure. Internal validity also ensured that the research design carefully followed the principle of cause and effect (Khalid et al. 2012:23). External validity has to do with whether the research results can be generalised across various characteristics of the subjects of the investigation, such as, people, industries and organisations (Cooper & Schindler 2006:318; Khalid et al 2012:23).

Reliability was used to define the degree to which a measure was free from random error and provided consistent results (Cooper & Schindler 2006:321; Khalid et al 2012:22; Kothari 2004:75). To maintain consistency and accuracy of the research instruments, reliability was expressed in terms of stability, equivalence and internal consistency.

1.7.3 Sample design

A survey population which is defined as the entire set of individuals or organisations from which survey data can be collected (Lind, Marchal & Wathen 2008:6), was used as a basis for this research. In order to collect the required data, industries in the South African context were considered and, in particular, 13 industries formed the survey population. A representative sample was selected carefully based on two units of measurement, namely, BI experts and BI users of the target population, so that a conclusion could be drawn about the entire population (Cooper & Schindler 2006:402). Sampling saved money and time, ensured minimum cost were incurred, ensured better and accurate results, ensured data was collected quicker and accessibility of the unit of measurement. A list of BI professionals from the sample constituted the sampling frame for the research (Kothari 2004:16). This was achieved using non-

probability sampling techniques, such as, purposive sampling and snowball sampling which ensured that the selected sample of the population constituted a representative sample of the whole population.

1.7.4 Data preparation and cleaning

The survey data for this research was stored, prepared and cleaned in Microsoft Excel. A few steps were followed to ensure that data was ready for analysis. Firstly, tables were created for coding and storing responses. Then, numerical values were assigned to all questionnaire items in the coding table. Thirdly, responses were entered in the storage table, and finally, to ensure data quality, data was cleaned where necessary.

1.7.5 Data analysis

Triangulation analysis ensured that a combination of literature review, quantitative analysis and qualitative analysis complemented each other to get two or three viewpoints of things being studied (Olsen 2004:4). Quantitative data analysis was used because much of the data collected from the survey was numerical while qualitative data analysis was used to analyse a small portion of data that was qualitative in nature. Since most of the data collected was quantitative, statistical data analysis in terms of descriptive statistics was used to categorise, summarise and present data in an explanatory way (Lind et al 2008:6). The data collected was relatively inexpensive to analyse and the results were used to generalise perceptions. More detailed information on data analysis methods is provided in chapter 7 and 8.

1.8 Preliminary literature review

The purpose of this preliminary literature review is to convey what knowledge and ideas have been established about data management systems and BI applications which are utilised to manage big data. This section briefly reviews the literature about the research problem by discussing the key aspects related to data management and BI applications.

Implementation of BI solutions is driven by clearly outlined mandating documents of the organisation, such as, strategy, goals, objectives, competition, commitment to profitability, shareholder value expansion, speed of delivery, and innovation (Ramamurthy, Sen & Sinha 2008:820; Khan et al 2010:3). Organisational commitment is the key element for an innovation, such as, implementation of BI solutions to be adopted and subsequently used. Organisational commitment, such as, support and participation from senior management and vision building at the senior management level is very important. It is imperative that when implementing a BI solution, the organisation understands and outlines its business objectives (Harriott 2013:38), and as pointed out by Kocakoç and Erdem (2010:172) the success of a BI project is dependent on having the right information and business requirements beforehand.

The goal of achieving success in implementing a fully-fledged and well-functioning BI solution has proved to be challenging. Thus, it is important to identify why many organisations fail even after making significant investments in BI solutions (Weber 2013:20). Gonzales (2011:25) suggests that most current measures or models do not assist in determining whether BI solutions will help achieve organisational goals and competitiveness. BI solutions are mostly needed by organisations which are aiming for better competitive position and are faced with the need to share and analyse business information (Chavan 2012:336). Managing and analysing quality information will lead to the generation of informative reports, dashboards, etc. that organisations can use in decision-making.

Nowadays, organisations can analyse and predict hurdles that they may face during major IT projects in advance and the benefit would be the elimination of a negative impact on the business (Khan et al. 2010:15). Organisations can only fulfil such predictions if they are using the correct BI and analytical applications. Analytical applications have the predictive capabilities that can simulate 'what if' scenarios, and the use of the correct applications can help organisations meet their strategic goals and retain competitiveness (Lopez & Antoni 2014:11). However, the use of such applications requires the careful consideration of key aspects and when misunderstood, as in most cases, it leads to the implementation of expensive converged BI and analytics solutions.

The need for organisations to improve their business processes constantly and the use of quality information for decision-making requires an investment in a suitable BI solution (Wixom, Watson & Werner 2011:68). There are various types of business value that are realised through an investment in a BI solution. This includes but is not limited to bottom-line saving, quality service, and the generation of revenue. It is, therefore, important that organisations use their business data in a manner that will ensure that business value is realised after investing in a particular BI solution. The procurement of an ideal BI solution is dependent on the type of information or model an organisation uses to make the correct procurement decision (Birst 2014:1; Ouf & Nasr 2012:650). There is limited information on data management and analytic models that can be used in a decision-making process and, or during implementation of a procured BI solution.

There are so many types of BI solutions that can be procured to fulfil different types of business requirements for organisations (TEC 2012:16; Summar 2012:1). Some BI solutions are procured through out-of-the-box cash transactions, others through some sort of subscription-based pricing agreement or some are developed in-house, and some are designed for local servers while others are cloud-based servers. Most organisations have invested in a costly information technology (IT) infrastructure to manage locally hosted BI solutions (Winkler & Brown 2013:20). This leads to high management and maintenance cost as most IT departments are overwhelmed with the large number of applications they have to manage (MicroStrategy 2012:7). To overcome such challenges, some organisations are opting for BI solutions hosted in the cloud. However, limited information is available to assist them implement such solutions with ease.

Mvelase, Dlodlo, Makitla, Sibiyi & Adigun (2009:215) argue that “cloud computing is a significant trend to increase agility and lower costs” as it provides the ability to adapt rapidly to changes in the business environment. Organisations considering the integration of cloud resources and BI solutions must have a clearly defined strategy to address all cloud computing capabilities and challenges (Marinela, Bogdan & Marian 2011:4). There is an unlimited capacity to expand in the cloud. This means that when planning for the initial capital investment, hardware expansions must be carefully thought over (Gash, Ariyachandra & Frolick 2011:264). When an organisation decides

to manage data in the cloud, data management applications are the best option for deployment in the cloud (Abadi 2009:1). Records show that most business users use descriptive analytics to manage data, and business analysts are helping them realise predictive analytics (Watson 2012:5).

Hobek, Ariyachandra and Frolick (2006:34) recommend that users must be trained after completing the implementation of the BI project as this will ensure that users are skilful and are able to use the new technology. The success of any locally hosted or cloud-based BI applications is user training. Effective training of the business user on BI applications will lead to effective management of big data.

The use of big data is becoming an important way to lead organisations to outperform their competitors (Manyika et al 2011:16). Organisations that want to sustain and operate at a strategic level by using BI and performance management must consider using proper data management systems and BI solutions (Hostmann 2007:7). BI solutions are often designed to enable organisations to remain competitive as they make use of informed decisions through the storing of quality data, and the analysis and production of informative reports (Karim 2011:197). However, Power (2008:153) maintains that most organisations are having problems retrieving historical data from their data stores due to an increasing amount of data coming from various sources. One of the solutions would be to implement a new BI solution that would consider locally hosted or in the cloud solutions but in order to do that, organisations need to use readily available information or a model as a guideline.

Dyché, Geiger, McGivern, McKnight, & Zeid (2011:44) recommend that the use of BI experts in a dedicated department, such as, a centre of excellence promotes effective collaboration and manipulation of business data. The utilisation of BI solutions is more effective if BI experts or business users have the ability to use available data and perform analysis that conform to, and support the business strategy. BI solutions encourages organisations to use various techniques, technologies, data systems, practices, methodologies, and applications to perform analysis on critical business data which results in better understanding of their business and market (Chen et al 2012:1166).

The discussion in the introduction and background and motivation sections revealed that there are a variety of business requirements that influence and challenge the deployment of a BI solution. This preliminary literature review further proposes that a further comprehensive study is required to establish what kind of a model could be used as a guide for the data management and successful implementation of BI applications. Formulating such a model could assist in achieving and meeting the main research objective and finding answers to the main research question. The findings of this study could also contribute to the limited body of knowledge on the use of BI solutions and whether it leads to better decision-making and business value realisation through managing and analysing big data.

1.9 Ethical considerations

Ethical clearance and considerations ensured that no respondent was abused or suffered adverse consequences from the research activities (Cooper & Schindler 2006:116). The following ethical considerations applied:

- Participants had the right to participate voluntarily and to be able to withdraw at any time without facing any penalties.
- A consent letter ensured that potential participants were fully aware of the procedure of the proposed survey, primarily the right to privacy, confidentiality, and anonymity.
- An ethical clearance was obtained from UNISA to ensure the legitimacy of the research being conducted.

All these ethical considerations ensured that the potential respondents do not suffer from physical harm, discomfort, pain, embarrassment, or loss of privacy when they participated in the research.

1.10 Significance of the study

The significance of this study is that:

- A data management and analytic model is of importance to enable organisations to make decisions from a BI perspective.

- Additional insights discussed throughout the study on BI, analytics, big data, on-premise and cloud-based BI applications, should contribute further to the enrichment of the current body of knowledge.
- Organisations seeking hybrid BI implementation would have base information on key concepts to consider before deciding on the potential solution to address their needs.
- Organisations would know beforehand about the value of managing data in a well-established BI deployment, and thus, be able to structure their projects accordingly.

1.11 Limitations of the study

The inadequacies of the study are not to be dismissed and it is imperative that the limitations of the study are clearly stated. The limitations of this study are discussed below:

- Due to the fact that BI is a broad subject, this research had limitations in respect of the particular research areas to be covered. The study only addressed issues relating to the defined research questions and research objectives.
- Time was a major factor to ensure that only the elements covered in the scope of the study were handled. Although data collection took a long time to complete in order to receive enough responses, it was difficult to obtain responses from more than 50 participants. This could have been due to the fact that people in the BI profession have a scarce skill, with very few people having acquired a BI qualifications or skills.
- Quantitative research methods often attract a large number of respondents but do not guarantee that it proceeds as it should. For example, one of the questionnaires was supposed to have been answered by those respondents who have a job title as BI expert, but the questionnaire was administered to participants who were performing a similar role with the job title of BI consultant, Business Architect, or Business Analyst, for instance. If all respondents were designated BI experts, this limitation could have been avoided.
- A limited amount of qualitative data was collected, analysed and used respectively. The qualitative research method in general, improves the efficiency of a research study (Kothari 2004:5).

1.12 Dissertation layout

This research study was presented in detail in the following chapters:

- *Chapter 1. Introduction* – covered the background to the research, statement of the research problem, objectives and research questions and outcomes.
- *Chapter 2. Business Intelligence Overview* – provides a general overview of BI adoption in organisations.
- *Chapter 3. On-premise Business Intelligence* – provides a specific investigation on the on-premise deployment of BI.
- *Chapter 4. Cloud-based Business Intelligence* – provides a specific investigation on the deployment of BI in the cloud.
- *Chapter 5. A Hybrid Deployment of Business Intelligence* – provides a specific investigation on the deployment of the proposed hybrid BI solution.
- *Chapter 6. Data Collection* – discusses data collection approaches adopted for the survey conducted in order to achieve the above-mentioned objectives.
- *Chapter 7. State of BI Deployment in Organisations* – presents and discusses the data analysis and the findings of part one of the survey conducted to determine the current state of BI utilisation in organisations.
- *Chapter 8. An Improved Hybrid BI Architecture* – presents and discusses the data analysis and the findings of the survey conducted on validating proposed BI architecture by experts. Furthermore, it discusses an improved hybrid architecture based on contributions from the experts.
- *Chapter 9. Summary and Conclusion* – summarises and concludes this research study conducted on a data management and analytic model for BI applications.

1.13 Conclusion

The purpose of this chapter was to introduce and explain the motivation for undertaking this research study. The chapter introduced some of the challenges and benefits of current BI applications. The background to the research covered the evolutions of data management and BI applications and related key aspects. These discussions informed the development of the statement of the research problem.

The purpose and the main objective and sub-objectives of the research were defined, followed by the main research question with sub-research questions and research outcomes. The data collection and data analysis methods used were also discussed highlighting key aspects for collecting, analysing and presenting research survey data. This was followed by the preliminary literature review, and definitions of the key concepts of the study. Ethical considerations and the significance and limitations of the study were also discussed followed by the layout of the dissertation.

Having achieved the purpose of this chapter, the next chapters will extensively review available literature to gain a general overview of BI and the adoption of BI in organisations in an effort to achieve research sub-objective 1.

CHAPTER 2 BUSINESS INTELLIGENCE OVERVIEW

2.1 Introduction

In the previous chapter, the motivation for embarking on this research study was presented and discussed and this informed the main research objective and question. The methodology, preliminary literature review and other critical aspects that would ensure the success of this research study were discussed.

The purpose of this chapter is to provide a general overview of BI, its use and the associated business value that could be generated by BI adoption in organisations. It covers in detail the key concepts and terminologies that are needed and, or commonly used in order to understand and implement BI successfully. As such, this chapter provides an answer to our first sub-research question:

“What is the current state-of-the-art of BI usage in organisations?”

There are so many changes taking place in the BI industry, and it is often hard to keep up-to-date. In order to get a comprehensive understanding of BI, this chapter will describe BI in detail using some of the definitions in the BI literature and explore the main features and functions that make it possible to utilise BI for better decision-making.

The chapter is outlined as follows: section 2.2 simplifies the understanding of the BI concept through various discussions on its definitions, followed by section 2.3 which explores the BI value chain. This is followed by section 2.4 which presents a naïve model for representing a BI problem. The model informed the development of an ideal architecture of BI in section 2.5. Also covered in this chapter are the application of OLTP and OLAP in relation to BI in sections 2.6 and 2.7 respectively. The total cost of ownership of BI is discussed in section 2.8 and the chapter concludes with a summary discussion of the chapter in section 2.9.

2.2 Definition of business intelligence

There are various definitions of BI in the literature of which some are explored along the following lines.

Manyika et al (2011:31) defines BI as “a type of application software that is designed to report, analyse, and present data”. Similarly, Bucur (2012:96) expands on this view describing BI as the “existence of operational databases, systems of decision assistance, and applications platforms that facilitate access to economic information to persons involved in business management”. These definitions explain that the process of accessing, analysing and reporting on business data is achieved through the use of BI software. This all happens through reading data that has been cleansed and stored in a central location and, or data is stored in a sub-set of data, better known as a ‘data warehouse’ and a ‘data mart’ respectively. A data warehouse and data mart are important components of the BI system and will be discussed in detail later in this chapter.

In another context BI has been defined as a “black box that transforms the raw data from data source to useful knowledge and allows the user to view business performance and hence improve decision-making” (Chuah & Wong 2013:1). The black box may collect information and, or be used to gather data or, “provide access to, and analyse data and information about the company operations” (Chavan 2012:335). From these definitions, the understanding is that BI involves acquisition and analysis of business data to support decisions in an organisation. Nowadays, it is important that organisations have the right answers readily available to support any decision-making process that aims at achieving their business objectives.

Furthermore, BI referred to as the “skills, technologies, applications and practices used to help a business acquire a better understanding of its commercial context” (Chavan 2012:335). This implies that users need to be equipped with the relevant skills required for Statistical Analysis System (SAS) developers in order to use SAS BI Suite to derive information that management can use to make informed decisions. And “the methods and technologies that gather, store, report, and analyse business data to help people make business decisions” are often referred to as BI tools (*Dictionary.com* 2014).

BI has also been used as an universal term to define the “technologies, processes, and applications for supporting decision-making” (Watson 2011:5). Gartner (2013) defines BI as an “umbrella term that includes the applications, infrastructure and tools, and best practices that enables access to, and analysis of information to improve and optimise decisions and performance”. This means that effective decision-making is

achieved when the appropriate technology, such as, data cleansing tools are used to ensure that the right data is being utilised for BI purposes. The right processes need to be followed to ensure effective data governance and consistent data gathering. Once the level of acceptable data quality has been reached, front-end reporting applications or tools, such as, dashboards, scorecards, and OLAP, need to be utilised to achieve and support better decision-making.

From the discussion above, it is understood that BI is the process of analysing business data stored within an organisation to determine patterns and trends that support better decision-making for the organisation. Some of the common themes that have been picked up from the various definitions discussed are: sourcing data, analysing stored data (i.e. historical, or real-time) in order to present information that business users (executives, business analysts, operational users, etc.) can use to make informed decisions.

Therefore, in the context of this study BI is defined as an optimal and effective collection of business data from various sources, which is then transformed through intelligent applications in order to understand information, aimed at supporting decision-making that helps organisations achieve their objectives. There are several BI applications on the market, such as, SAS¹ BI, SAP BI², and Cognos³ that have the capabilities of transforming business data into intelligence information that allows organisations to make informed decisions competitively.

In order to further describe and have a clear understanding of BI in the context of this work, the next section will explore the BI value chain.

¹ Statistical Analysis System (SAS) BI & Analytics software provides the most comprehensive, integrated and easy-to-use reporting and analysis features.

² Systems, Applications, & Products (SAP) BusinessObjects BI platform provides flexible systems management for an enterprise BI standard that allows administrators to confidently deploy and standardise their BI implementations on a proven, scalable, and adaptive service-oriented architecture.

³ International Business Machines Corporation (IBM) Cognos BI turns data into past, present and future views of your organisation's operations and performance so that your decision-makers can capitalise on opportunities and minimise risks.

2.3 Business intelligence value chain

The fact remains that more and more data will be generated by organisations, but its use to support decision-making and generate value still continues to be uncertain (Dupada, Gedela, Aryasri, Acharya 2013:24). Michael Porter introduced the concept of a value chain in 1985 in his classic book “Competitive Advantage”. A value chain is a collection of mutually supporting activities that an organisation perform to create value for its customers (Porter & Millar 1985:150). In the context of this study, a BI value chain comprises ideal value activities and components needed to capture, manipulate and channel the data necessary to support better decision-making. The value activities and components of the BI value chain are data, processes, analytics, and decisions. This study’s BI value chain explores how BI as a tool is transforming the way value activities are performed and the nature of their interdependencies in order to address specific business problems and support organisations in making appropriate decisions with respect to the problem at hand.

The idea of using a BI value chain in this study, implies that all value activities and components and their interdependencies will support and add value to the process of decision-making (Brohman, Parent, Pearce, & Wade 2000:4). Figure 2-1 below, depicts the perceived BI value chain comprised of a set of interrelated components aimed at supporting effective and efficient business decisions for an organisations, value creation.

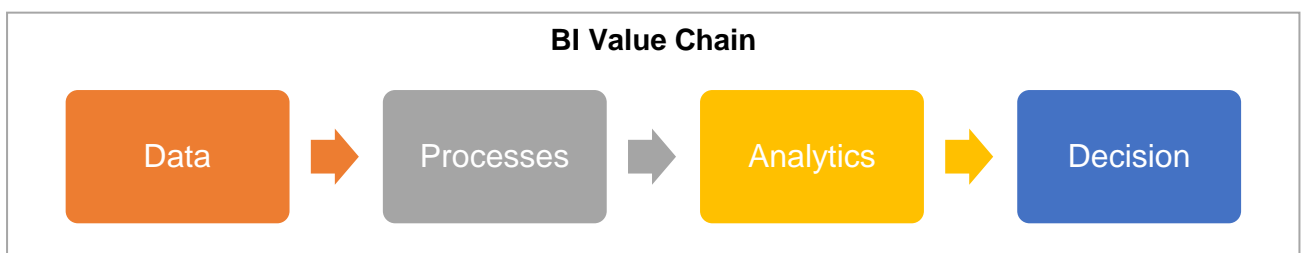


Figure 2-1: Business Intelligence Value Chain

Before discussing the BI value chain components in detail, it is important to explore and define the value realised from the BI value chain process and the beneficiaries or types of business users. The need for organisations to know what happened, why it happened and what will happen has increased the number of organisations that have an interest in investing in BI and analytics solutions. This comes with the view that

organisational data will become a key asset in leading organisations to outperform their competitors (Manyika et al 2011:16). Organisations only realise value when they can make better informed decisions from their business data or information.

There are several ways of defining business value emanating from an investment in BI solutions. Williams and Williams (2007:13) define business value in “economic terms as the net present value of the after-tax cash flows associated with the investment (an asset)”. They further assert that this means that the business value of BI investment depends on its use within management processes (planning, controlling, measuring, monitoring, etc.) that influence operational processes (fraud detection, sales campaign execution, customer order processing, purchasing, accounts payable processing, etc.) that drive revenue, quality services or reduced costs, as well as the use of BI within operational processes.

Melville, Kraemer & Gurbaxani (2004:287) describe business value as the efficiency and competitive impacts on the organisational performance as a result of the optimal utilisation of BI at both the in-between process level and the enterprise-wide level. The investment in BI not only produces business value as bottom-line cost savings and revenue generation, but also as better-quality business processes and enhanced decision-making (Wixom et al 2011:68).

It can be deduced from the above-mentioned definitions that there is more to realising business value than just the technical implementation of a BI solution. This ranges from effective processes (improved business processes), change management, information quality, improved business performance, and so forth (Popovi, Turk & Jaklic 2010:16; Williams & Williams 2003:3). Value realisation will be worthless if organisations do not fully understand who (in terms of the business users) in the organisation is responsible for creating value and who will consume it.

Business users in the context of this study are categorised into two types based on skills, that is, those that perform the much-needed analysis and those that consume the resultant quality information. On the one hand, analytical tasks are performed by power-users, such as, business analysts, advanced query developers, and data scientists who drive the adoption of BI solutions and, at the same time, identify the requirements, opportunities and problems that BI will resolve (Weber 2013:23). On the

other hand, end-users or information consumers are business users who are able to make strategic, tactical and operational decisions, ranging from executives, senior managers, and managers, down to basic end-users (Williams & Williams 2007:80).

The foregoing description and definition of the notion of business value (and, therefore, value chain) justifies the strategic importance of BI for organisations. It clearly shows that not only does it assist decision-makers in taking informed decisions for the benefit of the organisation, it also serves as an important instrument for driving optimisation based on efficient interaction between components that makeup the overall value chain. In the next subsections, each of those components is discussed in detail.

2.3.1 Business data

It is understood from the previous sections that business data is the most important part of any BI implementation. Business data are facts, datasets or pieces of information which are collected, generated and stored for analysis, reasoning and reporting purposes by organisation systems (Manyika et al 2011:1; Williams & Williams 2007:201). Business data examples include sales information, billing, customer names or types, categories, statistics; customers' buying patterns, employees' profiles, devices information, and so forth. Organisations have source systems that process operational data, and these systems come in different types, such as, transaction processing systems, operational systems, database management systems, OLTP, etc. (Plattener & Zeier 2012:27). These systems are discussed in detail in subsequent sections.

Source systems are systems that process and provide the much needed data for BI either through scheduled intervals, or in real-time, and update data as, and when needed and are usually made up of systems, such as, Enterprise Resource Planning (ERP), Supply Chain Management (SCM), and Customer Relationship Management (CRM) (Ong et al 2011:4). Data generated from source systems can be used to perform any required analysis and yield results which can assist organisations in making informed strategic objectives, business models, or learning about market trends and risks (Jiang, Barone, Amyot & Mylopoulos 2011:429). An organisation's data is crucial as it forms the basis for performing analytical functions.

It is crucial that the data being used for performing analytical functions is clean and validated using business rules and cleansing tools (Davenport & Harris 2007:7). Furthermore, business rules are conditional statements or sets of rules used to address logical questions in order to obtain consistent formats for analysis and reporting. For instance, one way to implement a rule would be to have a four-number postal code for all South African physical and postal addresses. On the other hand, data cleansing tools, such as, InfoSphere⁴, DataFlux⁵, DataCleaner⁶, and Trillium⁷ use a specialised technique also known as the Extract, Transform and Load (ETL) process to correct and standardise data (Manyika et al 2011:32). This process of ensuring that the correct (cleansed) data is stored in the correct repository is discussed in detail in the later sections.

2.3.2 Processes

Business processes are specific tasks or step-by-step activities that are carried out in order to produce the intended outcomes for business users. These processes are often automated and optimised. For example, to ensure that they convert data into usable information enabling organisations to make fast data-driven decisions (Kolb 2012:4). The process of importing data from the sources systems and preparing it for analysis ensures that data is controlled and consistent (Olszak & Ziemia 2007:144). Processes in the BI value chain are all automated data processes that prepare data for analysis. Data processes are often triggered by data coming from day-to-day business activities which will need to be analysed and generate reports for further business-related decisions and actions.

It is important that data is properly organised before it is presented to the business user, particularly if the data emanates from different source systems (Plattener & Zeier

⁴ IBM InfoSphere Information Server is a market-leading data integration platform which includes a family of products that enable you to understand, cleanse, monitor, transform, and deliver data, as well as to collaborate to bridge the gap between business and IT.

⁵ DataFlux Data Management Studio is a data management suite that combines data quality, data integration and master data management.

⁶ The heart of DataCleaner is a strong data profiling engine for discovering and analysing the quality of your data. It finds the patterns, missing values, character sets and other characteristics of the data values.

⁷ Trillium Software's data quality products certify that the most critical data assets are accurate, complete and aligned with your business goals.

2012:4). Most of the data preparation for analytics is applied to only a portion of the whole organisational database, limiting the data granularity of possible reports. This can sometimes be time-consuming and costly depending on the technology being used as it may take several hours or even days (Sen et al 2012:336). For example, processing 1TB of transactional data can also lead to performance issues of the entire system. Hence, most organisations have separate systems, that is, BI and transaction processing applications.

The process of importing, cleansing and storing, commonly referred to as extracting, transforming and loading of the data, will be discussed in much detail in the Data Layer section. However, once the data is ready for use, analytical functions can then be executed as discussed in the next section and possibly meet the intended outcomes.

2.3.3 Analytical function

Once the right business data is in place and the correct processes have been followed, analytics can start. Analytics are functions carried out by BI tools and applications that turn raw data into information so that it can inform and support decision-making as well as drive informed actions (Kolb 2012:3). For example, instead of simply knowing where the trend is headed, analytics tell you what you need to know to influence the trend in the direction you desire. Analytics tools, such as, Customer Analytics⁸, Web Analytics⁹ (Google Analytics), Siebel Analytics¹⁰ (Oracle BI application), Microsoft Excel, Enterprise Resource Planning, and Social Media Analytics¹¹, are used by organisations to offer data as an asset by finding patterns and data trends through data mining. It anticipates the future outcomes through predictive modelling,

⁸ Customer Analytics enables businesses to take information gained about consumers gathered either internally or externally and use predictive modelling in order to gain useful insight about customer behaviour.

⁹ Web Analytics enables organisations to increase their knowledge about the website's performance in a wide range of areas.

¹⁰ Siebel Analytics helps organisations obtain, view, and analyse their large amount of data about products, customers, prices, contacts, activities, assets, opportunities, employees, etc.

¹¹ Social Media Analytics delivers key insights that can be used to understand the potential lifetime value of customers from their communication and collaborations on social media sites, such as, Facebook, LinkedIn, blogs, etc.

forecasting, and so forth. Different types of analytic tools are discussed in detail in the subsequent sections.

To generate useful information from business data, analytical functions are often executed through vendor specific applications like the few mentioned above. In most cases, business data is used for analysing and reporting so that operational performance can be measured as well as measure the execution of the organisation's strategy and objectives (Barone et al 2012:504). It all starts with finding the relevant business data and discovering the best parts of it to answer the questions raised. Most organisations are using specific analytical tools because of the ease of use features; the power to generate reports with minimal technical skill and deliver the reports quickly, which enables the business users to make highly informed decisions (Kolb 2012:6).

Analytical functions are executed through easy to use tools, such as, Microsoft Excel prior to presenting data in reports or dashboards for key decision-makers in the organisation (Davenport & Harris 2007:9). Analytical functionality is also executed through advanced analytical tools for reporting with capabilities, such as, data mining, predictive and prescriptive analysis, multivariate testing, pivoting, quick response analysis, data visualisation, and infographics (MicroStrategy 2012:22; Kolb 2012:4). Advanced analytics, such as, Siebel Analytics have the capabilities of letting the organisation know how to do a better job of, for example, retaining customers, and attracting new customers through effective and efficient utilisation of information (Schauer 2004:28). In the current competitive business environment, the possession of the right information is crucial for decision-making and, therefore, the survival of the business.

2.3.4 Decision

The previous stages in the BI value chain have shown the importance of collecting appropriate data, and the need to follow the correct processes to prepare data for the intended analysis. Then follows the crucial stage of making sense of the information presented to the business user in order to influence the course of the challenge at hand referred to as decision-making (*BusinessDictionary.com* 2015). This thought

process ensures that organisations have the most effective ways to decision-making based on the available information (Kruger 2010:36). The availability of correct information also ensures that the majority of decision-makers are empowered with data-driven business decisions which are based on dynamic BI reports or queries that allow slicing and dicing of data, so that the information can answer difficult business questions (Weber 2013:19).

Thus, high-quality information (and, therefore, powerful results emanating from analytics) is one of the most important elements in the BI value chain. This ensures ease of use by the business users when it comes to making better data-driven informed decisions (Power 2008:152). Decision-making is fundamental as it leads the business users to generate well-timed plans of action based on intelligence gathered, as well as to respond to emerging threats and opportunities (Chavan 2012:340). The ability to make the right decision immediately when presented with information is crucial in this dynamic environment.

Therefore, it is important that if organisations are to make informed decisions the right data must be acquired, and efficient and easy to use analytical tools must be in place. This will ensure the realisation of business value from implementing a BI solution, as business users will make sound decisions that have an impact on the organisation's performance.

In the next section, a 'naïve' or simple BI model is discussed that is aimed at supporting the BI value chain discussed in this section and can form the basis for its operationalisation from a technical perspective.

2.4 Naïve business intelligence model

In support of the earlier discussion on the BI value chain, in this section a naïve BI model is suggested, which forms the basis for the technical implementation of a simple BI solution. The intention of a naïve BI model is to demonstrate how a BI solution can be easily implemented based on its original value chain. Although the model has been intentionally termed a 'naïve' one, it should be noted that the derivation of a technical architecture from such a model could yield a sound BI solution capable of supporting decision-making within an organisation with little complexity. The naïve model will

demonstrate in its simplest form how other elements contribute to the process of supporting better decision-making. The model as depicted in figure 2-2 was derived by aligning each component of the value chain with an appropriate technology for facilitating the realisation of such components. Therefore, a data source system is required for the materialisation of the data component of the value chain. Once it is collected it is subjected to a range of processes for data cleansing before it is stored. By using the appropriate tools, data analysis can be performed of which the outcome is appropriate information to support decision-making. A discussion of each component of the model follows:

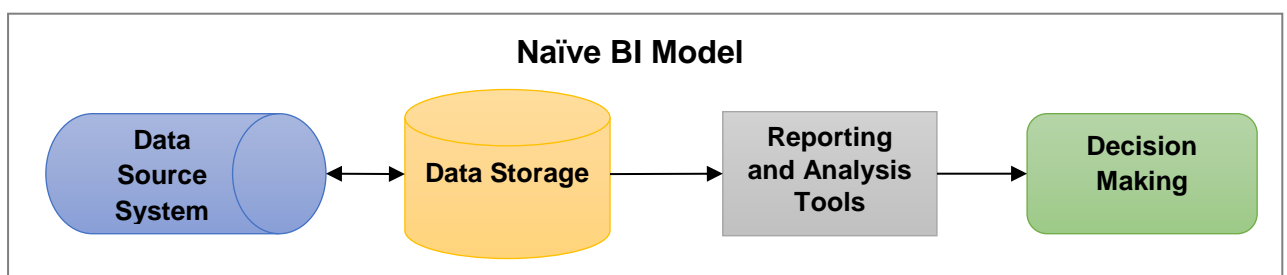


Figure 2-2: Naïve BI Model

2.4.1 Data source systems

Previous discussions revealed that BI depends on data that must be collected or produced from various sources. Data sources systems discussed earlier, such as, ERP, SCM, and CRM, are transaction systems responsible for gathering and generating business data needed in the BI value chain (Zhang & Tu 2009:182). Data source systems in the naïve BI model come in two types: internal and external systems. Internal transaction systems are responsible for generating business data within the organisation and are also known as OLTP (Ong et al 2011:4; Plattner & Zeier 2012:27). Furthermore, external sources systems are those that transmit data into an organisation from outside, such as, the Internet, a business partner or the government, enterprise applications, web services, application servers, content management systems, and the like.

Data source systems are crucial in the sense that without data, BI tools become unusable, and it is important to clearly identify data sources in order to address specific business questions and requirements (Qionqwei, Yaotang & Qiuyun 2012:909; Ong et al 2011:4). It is important that source systems generate the right quality of data that

can be used in the BI value chain activities, and be stored for future use for the purpose of understanding the organisation's performance, supporting decision-making, and competitive analysis.

2.4.2 Data storage

Once business data has been generated, it needs to be stored where it is accessible as, and when needed (Lopez & Antoni 2014:12). The naïve model shown in figure 2-2 above includes a simple data storage component that acts as a database and stores data emanating from the various source systems which are later used to perform analytic functions (Bucur 2012:102). The data stores come in different types and sizes, such as, data warehouse, operational data stores, and data marts, depending on the technology being used (Pant 2009:4). Further details are provided in the Data Layer section 2.5.3.

The data storage component is the central repository of data from where data can be accessed by business users in order to perform any necessary reporting and analysis (Chaudhuri, Dayal & Narasayya 2011:90). It has the capability of storing historical data and current data in order to meet business needs (Ong et al 2011:5). The value chain suggests that once data has been processed, it becomes available for performing analytic functions. This can only be achieved if data is stored within the organisation or somewhere else in a remote location, such as the cloud. Organisations that keep their data in the locally hosted servers often feel comfortable and secure with the integrity of the data as the ownership lies with them (Lonnqvist & Pirttimaki 2006:38). On the other hand, if the data is stored in the cloud, organisations may feel that their business data may be compromised.

Therefore, the crucial element in the data storage element is that once data has been stored using the relevant technology, it must be accessible as, and when required. The reporting and analysis tools discussed in the next section rely heavily on the availability of sound data in order to produce the necessary outputs. For example, business users may want to pull a previously generated report on the most popular items that sell most and compare it to the current sales figures. This can only be

achieved by gaining access to stored data and the capability to perform the necessary analytic functions.

2.4.3 Reporting and analysis tools

The naïve BI model has demonstrated the need to store accessible data so that the analytical functions in the BI value chain can be executed to achieve the intended outcomes. The reporting and analysis tools are BI applications which are designed to retrieve, analyse and produce reports based on data in the data storage component (Plattener & Zeier 2012:198). In general, these tools range from simple to advanced tools, such as, Microsoft Excel Spreadsheets, Google Sheets, data mining, querying and reporting applications, and dashboards. The task is to keep the correct data so that reliable reports for decision-making can be generated, to avoid a situation of 'garbage in garbage out' (Safeer & Zafar 2011:268).

It is crucial that data stored to be used with the reporting and analysis tools has the best ever quality as possible, because the output is critical for decision-making (Hazen et al. 2014:73; Kwon et al 2014:389; Ong et al 2011:4). The reporting and analysis tools have the capacity to read data with the intention of applying the necessary intelligence to it and generating meaningful output on which business decisions can be based. Furthermore, it is often easier to perform analytic functions when using the right reporting and analysis tools as the output for these BI value chain activities are crucial for supporting the decision-making process.

2.4.4 Decision-making

As discussed earlier in the BI value chain, the ultimate goal of implementing BI solutions is so that better decisions can be made based on the information coming from performing analytic functions through the use of reporting and analysis tools. Hence, improved decision-making which is made based on reports generated from business data is a sign of the correct deployment of a BI solution (Moghimi & Zheng 2009:77). These decisions are important in an organisation as they assist business users in using the information at their disposal to implement the necessary actions and to improve operational efficiency.

These days, business users have the opportunity to determine how the organisation is performing and where it is heading through the use of self-service BI. Self-service BI removes the challenges of having speedy insight and decision-making by cutting off the labour cost and the need to use the middleman (e.g. IT personnel), as data is often accessible to the business user (Weber 2013:19). Furthermore, the ability to make prompt decisions contributes to gaining competitive advantage and attaining consistent returns. It is understandable that in future this will be the way business users need to operate with less reliance on IT or BI developers to generate needed data-driven reports for decision-making.

In the next section a BI architecture derived from the proposed naïve model is discussed.

2.5 Business intelligence architecture

A BI architecture is the complete design of a BI system which demonstrate the interrelationships among its components (Gartner 2015a). It acts as the foundation for any BI initiative for delivering business information that will assist improved business performance. In the context of this study, the BI architecture in figure 2-3 demonstrates the understanding of the general BI and how to implement the earlier discussed BI value chain and naïve BI model. It demonstrates the dynamic components related to data source systems, the storage capabilities, data analysis services, reporting and delivery (Popovi et al 2010:9). The architecture needs to be more seamlessly integrated and allow for more re-use of its components (Evelson 2014:11).

The subject literature on BI refers to, or reports on, or offers various theories on, or puts forward several BI architectures with different structures or layers, but all have common components, such as, reporting tools, data storage and sources systems (Ong et al 2011:2). Considering the discussion on the BI value chain and the naïve BI model, an architecture is suggested consisting of four (4) main layers namely: the presentation layer, the service, the data and source systems layers. Figure 2-3 below depicts the overall architecture of which each layer (and components therein) is described in the next subsections.

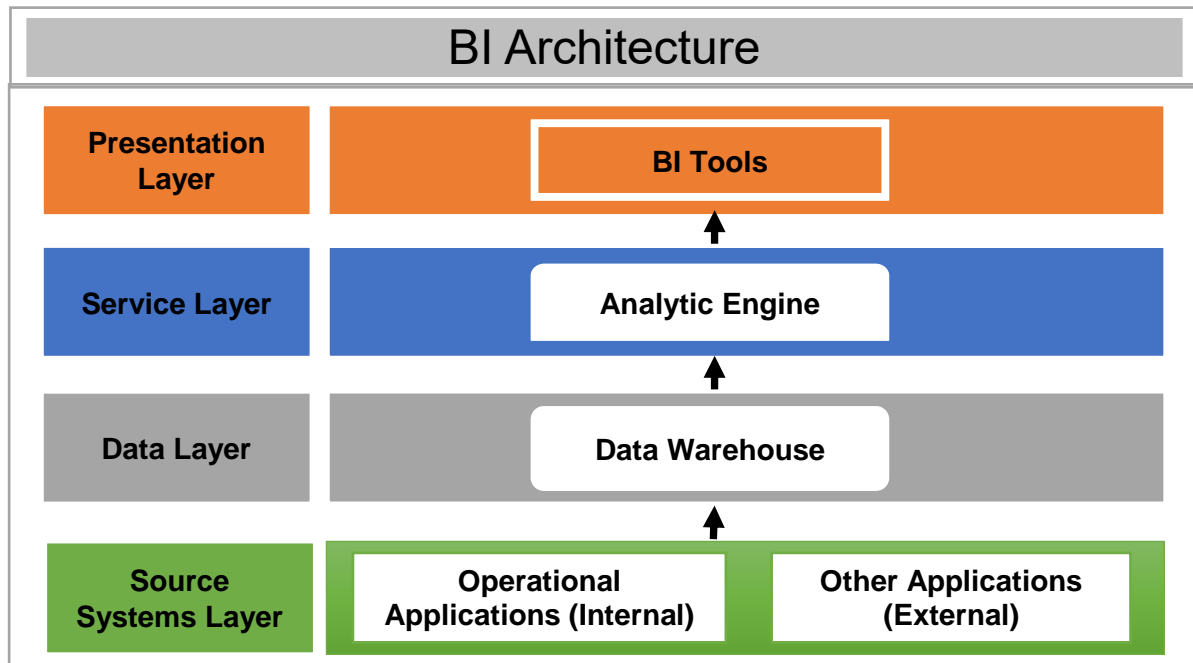


Figure 2-3: BI Architecture

2.5.1 Presentation Layer

The presentation layer acts as the front end in the BI implementation from where the kinds of business users discussed earlier, such as, power-users (business analysts, data scientists, etc.) and end-users (executives, senior managers, etc.) can access the data in order to report on the organisation with reporting BI tools (Weber 2013:23; Williams & Williams 2007:80). In other ways, this layer provides access to the use of all relevant analytic engines found in the service layer in an integrated manner (Ouf & Nasr 2012:655). In this layer, business users have a single point of entry to all analytic functions and applications which were discussed in the BI value chain and the naïve BI model. The access to this single-entry point can be through various types of communication media, such as, web browsers, desktops, mobile applications, online portals, and so forth (Plattener & Zeier 2012:180; Asghar, Fong & Hussain 2009:676). Gollapudi, Jangeti & Kotapati (2012:408) illustrate this process further with a specific web browser-based user interface which uses the IBM Cognos¹² Platform to handle all interactions and deliver BI capabilities.

¹² The IBM Cognos platform is an open, enterprise-class platform that can provide users with the BI and performance management content they need to understand the business and help improve results.

The presentation layer hosts a number of BI reporting tools that range from dashboards, scorecards, and reports, to mobile applications and alerts (MicroStrategy 2012:12). MicroStrategy Report Services as shown in the diagram below (figure 2-4) illustrates a key performance indicators' dashboard for comparing current performance against established targets in order to identify opportunities or issues. SAP Business Explorer Suite (BEx Suite) offers similar functionality. The benefit of using such reporting tools is that they have the ability to explore the data without the possibility of unintentionally changing the fundamental model. Furthermore, these reporting tools allow business users to create unplanned reports and to be informed about exceptions through communication tools such as text messages, email, etc.



Figure 2-4: Corporate Sales Overview (sourced from MicroStrategy 2012:369)

2.5.2 Service layer

The service layer provides users with the analytic engines that allow them to perform any required analysis. The analytic engines come with various analysis capabilities for

managing multiple data types which include visual OLAP¹³, slice and dice¹⁴ analysis, data mining, predictive analysis, etc (Dyché et al 2011:10; Davis, Imhoff & White 2009:9). The components in this layer are important as they enable the users to plan and forecast as they know what has happened, why it happened, and what will happen should the situation change or remain the same. Following the earlier brief introduction and discussion on the BI value chain, the following are some of the most common analytical capabilities and functions utilised by analytic engines:

- **Data mining** is a method that looks for new, valued and noticeable information in big data by applying classification, association and forecasting algorithms and enables optimal utilisation of descriptive, prescriptive and predictive analytics (Kocakoç & Erdem 2010:176). Descriptive analytics has the ability to slice, dice or drill down historical data to yield descriptive insight for business users to know what happened in a particular situation (Barnes et al 2012:1). In order to know about, or answer questions as to what will happen in the future?, predictive analytics are used which allow business users to identify meaningful patterns from historical and potential future big data, in order to make time-sensitive decisions to predict specific events (Evelson 2013:6; Barnes et al 2012:21). Furthermore, in order to answer 'what if' scenarios, prescriptive analytics are utilised. These analytics have the capabilities of building 'what if' scenarios to explain why things happen in a particular way and allows business users to plan and propose changes that will optimise performance (Watson 2012:5; TEC 2011:66).
- **Text analytics engines** have the capacity to analyse large amounts of text (unstructured) data for the purpose of extracting useful information that usually need manual intervention, such as, customer feedback comments that would come from customer surveys (Chaudhuri et al 2011:3). The other analytics engine that handles unstructured web intelligence content is Web analytics, such as, Google Analytics. Web analytics enables the tracking of users' online activities emanating from social media sites, and online transactional sites by

¹³ OLAP systems are used to report on the captured data once the ETL processes and data is residing in the data warehouse

¹⁴ To 'slice and dice' implies a systematic reduction of a body of data into smaller parts or views that will yield more information and present it in a variety of different and useful ways so that business users can understand it better.

providing specific site interaction details, such as, browsing history (knowing the frequency of visits to specific web pages), customer transaction analysis, and product recommendations (Chen et al 2012:1167; Chaudhuri et al 2011:90).

- **Diagnostic analytics engines** are process-level performance measurement tools that inform business users (management) as to why organisational performance is, or is not improving, or activities are not performing effectively and, therefore, have a significant impact on the expected organisational benefits (Elbashir, Collier, Sutton, Davern, & Leech 2008:150).
- **Mobile analytics** are the emerging and most popular analytics engines that are capable of performing operational intelligence and location intelligence on smart devices through the collection of location-specific and personalised content (TEC 2014:11; Chen et al 2012:1178). It is also viewed as a better way of ensuring that the organisations' business users remain productive and efficient.

The above-mentioned analytic functions have the capacity to mine data (structured or unstructured) with the view to analysing it for unique patterns on customer needs, business performance improvement and realising new business opportunities.

2.5.3 Data layer

The data layer presents the data warehouse, one of the crucial components in the BI architecture. Data warehousing is a process of gathering all organisational data from multiple data source systems into a structured central repository for the purpose of optimal utilisation in decision-making processes (Boateng et al 2012:2; Manyika et al 2011:32). The repository stores data that has been cleansed for reporting and analysis purposes in order to support managerial decision-making (Sen et al 2012:336).

The data cleansing process is performed by ETL software tools. ETL tools extract data from multiple source operational systems (internal or external), and then transform data by cleansing and reformatting it to fit operational needs and eventually loading the data into a targeted data warehouse (Manyika et al 2011:32).

Williams and Williams (2007:126), Boateng et al (2012:227) and Chavan (2012:341) offer a further explanation in regard to ETL and assert that the data extraction process has the capability of automatically capturing metadata or reading data from an infinite number of unstructured or poorly structured data sources. The transformation process is initiated to make sure that the data entering into the data warehouse is in the right format for the intended BI applications. It is common knowledge that data entered by users into internal systems or from external transactional systems is prone to errors or incomplete entries, and the transformation process ensures that there are no 'garbage-in, garbage-out' scenarios. The data loading process ensures that data loaded into the data warehouse adhere to the business rules by keeping track of changes and formats of data.

Upon completion of the ETL processes, the data layer will be ready and enable users to use BI tools to perform extensive analysis on large volumes of data in real-time through three commonly used components, that is, data warehouse, operational data store (ODS) and data mart (Boateng et al 2012:223). The data may first be stored in an ODS which addresses operational data needs by providing real or near real-time information direct to users for them to report and respond instantly to changing business environments and requirements (Ouf & Nasr 2012:655; Ong et al 2011:4). Furthermore, information in the ODS is constantly updated with new data ensuring that operational processing and reporting needs for middle management are supported. This essentially means that ODS does not keep historical data but holds the most recent data before it is migrated to the data warehouse (Boateng et al 2012:223).

On the other hand, data marts are designed particularly to support the analytical business needs of specific business units which often comes as a sub-set of a data warehouse (Boateng et al 2012:223). However, the use of additional physical data marts or an ODS increases the project and on-going costs (Bhatti 2013:29).

In other ways, a data warehouse stores data for strategic analysis based on historical data while ODS store data for operational analysis based on current data. Data marts store historical data but not for a long time as the data warehouse does.

2.5.4 Source systems layer

The source systems layer is often comprised of operational and transactional IT applications that are either internal or external and are responsible for supplying structured or unstructured data to the data layer. Source systems are responsible for capturing and maintaining operational data within the organisation (Ong et al 2011:4). Source systems are usually a variety of IT applications from either a specific vendor or different vendors all with the purpose of enabling the fulfilment of specific functions in the organisation.

These IT applications are often referred to as OLTP systems which include ERP, SCM, CRM, Sales and Marketing, and so on. These IT applications may store their respective data in different types of systems, such as, Relational Database Management Systems (RDMS), XML, CSV file, MS Excel (XLS) file, and so on (Chavan 2012:341). The ETL tools play a major role in data integration by reading the data from these different source systems, transforming the data into the required format and eventually migrating the data into the data layer (Pant 2009:15).

In summary, BI architecture depends on many critical factors needed for easier, faster and informed decision-making. These factors include strategic, tactical and operational needs, type of analytical functions required, the rate at which a decision needs to be made, the type of users, data types and how long it needs to be stored, and so on.

2.6 Application of on-line transaction processing

The term 'on-line transaction processing' (OLTP) is often used to refer to transaction processing systems rather than analytical systems as they reside in the source systems layer of the BI Architecture. OLTP are operational systems because they have the capabilities of processing operational data simultaneously as well as updating data continuously (Ong et al 2011:4; Williams & Williams 2007:201). For instance, they record a bank customer's on-line transaction of paying a supplier using a smart phone app.

In order to handle large data entry volumes and speed up the processes of inserting, updating and deleting, OLTP systems are highly normalised (SAP 2012:5; Plattner & Zeier 2012:124). The large data volumes emanate from various OLTP source systems and other data sources and it is loaded into the data warehouse through ETL processes (Sen et al 2012:336). The ultimate goal of OLTP transactions is the flawless execution of applications within the organisation as they form the basis for most operations. Although OLTP are more focused on transactions, they play a key role in providing the much-needed data for the execution of analytical functions in real time.

2.7 Application of on-line analytical processing

On-line analytical processing (OLAP) unlike OLTP resides in the service layer of the BI Architecture. OLAP has capabilities of executing analytical functions on data found (captured using OLTP) in the data warehouse, data mart or ODS in order to yield actionable BI (Williams & Williams 2007:201).

OLAP is a user-friendly graphical tool and an essential component in the BI architecture because of its ability to perform extensive analysis (exploratory, time series, multi-dimensional cubes, etc.) on the data and provide information for making informed decisions (Bucur 2012:103; Ong et al 2011:7). It is only through OLAP that organisations are able to understand their business, and transform their operations to fit in with the changing environment (Plattner 2009:1).

The organisation's OLAP capabilities are often significantly enhanced when data is distributed and will require mixing internal and external data (Mate et al 2012:112). To achieve this, in most cases, the main data warehousing service providers are offering OLAP servers (such as, IBM Cognos, Microsoft SQL/Excel, Oracle Hyperion, and SAP Business Objects BI Platforms).

Chaudhuri et al (2011:90-93) and Asghar et al (2009:1) agree that data warehouse servers are usually accompanied by a set of mid-tier servers, that is, OLAP servers that provide specialised analytical features to manage a variety of BI circumstances. The good thing about OLAP servers is that they are able to interpret the multi-dimensional view of data to applications or users so that most used BI operations, such as, data mining, filtering, aggregation, 'drill-down', 'rollup' and 'pivoting' can be performed.

There are various ways of implementing OLAP servers, for example it can be done through “a multidimensional storage engine (MOLAP), a relational DBMS engine (ROLAP) as the backend, or a hybrid combination, called HOLAP”.

In summary, in a normal business environment, end-users record their data using OLTP systems while OLAP systems are used to read on the written data after the ETL processes and data has been stored in the data warehouse, data mart or ODS. The OLTP and OLAP systems are normally deployed independently, “using different end-user interfaces, and target different scenarios” (MicroStrategy 2012:179). However, there is a desire to deploy both OLTP and OLAP in one system, as the use of both servers could be more appreciated than using different systems since they both use the same data (Plattner 2009:1).

2.8 Total cost of ownership of BI deployment

In order to get a broader understanding of BI total cost of ownership in different organisations, some of the organisational characteristics that are often considered are as follows:

- Industry
- Number of employees or users
- Size: small, medium, or large
- Departments or units: specialisation types
- Branches
- Sources systems: ERP or CRM
- Number of IT, BI professionals
- Sensitivity of data in use
- Data volumes

In addition to these characteristics, the following key procurement options are also considered when determining the TCO for the BI solutions (Daneshgar, Low & Worasinchai 2013:1743; Cortellessa, Marinelli & Potena 2008:3091):

- *In-house BI solutions (custom developed solutions)*: these are BI solutions that are developed within the organisation using their own subject matter expertise which ensures that the new solution meets the functional requirements.

- *Buying packaged BI solutions (commercial-off-the-shelf purchase)*: buying of commercially available BI solutions which normally require the organisation to adjust its critical features so that the BI solution can meet the functional requirements of the business.
- *Contracted BI solutions (rent/leased)*: these are BI solutions that are used on a 'when and, as needed basis' and are often cloud-based solutions. In most cases organisations do not need to invest upfront, but rather subscribe to the necessary BI solution when the need arises.

Furthermore, to determine successfully the hidden cost factors of acquiring, deploying and maintaining a BI system in the various models to be discussed, the TCO analysis will be used to help obtain an accurate and transparent cost model for BI investments by using direct, indirect, and related costs (King 2007:8; Plattener & Zeier 2012:16). Direct cost are costs associated with deploying a new BI system, such as, software licenses, development, subscription, integration, customisation costs, professional service, and hardware costs (Bibi, Katsaros & Bozanis 2012:88). Indirect costs are associated with ongoing management costs which ensure that procured and deployed BI systems function according to expectations. Related costs are costs associated with the introduction of the new BI system.

The costs of successful BI implementation extend beyond the initial direct and indirect costs associated with acquisition, deployment and maintenance. For example, the employment of IT professionals stand "between 60% and 86% of BI total cost of ownership" (MicroStrategy 2008:5). The following are some of the identified key metrics for measuring direct, indirect and related costs which are associated with the three broad TCO cost categories identified in table 2-1, namely, software, hardware and personnel costs (King 2007:8; Geiger 2006:11; Birst 2013:5; MicroStrategy 2008:6; Reddy & Rao 2012:41):

Table 2-1: Naïve TCO Cost Categories

Key metrics	Naïve
Direct costs (DC)	
• Cost of hardware (e.g. new servers)	X
• Cost of middleware (operating system, database servers, Web servers)	X
• Cost of software licenses	X
• Cost of consulting services for implementation	X

Key metrics	Naïve
• Cost of staff hours to implement the project	X
• Cost of software quality	X
• Costs of changes needed in the network	X
• Cost of additional storage	X
Indirect costs (IC)	
• Cost of upgrading related software	X
• Cost of upgrading related hardware/middleware	X
• Cost of customisation/Continued application evolution (new end-user reports/dashboard)	X
• Impact on use of existing servers	X
• Impact on existing network traffic	X
• Impact on existing storage	X
• Increase in support calls to IT	X
• Increase in maintenance of hardware/middleware	X
• Downtime cost (loss on productivity costs)	X
Related costs (RC)	
• Staffing costs	X
• Training end-users	X
• Training BI professionals (business analysts, data science)	X
• Training IT staff	X
• Creating governance rules	X
• Implementing governance rules	X
• Maintaining governance rules	X

It is understood that naïve initiatives are often performed by informal organisations that hardly have any BI business users who could even attempt to invest in BI in order to acquire the software. BI users in the naïve BI deployment are usually casual or student users who would choose a free or trial software acquisition just to familiarise themselves with a certain BI solution in preparation for moving into a more mature category of organisations that are more dependent on BI for their day-to-day operations.

Therefore, the TCO in such organisations is often near zero. Furthermore, such category of business users may also opt for a cloud solution by temporarily creating an account from cloud BI solutions service providers and making use of it only for personal reasons. As such, there cannot be a sound argument or recommendation to be offered to this category of BI users, apart from that of advising them to explore BI solutions on a trial basis, and also by accessing some free software in the BI landscape. But the ultimate recommendation is to suggest that they do not have any TCO at all.

2.9 Conclusion

The purpose of this chapter was to provide a general understanding of BI, its use and the associated business value. It covered in detail the key concepts and terminologies which are commonly used in order to understand and implement BI successfully. In order to do that, a literature review on the definition of BI in general was conducted. This contributed to an understanding of the main features and functions that enable BI to be used as a business management tool, consisting of applications and advanced technologies that gather and analyse information in order to support decision-making and realise business value.

The value generated from the use of BI was demonstrated through a BI value chain. This was supported with a naïve BI model which covered data sources systems, data storages, analysis and reporting tools and decision-making. The naïve model later informed the development of a BI architecture which further covered the different ideal layers, that is, the presentation, service, data, and source systems layers. Then the discussion finished with an explanation of the application of OLTP and OLAP in relation to BI.

Clearly, having achieved the purpose of this chapter, the various deployment methods of BI applications should be explored. This should contribute to a comprehensive understanding of how BI is implemented in organisations. Therefore, the next chapter discusses the first type of BI deployment, that is, on-premise BI deployment.

CHAPTER 3 ON-PREMISE BUSINESS INTELLIGENCE

3.1 Introduction

In the previous chapter, a general overview of the definition of BI and its value chain activities and components leading to the realisation of a BI value proposition was presented. A naïve BI model representing a BI problem was discussed. The model informed the ideal architecture for implementing a BI solution which can support decision-making.

The purpose of this chapter is to provide a more specific understanding of BI through an investigation of 'on the on-premise' deployment of BI and the resultant business value. The various on-premise BI deployment modes are discussed in order to gain a comprehensive understanding of the function of BI in an organisation. As understood from the previous chapter, organisations now seek to deploy BI solutions in different ways with the availability of various communication infrastructures in mind. Therefore, this chapter partly provides an answer to the second sub-research question:

“What are the most currently used models at organisation level and the limitations thereof?”

The remaining sections of the chapter are as follows: Section 3.2 provides a detailed discussion on the on-premise BI deployment, including the various deployment modes. Then section 3.3 proposes and discusses a potential on-premise BI model to explain on-premise BI deployment which is followed by an ideal architecture for implementing such a model in section 3.4. Also covered in this chapter are the discussions on the advantages and challenges of on-premise BI deployment in sections 3.5 and 3.6 respectively. TCO for an on-premise BI deployment is discussed in section 3.7 and then the chapter is concluded with a summary discussion in section 3.8.

3.2 On-premise business intelligence deployment

Organisations face a variety of choices when it comes to deploying BI solutions in the business environment, and one of the most commonly used types of deployment is on-premise deployment (Gastermann, Stopper, Kossik & Katalinic 2015:577), as well as cloud and hybrid deployments which are discussed in later chapters. From a

general perspective, on-premise deployment has to do with the way in which an IT system or solution is implemented and executed on technological devices located within an organisation, instead of being hosted at a remote facility, for example in the cloud (Wikipedia 2014). This may also mean that the locally installed or acquired software could be managed by the organisation internally, or by an external service provider.

In the context of this study, on-premise BI involves deploying the hardware, software and application systems which enable data to be sourced; and historical or real-time data to be analysed in order to present information for supporting informed decisions inside the organisation's premises (Winkler & Brown 2013:20). Most organisations choose this option as it is secure and has low risks as far as business data is concerned. Furthermore, this option is able to perform its own backups and in this way ensure disaster recovery (Srinivasan 2011:341). This also ensures that decision-makers feel safe with their data housed within their own data centre premises and, in certain cases, it also ensures that organisations comply with regulations that require data to be kept on the premises (Kaur, Azad & Sing 2013:25). For example, the South African Revenue Services (SARS) requires that organisations keep hard-copy documents on-premise, such as, employees' IRP5s for 5 years so that, should there be an audit query, the organisation can then consult the physical file.

Most organisations are using on-premise BI solutions (for mission-critical or core applications) to derive business value from the ever-growing business data mainly because of security and privacy concerns (Bibi et al 2012:86). However, the on-premises deployment requires that organisations own their own infrastructure, and this approach has been understood to relate to higher investments (Gastermann et al 2015:576). This approach has led to an on-going need for different departments or divisions within an organisation to use locally hosted data and legacy BI enterprise applications in order to fulfil their BI needs (Mateescu & Vladescu 2013:616). In addition, with this approach, organisations have opted to deploy BI in various deployment modes which ensure that their departmental data is utilised to fulfil the respective departmental BI needs MicroStrategy (2012:8).

For the purpose of this study and in order to provide a comprehensive explanation of on-premise BI deployment, three deployment modes, namely, isolated, overlapping and centralised or shared deployment are discussed in the next sections.

3.3.1 Isolated BI deployment

Isolated (silo) BI deployment is applicable in cases where organisations have deployed BI applications as departmental solutions (Harriott 2013:38). MicroStrategy (2012:8) emphasises that many organisations deploy departmental BI applications in order to meet departmental BI needs, and as a result, they have a large collection of different BI technologies. Most organisations adopt departmental BI deployment so that they can reliably and speedily use their departmental data to make informed decisions (Davenport & Harris 2007:3). As illustrated in figure 3-1 below, in most cases this kind of deployment allows organisations to use reliable, valid and actionable data, and the resulting organised information is privately owned and utilised within organisations (Mate et al 2012:111; Kruger 2010:257).

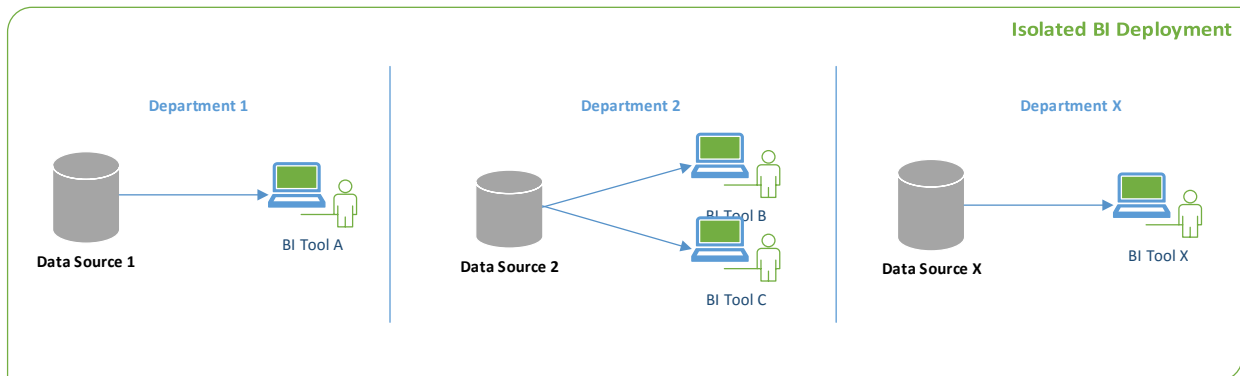


Figure 3-1: Isolated BI Deployment

It is clear from figure 3-1 above, that in an isolated BI deployment, each department in an organisation has the ability to use its own data source and perform BI tasks based on the requirements of the department. This means that department 1 uses only data source 1 and its own BI technology to achieve its own BI needs. However, department 2 has different business users that may be using different BI applications to perform the much-needed BI to achieve their needs. Some initial business needs, such as, the ability to generate weekly or monthly sales reports, are met through departmental BI

deployment as each separate BI solution is installed to manage unique business requirements for the business user (MicroStrategy 2012:8). However, some challenges have emerged, such as, the inability to communicate across departments, system incompatibility, etc., and these problems are discussed later in detail.

3.3.2 Overlapping BI deployment

Overlapping BI deployment occurs where BI applications deployments are successfully expanded within the organisation and have reached a point where they are no longer isolated (MicroStrategy 2012:8). The success of this kind of deployment has led to the expansion of BI applications that overlap in application usage as depicted in figure 3-2 below. As a result, this has led to unsustainable situations where different used BI applications are producing inconsistent versions of the truth and the only way to arrive at the same truth is to manually collate the results (Kwon et al 2014:389; Pant 2009:16). The other challenge is that whenever organisations want to perform BI tasks on their multiple data sources, the only way to be successful is to use different BI applications. Therefore, despite the advantages of on-premise BI deployment, the implementation of several BI applications from different vendors is often very expensive (Sun et al 2012:512).

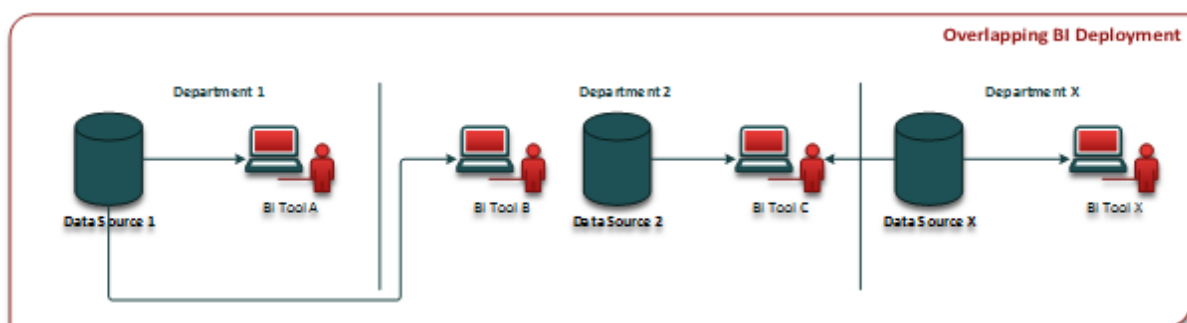


Figure 3-2: Overlapping BI Deployment

The overlapping BI deployment model in figure 3-2 above, shows that although department 1 has a dedicated data source 1 and BI tool A, department 2 also has access to their data source. Department 2 has two different BI tools that access different data sources, and in particular BI tool C accesses or acquires additional data

from department X in order to complete its BI tasks. Schlegel, Milbury, Buytendijk & Sommer (2014:3) emphasise that in most organisations, data analysis occurs in various departments at different levels. For example, analysis occurs within horizontal units, such as, sales, marketing, HR, financial or other departments. Although these departments achieve the intended results, clearly there is a great amount of overlapping in this deployment mode.

3.3.3 Centralised or shared BI deployment

Centralised or shared BI deployment occurs in an organisation which has a dedicated department that integrates, reports and analyses data from across the organisation (Schlegel et al. 2014:2). The advantages of this approach are that data from different sources is stored in a central and shared database management system, and all business users with access privileges have access to a more stable and more secure database system (Dyché et al 2011b:15; Pant 2009:15). Furthermore, Boateng et al (2012:229) maintain that this deployment mode offers a cheaper approach to managing a database management system that provides access to BI tools as depicted below in figure 3-3. Although this deployment mode has many advantages, the drawback is that it can lead to user disengagement and an unsustainable workload for BI users (O'Donnell et al. 2012:211). For example, O'Donnell et al (2012:211) point out that BI users may land in a situation where they are unable to identify and consolidate duplicate reports.

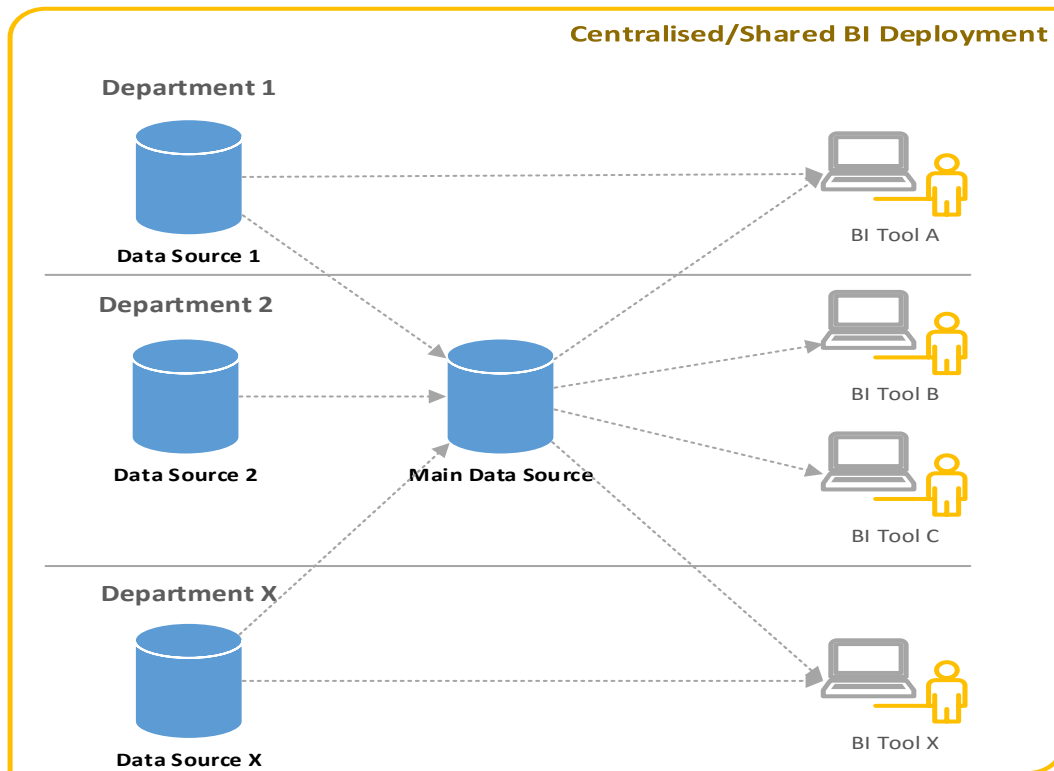


Figure 3-3: Centralised BI Deployment

Although different departments in the centralised BI deployment model as illustrated in figure 3-3 above have dedicated data sources, data is collected and stored in a central main data source. The main data source provides data access to different BI tools residing in different departments. Although different departments (1, 2, X) have the advantage of working from the same data, they still face the challenge of using various independent BI tools (A, B, C, X) which could be a costly exercise.

The use of different BI tools in the three different deployment modes discussed above is motivated by the fact that vendors design and develop BI tools to meet specific business needs, and functionalities, such as, architecture, data management, analytics and information delivery (Birst 2013:1). For example, BI tool A could be a Pentaho¹⁵ product with service-oriented architecture that supports application programming interfaces for data access and integration (Sallam, Hostmann, Schlegel, Tapadinhas & Oestrelch 2015:30). BI tool B could be a Birst¹⁶ product mainly focusing

¹⁵ Pentaho, Inc. provides a Business Analytics platform which is unified, open source, embeddable, and has pluggable capabilities.

¹⁶ Birst is the only enterprise-caliber BI platform born in the cloud which is more agile than legacy BI and more powerful than Data Discovery

on data management, that is, data extraction, transformation and loading (Khan & Quadri 2014:25; Birst 2013:1). BI tool C with a combination of data management and analytics and information delivery, that is, ETL and predictive or descriptive analytics, dashboards, reports, visualisation and data discovery, and BI tool X could have all capabilities in one product, such as, SAP BI suite.

In summary, the on-premise utilisation of the emerging latest BI and analytics technologies has not only helped support faster decision-making, they have decreased inactivity in data warehouse updates and without impacting performance and resources utilisation (Rahman et al 2013:543). The success of on-premise BI deployment and utilisation is dependent on user training as well. Hobek et al (2006:34) recommend that users must be trained after completing the implementation of the BI project. This will ensure that users are skilful and are able to use the new technology.

In order to support the various BI deployment modes discussed above, a model can be developed so as to accommodate all of them to be discussed in the next section.

3.3 On-premise business intelligence model

Based on the BI deployment modes discussed in the previous section, on-premise deployment of BI is believed to be largely dependent on the organisational structure, departmental needs, and size. It is understood that BI can be deployed in isolation, overlapping, and centralised modes which come with advantages and disadvantages (MicroStrategy 2012:8). The model in figure 3-4 below demonstrates how a BI system can be deployed in an organisation with the intention of managing data coming from various sources and to support departments which need the information whenever the situation arises (Davenport & Harris 2007:3). The on-premise BI model shown in figure 3-4 below demonstrates the utilisation of four main components, namely, multiple source systems from the independent departments, different data stores, isolated or centralised BI tools and various types of reports.

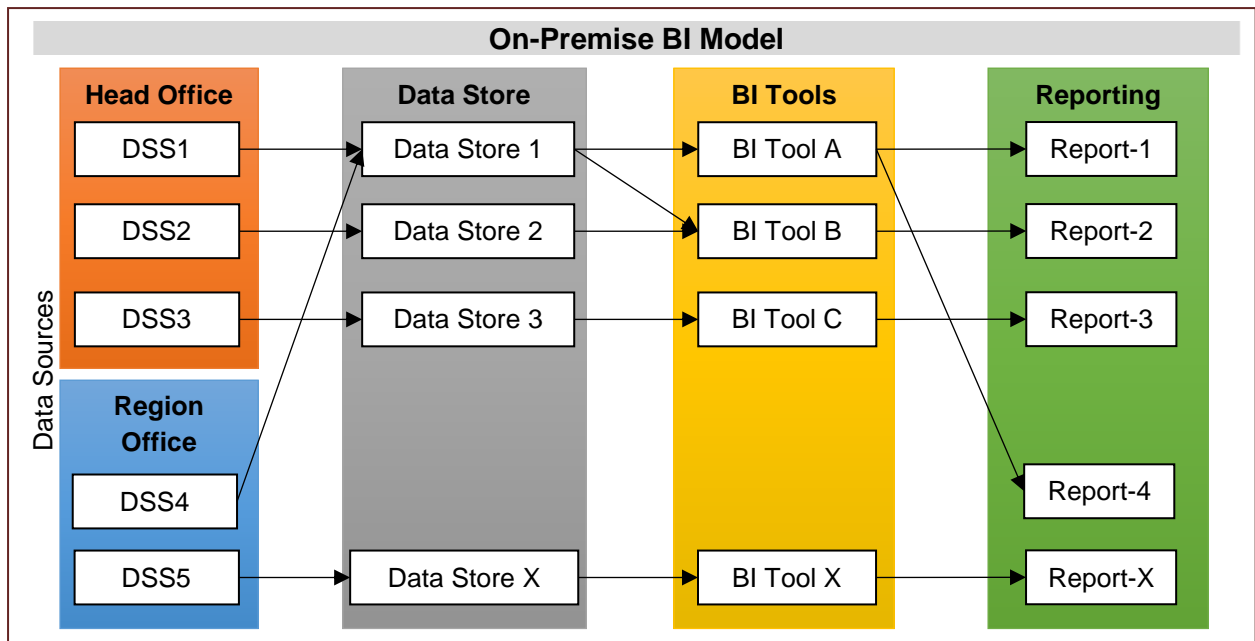


Figure 3-4: On-Premise BI Model

The following sections discuss the components of the on-premise BI model in detail.

3.3.1 Multiple sources systems

The on-premise BI models above demonstrate that the head office and regional office uses different source systems (DSS); for example DSS1¹⁷, DSS2¹⁸, DSS3¹⁹, and so on, that captures and generates the business data in their respective departments. These source systems are all OLTP based operational systems and provide the relevant data for BI. The crucial part about these OLTP based systems is their ability to process large amounts of data which is then stored in individual databases for analysis (Ong et al 2011:4; Williams & Williams 2007:201). This example illustrates and confirms that multiple source systems can exist in the business environment (Van der Linde 2013:118; Ghazanfari, Jafari & Rouhani 2011:1580). These systems also generate assorted data which may be in the form of structured data or unstructured data, such as, digital, multimedia, text file, and email (Bucur 2012:103). The critical and challenging task in such a multiple source system set-up is integrating data for reporting purposes. Organisations often find themselves with departments that

¹⁷ DSS1 - SAP: Customer Relationship Management

¹⁸ DSS2 - Salesforce.com: Sales Application

¹⁹ DSS3 - IBM: Human Resources Application

operate in silos and data never becomes integrated so that informed decisions can be made and reported across the organisation.

The advantage of using DSS1 and DSS4²⁰ systems is that they share the same Data Store 1 from the same software vendor; and hence the data format coming out of the source system can easily be integrated for reporting purposes. On the other hand, one of the drawbacks of using the data coming from multiple sources, that is, DSS2, DSS3, DSS5²¹, and so on, is that, they use source systems from different software vendors and are extremely likely to have a problem with different data formats (Sen et al 2012:341). In addition, Van der Linde (2013:116) maintains that managing and reporting on data emanating from various data source systems poses serious challenges. This also means that most of the business users will spend time trying to consolidate data from the other departments so that they are able to obtain a comprehensive view of the organisation.

3.3.2 Data Stores

The on-premise model above illustrates that there are several different data stores which are all linked to their respective source systems. These data repositories store data through the ETL process for analysis and reporting purposes to support informed decision-making (Sen et al 2012:336). The limitation of the data stores in the model is that they enable each department to use only its corresponding BI tool to perform analysis and reporting based on data available.

There are minimal advantages and more disadvantages in BI deployment that has data stored in isolated data stores, that is, 2, 3, X, etc. or shared in data stores 1 and 4, which are discussed in much detail in the later sections. One of the key issues emanating from the model is that an organisation with such a BI deployment will still face the challenge of having multiple data stores that are configured in multiple file formats (Birst 2015:1). This, in turn, will limit the use of enterprise-wide data to

²⁰ DSS4 - SAP: Enterprise Resource Planning

²¹ DSS5 - Oracle: Supply Chain Management

satisfying only departmental needs as there is no main BI tool that has the ability to use multiple source systems and data stores.

3.3.3 Business intelligence tools

In the model, the BI tools that business users utilise to perform various analytical functions to support strategic, tactical and operational decision-making all have different independent data stores (Davis et al 2009:3). However, BI tools, 1 and 4 have the same BI tool implementation but all reside in different geographical locations, resulting in overlapping BI applications deployment. The rest of the data stores have isolated BI tools making sharing data for reporting purposes very difficult, if not, impossible.

This type of BI deployment has kept some organisations wondering whether their current BI tools are good enough to support their decision-making processes (Birst 2015:2). In addition, isolated BI tools deployment disseminate new problems when the applications grow as they are unable to accommodate incremental growth of the BI infrastructure (MicroStrategy 2012:9).

3.3.4 Reporting

In order to make informed decisions and realise the business value of having business data as an asset, the on-premise BI model shows that business users generate, access or receive different types of reports from BI tools – A, B, C, and X (Watson & Wixom 2007:97). This has been made easier because each BI tool has direct access to the data store and there is no need to wait for data to be transformed so that the necessary analysis and reporting can be performed.

In addition, most of these BI tools have OLAP capabilities e.g. SAP BusinessObjects Analysis edition for OLAP helps organisations better understand their business data and provide intelligence that can be used to improve the performance of the business (SAP 2014:5). IBM Cognos BI system has the Cognos Dynamic Cubes feature that performs interactive analysis over large data volumes in a way that allows for greater

performance over high data volumes (ZiffDavis 2012:1). Oracle OLAP - Database 11g Release 2 delivers advanced multidimensional analytic capabilities and was “designed to provide excellent query performance, fast incremental updates of data sets, efficient management of summary data and rich analytic content” (Oracle 2012:1).

In summary, it is understood that BI deployment can be in isolation, overlapping and, or centralised or shared and the model has demonstrated the possibility of implementing such deployments. Although on-premise BI deployment solutions are able to support the decision-making process, more can be done with the current advancement in technology. To achieve this, in most cases, dedicated and specialised BI departments are used, such as, the BI Centre of Excellence, that act as an official department with expertise and responsibilities for ensuring that BI supports the business decision-making process (Dyché et al 2011:44). The on-premise BI model also shows that it has limitations when it comes to expansion which is required in any BI deployment as organisations are constantly growing and so is their data. There is a need for a BI solution that can integrate and analyse data managed in different enterprise applications and has the ability to scale up to meet future organisational needs (Birst 2015:2).

The following section discusses the architecture for implementing an on-premise deployment of BI.

3.4 Architecture for on-premise deployment of business intelligence

The on-premise BI model can only be complemented with an architecture or framework that accommodates all those elements which can best deliver the organisation’s BI needs. The on-premise BI architecture below demonstrates the ideal structure for delivering enterprise-wide on-premise BI deployment. Based on the previous discussion on the on-premise deployment modes and model, the architecture below illustrates the key components for implementing the on-premise BI solution.

Although various implemented on-premise BI solutions generate the required results, BI technology is evolving at a very fast rate. It is recommended that new BI technologies can increase agility and lower costs as it has the capability to adapt rapidly to changes in the business environment (Mvelase et al 2009:215). The use of various on-premise BI solutions helps organisations to analyse and predict hurdles in

advance that they may face during major projects, which eliminates the negative impact on the business (Khan, Amin and Lambrou 2010:15). With this in mind, most organisations struggle to determine which BI model and architecture would best fit their requirements to effectively run their business and meet BI objectives.

Considering the discussion on the on-premise BI model, an architecture consisting of five (5) main layers is suggested namely: the presentation layer, the network layer, the service layer, the data layer, and the source systems layer. Figure 3-5 below depicts the overall architecture each layer (and its components) of which is discussed in the following subsections.

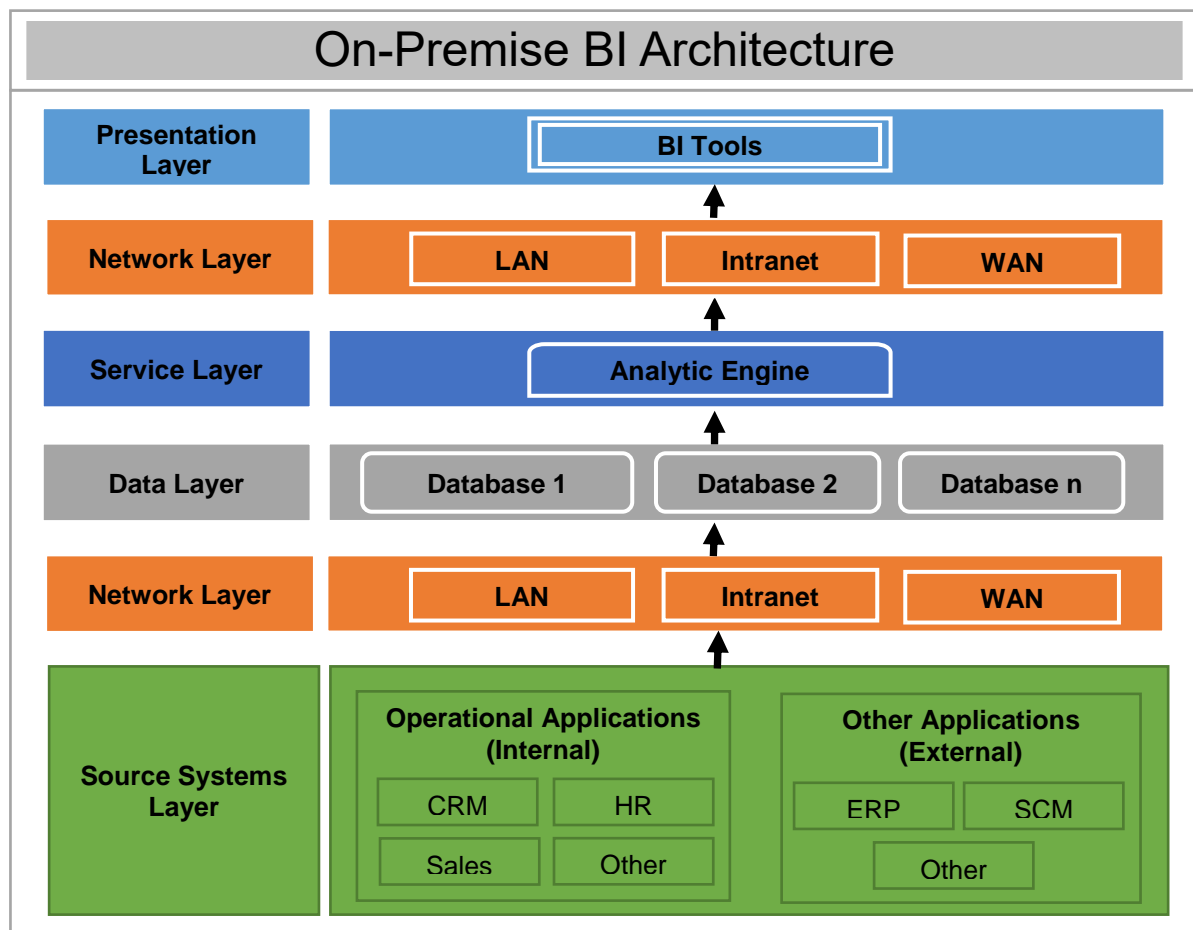


Figure 3-5: On-Premise BI Architecture

3.3.1 Presentation layer

The presentation layer supports business by linking information to business users by presenting reports according to the business user's locality which in this case are the

head office and regional office, respectively. Based on the discussion in the on-premise BI model section, the common BI tools that are used in the presentation layer are SAP BusinessObjects BI solution, Cognos, Oracle BI Publisher, and so forth. These tools allow power users to configure reports and analyse data according to business users' information requirements (Russom 2013:34). Furthermore, business users with non-IT background are able to access or generate needed reports easily without having to always request reports through power users.

In the context of this study, the BI tools discussed in the presentation layer, namely, BusinessObjects, Cognos and BI Publisher are flexible, scalable, pre-configured and reliable for providing reports to business users through the Local Area Network (LAN), Intranet, Wide Area Network (WAN), and mobile technologies, etc. (TEC 2012:16). This would allow those business users who utilise the data source systems discussed in the on-premise model to access the data stores easily and perform all necessary BI tasks in order to support decision-making processes and realise tangible benefits across the organisation. In addition, these BI tools provide a powerful, in-built interface between the organisation and critical business information (Evelson 2014:9; MicroStrategy 2012:22).

3.3.2 Network layer

The network layer ensures that data, information and reports are made available to both power users and other business users. The network layer facilitates the transfer of data from the source systems into the analytic engine for report preparation (SAP 2012:5). In the on-premise BI model business users are able to access the various BI reports (that is, reports 1, 2, and 3) through the use of a LAN and, or Intranet infrastructure. On the other hand, other business users access their data and generate BI reports (4 and X) through the WAN infrastructure connected to the organisation's LAN and can share with the rest of the organisation should the need arise.

3.3.3 Service layer

The service layer houses the analytics engine for the model. The data source systems in the model individually utilise the analytics engine to perform analytical functions, such as, descriptive, prescriptive, predictive, text, diagnostic, and mobile analytics in order to achieve and support a better decision-making process (Watson 2012:5; TEC 2011:66). The data for analytic purposes is received through the LAN, intranet and WAN infrastructure and made available for reporting purposes through the same infrastructure. The service layer forms part of the key components for facilitating the process of supporting better decision-making by enabling the BI analytical capabilities (Siller 2013:111). It supports a large amount of flexible configurable data for different BI tools in order to enable centralised or shared BI capabilities (Liyang, Zhiwei, Zhzngjun & Li 2011:1026). These analytical capabilities are accessible through the presentation layer by means of desktop applications, web portal, mobile clients, external application programming interface, etc.

3.3.4 Data Layer

The data layer in the architecture has the capability of accommodating several databases to store data from the different data source systems discussed. Based on the on-premise BI model, this means that it can accommodate several independent databases for SAP, Salesforce.com, IBM, Oracle applications, and so forth. As more departments are added to the organisation, more and more individual databases can be added in order to fulfil the BI requirements which can support better decision-making (Chaudhuri et al 2011:98). The outcome of this isolated BI application arrangement is that often more strain is placed on the IT departments in managing and maintaining the individual databases (MicroStrategy 2012:7). However, it also has the capability of hosting a centralised or shared database which can be facilitated through the use of available communication infrastructures, such as, LAN, Intranet and WAN (Minoli 2008:106).

3.3.5 Source systems layer

The source systems layer constitutes all internal and external (regional) applications used in the on-premise BI model. The source systems layer happens to be flexible, available, reliable and scalable with the capability of hosting a number of applications from different vendors (Liyang et al 2011:1026). Ong et al (2011:4) assert that internal applications are responsible for managing operational data in the organisation and can include OLTP applications, such as, CRM, Sales Application, HR, and other applications; while external applications reside outside the organisation (for example in the regions) and include applications, such as, ERP, SCM and other applications. Furthermore, they emphasise that it is useful to know the data source systems for the data needed for a particular business need as this will eliminate time wasting and will accelerate the provision of information. These applications often contain real-time data utilised to support operational activities and decision-making processes.

In summary, the architecture discussed above suggests that the on-premise model can be supported by a flexible architecture that specifically focuses on addressing BI requirements for the organisation. However, much can be done to do away with isolated approaches to meeting BI requirements by utilising the latest and standardised approach which can better support an enterprise-wide BI implementation.

The next section discusses the advantages of deploying an on-premise BI solution.

3.5 Advantages of on-premise business intelligence deployment

BI offerings delivered through on-premise deployment have a number of advantages.

- **Security of data:** On-premise BI deployment is less risky when sensitive data which require data transfers and manipulations have to be handled (Yang 2012:6). Organisational data is secure and easily accessible as it is located within the organisation. Organisations have the advantage of being able to utilise their on-premise data storage and information optimally by means of a well-organised data and information delivery system (Lonnqvist & Pirttimaki 2006:38). With on-premise, hacking or other forms of unauthorised access to data is usually well managed

when appropriate security measure are in place, such as using a vault with high end security features to manage data storage systems (Nunan & Domenico 2013:5). Organisations utilising data stores in the on-premise model enjoy these benefits although their data stores and data source systems are located in different departments and geographical regions.

- **Ownership of BI systems:** There is usually full ownership and control of on-premises BI systems by the organisations as most of these applications are acquired through purchasing an off-the-shelf solution or are custom developed according to Najjar and Kettinger (2013:215). Furthermore, they assert that full control of on-premise BI systems enable organisations to manage information which allows a lot of business executives to make better data-driven business decisions easily. After-sales maintenance and support costs are the only aspect that would concern most organisations. All the data stored in the on-premise BI model are owned by the organisation and none of it comes from external third party source systems apart from the regional offices.
- **Software license cost:** Depending on the vendor selection, there could be no on-going license cost once the BI application has been purchased or custom developed. The cost of licenses in some on-premises BI implementations will be different depending on the number of users and specific features required (Menon & Rehani 2012:11). In some cases it might be cheaper depending on the BI vendor utilised. There would be a number of different licenses in the on-premise BI deployment as it involves a number of different BI tools and BI deployment modes.
- **Customisation is easier:** On-premise BI systems are easier to customise to meet ever changing business requirements (Winkler & Brown 2013:21). Organisations need BI systems which are scalable in order to meet dynamic business needs and, at the same time, maintain competitive advantage. Olszak and Ziemia (2007:143) assert that although the creation of customised BI applications requires a lot of time, depending on the system's complexity, the benefit of a logical and consistent BI system provides competitive advantage to the organisation in the long term. The

model can easily attain competitive advantage, among other benefits, as most of the data source systems, such as, ERP, CRM, and SCM are fully customisable.

- **Easier integration:** On-premise BI deployment is often used so that organisations can easily manage data coming from various mission critical enterprise systems, such as, ERP, CRM, and HR) which must operate in a closed environment and be managed by the organisation's own IT departments (Plattener & Zeier 2012:219). This arrangement ensures that reliable business data originates from a particular source in real-time to meet BI business requirements. On-premise BI allows for seamless linkage among enterprise systems enabling easier data and application integration (Funahashi & Yoshikawa 2011:286). It is often easier to integrate different data source systems in the on-premise BI since most of the systems reside within the organisation's premises with a small portion in the regions.
- **Reduced infrastructure cost:** With on-premise BI implementation, there is often a reduced IT infrastructure cost that comes as a result of data mart consolidation and data fixing as it reduces unnecessary retrieval of data and data stored in duplicate across the organisation (Watson & Wixom 2007:97). The on-premise model illustrates a centralised or shared BI deployment which can ensure that there are reduced infrastructure costs.
- **Improved business processes:** On-premise BI implementation has the capability of yielding better quality information emanating from performing departmental analytical functions (Popovi et al 2010:12; Watson & Wixom 2007:97). Furthermore, this quality information is often used to plan and improve, or change departmental business processes that support decision-making and, simultaneously, realise business value. Departments utilising quality information according to the on-premise model, will be able to share informed decisions.

The next section discusses the limitations associated with implementing an on-premise BI solution.

3.6 Limitations of on-premise business intelligence

The following are some of the disadvantages or limitations of implementing an on-premise BI system:

- **Managing and maintaining large numbers of BI applications:** The IT department faces a daunting challenge to manage and maintain isolated departmental BI applications as, and when the number of applications powered by different technologies grows (MicroStrategy 2012:7). This normally calls for a more convenient and standardised way of managing such BI applications.
- **Costly IT infrastructure investment:** On-premise BI implementation involves huge IT infrastructure investments which also means that should the project be unsuccessful, there would be huge losses (Winkler & Brown 2013:20). It is, therefore, imperative that organisations utilising the on-premise BI model have a strong governance framework to handle all decision-making processes and approval in order to mitigate such investment risks.
- **Inconsistent reporting leading to manual interventions:** On-premise duplication of BI implementation leads to an unmanageable situation because different BI applications often produce poor quality of information and the only way to correct the situation is to have a manual intervention to organise the results (MicroStrategy 2012:8). Furthermore, whenever organisations need to perform BI tasks on their multiple data sources, they have to use different BI applications in order to achieve the intended business requirements. This often leads to not having a single version of the truth that can support informed decision-making.
- **Managing Big data challenges:** Organisations face big data challenges as more and more electronic data is generated, which cannot be handled by on-premise relational database management systems (Chaudhuri et al 2011:90). As a result, there is an increasing desire to design and build low-cost data BI solutions that can support much larger data volumes. Organisations that would implement the proposed model would be greatly affected by the big data challenge. Thus, they are always assessing and reviewing the practicality of the on-premise BI solution.

- **High maintenance and transaction costs:** On-premise BI is often complex and costly considering that organisations need to have significant upfront infrastructure investments (Kaur et al 2013:24; Gash et al 2011:263). This also means that, the management cost of the proposed on-premise BI systems will be high (Plattener & Zeier 2012:219). This also means the transaction cost of executing the applications that generate the much needed business data is the responsibility of the organisation on which the on-premise BI resides and not that of the service provider. Winkler and Brown (2013:23) further assert that there are also internal transaction costs when the on-premise BI applications have to be customised in order to meet business requirements as they are usually required to configure complex tables or have custom developments.
- **Lack of real-time access to critical business information:** Isolated and repeated on-premise BI deployments are impractical and unattractive in conditions where the organisation needs to use all the information opportunities to maximise performance and minimise cost (Marinela et al 2011:2). This results in critical business information being missed since manual processes are often used to correct the situation.
- **Lack of application scalability:** On-premise BI deployments lack the ability to scale up and down to meet the demands of the ever-changing business environment (Ouf & Nasr 2012:653). This also means that certain targets cannot be met by the organisation which will have a negative impact on its performance. Although the proposed on-premise architecture has the flexibility to grow, on-premise BI deployment cannot expand more as is the case with other cloud-based solutions.
- **Lack of geographical scalability:** Geographical scalability is often a challenge as in most cases users are unable to access the BI applications remotely (Menon & Rehani, 2012:6). This normally calls for a costly network infrastructure that can enable remote users (such as regions in the model) to effectively communicate and share information with the head office departments.

3.7 Total cost of ownership of on-premise BI deployment

TCO in on-premise deployment is largely dependent on the organisational strategy, structure, departmental needs, and size. There are number of strategies an organisation can adopt to keep the TCO cost down, for instance, if the strategy is minimal in nature within an organisation of small size, the TCO will reduce because the indirect cost or related cost could be reduced significantly. Another example is by using open source which can have an impact on software quality, development and licensing costs. In contrast to the naïve model described above in which costs are not necessarily involved, the table below summarises the TCO associated with on-premise BI deployment.

Table 3-1: TCO On-premise Deployment

Key metrics	Naïve	On-premise
Direct costs (DC)		
• Cost of hardware (new servers)	X	✓
• Cost of middleware (operating system, database servers, Web servers)	X	✓
• Cost of software licenses	X	✓
• Cost of consulting services for implementation	X	✓
• Cost of staff hours to implement the project	X	✓
• Cost of software quality	X	✓
• Costs of changes needed in the network	X	✓
• Cost of additional storage	X	✓
Indirect costs (IC)		
• Cost of upgrading related software	X	✓
• Cost of upgrading related hardware/middleware	X	✓
• Cost of customisation/continued application evolution (new end-user reports/dashboard)	X	✓
• Impact on use of existing servers	X	✓
• Impact on existing network traffic	X	✓
• Impact on existing storage	X	✓
• Increase in support calls to IT	X	✓
• Increase in maintenance of hardware/middleware	X	✓
• Downtime cost (loss on productivity costs)	X	✓
Related costs (RC)		
• Staffing costs	X	✓
• Training end-users	X	✓
• Training BI professionals e.g. business analysts, data science	X	✓
• Training IT staff	X	✓
• Creating governance rules	X	✓
• Implementing governance rules	X	✓
• Maintaining governance rules	X	✓

The cost elements for the naïve model are not all applicable, as users in this category are not so keen to invest in BI since they are casual users. However, the cost elements for on-premise BI deployment are applicable in most instances depending on the size of the organisation. As an example, license costs are applicable to on-premise BI users but not for naïve users since they mostly use available free software. Therefore, naïve and on-premise BI are fundamentally different. The naïve model cannot even be used for cross-comparison in real life because it is not centred on the organisation as opposed to the on-premise model that can be adopted by organisations.

3.8 Conclusion

The purpose of this chapter was to provide a more specific understanding of BI through an investigation of the on-premise deployment of BI and the resultant business value. It covered the various on-premise BI deployment modes, that is, isolated, overlapping and centralised or shared that help us to gain a comprehensive understanding of how BI works in an organisation. As understood from the previous chapter, organisations now aim to deploy BI solutions in different ways considering the availability of various communication infrastructures.

The chapter discussed the on-premise BI deployment model which illustrated the utilisation of four main components, namely, multiple source systems from the independent departments, different data stores, isolated or centralised or shared BI tools and various types of reports. The model informed the development of an ideal BI architecture consisting of five (5) layers namely: the presentation, network, service, data, and source systems layers. Finally, the advantages, challenges, and related TCO for implementing the on-premise BI deployment were discussed.

Although the BI market still appears active and the importance of BI applications is more widely accepted, there are a limited number of studies providing guidelines on the on-premise BI model as it is such a resourceful initiative (Chuah 2010:307). Having achieved the purpose of this chapter, the next chapter seeks to bridge the gap that exists between organisations and practitioners by discussing another type of deployment known as cloud-based BI deployment.

CHAPTER 4 CLOUD-BASED BUSINESS INTELLIGENCE

4.1 Introduction

In the previous chapter, the subject literature on the topic of business intelligence was reviewed and three specific modes of on-premise BI deployment, namely, isolated, overlapping and centralised or shared deployment and their associated business value emerged from the discussion. It provided a comprehensive picture of how on-premise BI deployment supports decision-making processes. These days, organisations are constantly seeking for the best ways to deploy BI solutions considering the availability of various channels of communication.

The purpose of this chapter is to explain another dimension of BI deployment by means of an investigation of cloud-based BI deployment as well as its associated business value. The various cloud-based BI deployment solutions will be discussed in order to contribute to a comprehensive understanding of BI. Furthermore, the chapter sheds more light on the second sub-research question.

This chapter is structured as follows: section 4.2 provides a general discussion on the foundation technology of cloud-based BI deployment, that is, cloud computing, standard cloud service models, cloud deployment models, and cloud pricing models. Then section 4.3 proposes and discusses a potential cloud-based BI model which is followed by an ideal architecture accompanied with sub-architectures for possible implementation options in section 4.4. Also covered in this chapter are the discussions on the advantages, challenges, and related TCO of a cloud-based BI deployment in sections 4.5, 4.6, and 4.7, respectively. This chapter will be concluded with a summary discussion in section 4.8.

4.2 Foundation of cloud-based business intelligence

4.2.1 Cloud computing

Cloud computing has been defined differently by technology experts, service providers, IT research firms, and so on. (Kalyvas, Overly & Karlyn 2013a). It has been defined as delivery of on-demand IT resources, such as, software, platform, or

infrastructure resources provided as services over the Internet. Cloud computing provides on-demand software services, application platforms, and infrastructure resources over the Internet (Marinela et al 2011:2). It is an Internet-based provision of shared computing resources to client organisations through technological devices on request (Ouf & Nasr 2012:651).

Cloud computing provides reliable computing resources with multi-tenancy capabilities at reduced cost and has robust security features when users access its Internet-based resources (Paul, Govindarajan & Chatterjee 2013:24). Furthermore, cloud computing provides elastic resources that allow scaling up or down quickly and easily to meet demand with self-service IT resources that organisations can pay for based on what they use (IBM 2015:1). This essentially means that organisations need not worry about the installation and maintenance of computing resources, such as, Google Apps Engine, Amazon Web Services and Microsoft Azure Services as they are managed by the service providers and are readily available from the Internet. To further illustrate cloud computing, the next section discusses the three types of cloud service models.

4.2.2 Cloud service models

In the context of this study, cloud computing provides a variety of computing resources or standard service models for delivering BI over the Internet in three categories as illustrated in figure 4-1 below, namely, Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS) (Yang 2012:4; Garrison et al 2012:63). These cloud service models are the current state of the art technologies that ensure that organisations move their business technology resources from on-premises to the cloud.

As with any emerging IT area, the move to the utilisation of cloud service models must be done cautiously by paying attention to the sensitivity of data, security and privacy requirements (Jansen & Grance 2011:7). The organisation's scope and control over the cloud computing resources (such as applications, platform architecture, virtualised infrastructure, hardware and facility) is determined by the service model that they choose. Figure 4-1 illustrates the cloud service models that organisations can consider for implementation.

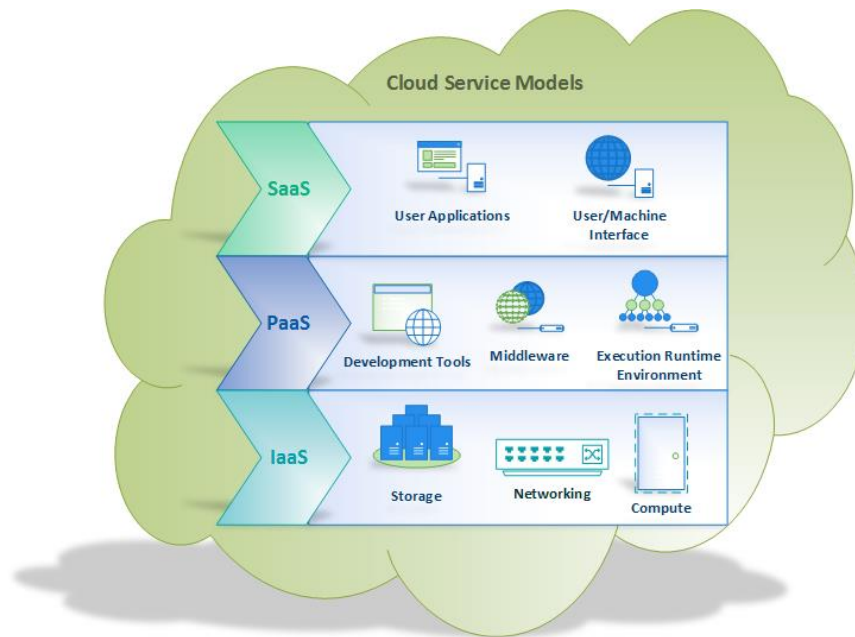


Figure 4-1: Cloud Services Models

4.2.2.1 Software-as-a-Service

SaaS provides pre-defined applications as a service that are accessible through technological devices connected to the Internet (Boniface, Nasser, Papay, Phillips, Servin, Yang, Zlatev, Gogouvitis, Katsaros, Konstanteli, Kousiouris, Menychats & Kyriazis 2010:155), where these applications are owned and managed in the cloud by one or various software vendors (Gartner 2014). In a SaaS service model, cloud-based BI applications or system software are provided as services over the Internet through specialised and centralised servers residing in a data centre (Bucur 2012:11; Kalyvas et al 2013a:9). Initially such kinds of cloud-based solutions were designed and best suited for small- and medium-sized organisations in order to have access to fast computing resources and increased response time at low cost (Marinela et al 2011:2).

SaaS services are ideally provided to reduce the total cost of managing similar services in an on-premise deployment (Jansen & Grance 2011:4). SaaS consists of computing resources that removes the need to constantly support and maintain software on-premise. Some of the good examples are MicroStrategy's MicroStrategy

Cloud, IBM's Cognos Express, Salesforce.com's Unlimited CRM, and Google Apps from Google (Phatak & Kamalesh 2010:12; Chen et al 2012:1177).

In addition to this, the SaaS-based applications discussed above are believed to have potential cost reductions when it comes to up-front IT infrastructure investments as the service provider will set up and maintain the overall infrastructure (Bibi et al 2012:87). Although most organisations prefer not to move their mission critical applications to the cloud, SaaS-based solutions offer simpler ways to implement and manage critical business applications, such as, ERP, CRM, e-commerce, and so forth (Bibi et al 2012:87; Suciu, Halunga, Apostu, Vulpe & Todoran 2013:110). Furthermore, this is because the burden of handling any application installation, configurations, modifications, on-going operations and support, be it small or large, becomes the responsibility of the application's vendor (Liyang et al 2011:1025), and in most cases such changes may require significant planning and testing to be done by different IT departments.

4.2.2.2 Platform-as-a-Service

PaaS provides a cloud-based middleware, development platform and environment that supports building and delivering web-based applications, but users have no control of all these computing resources (IBM 2015:2; Boniface et al 2010:155; Mavodza, 2013:134). PaaS provides the platform that allows BI service providers or business users to customise reports and applications through the web browser connected to the Internet from anywhere in the world (Kaur et al 2013:24). It removes the burden of purchasing and managing costly software development environment in-house (Jansen & Grance 2011:4).

Furthermore, PaaS enables organisations to fully manage the platform's application configurations. Google App Engine, and Microsoft's Azure are some of the best examples that PaaS delivers as a computing platform that allows organisations to support cloud-based applications (Phatak & Kamalesh 2010:12; Chen et al 2012:1177).

4.2.2.3 Infrastructure-as-a-Service

IaaS allows users to have maximum control of the infrastructure by having the ability to control and manage storage, networks, computing power, and so on as if they own the underlying physical infrastructure which is managed by the service provider (Han 2013:88). It removes the need to acquire such computing resources by accessing them virtually through a service interface (Jansen & Grance 2011:4). As a platform for the virtualised environment, IaaS offers and delivers a computing infrastructure, such as, Amazon.com's Elastic Compute Cloud (EC2), IBM's Smart Analytics Cloud (Phatak & Kamalesh 2010:12; Menon & Rehani 2012:7; Parziale, Avramenko, Chan, De Valence, Dziekan, Dziekan, Greggo, Hagen, Lin, D, Machung & Roik 2010:24).

Organisations are constantly faced with the market pressure to launch new products and services quickly and to have applications that have functionality which meets the on-going demand. Cloud-based BI offer applications with the ability to scale up and down without restricting customisation on applications and keeps the cost of operations low (Bibi et al 2012:87; Kalyvas et al 2013a:10). Cloud-based BI form part of leading SaaS BI packages, such as, SAP Business Suite that have the capability to integrate information processing in the same, or various organisations to cater for a wide range of different customer groups and demands (Plattener & Zeier 2012:22). These packages come with the ability to create industry specific solutions as organisations are able to determine how to query their data, modify the underlying data stores and customise their business processes to meet customer needs.

Most of the small and medium-sized organisations choose cloud-based BI because they are cheaper and often have insufficient IT computing resources such as BI software, storage servers or have minimal BI expertise (Kaur et al 2013:24). However, large organisations use the emerging technology in order to stay ahead of their competitors and avoid implementation of costly traditional on-premise BI solutions (TEC 2012:6).

Organisations are often faced with the challenge of choosing the right deployment model for implementing cloud-based BI solutions. In the next sub-sections the different types of cloud deployment models are discussed.

4.2.3 Cloud deployment models

Cloud computing solutions are accessible to organisations through four possible deployment models, that is, public, private, hybrid, or community deployments. These deployment models illustrate the way in which cloud computing resources can be provided to organisations by way of defining the business model between the user and vendor (Gastermann et al 2015:577; Jansen & Grance 2011:3).

4.2.3.1 *Public cloud deployment*

A public cloud provides its cloud computing resources to the users over the Internet (Gastermann et al 2015:577; Mell & Grance 2011:3). It does so in three broad classes as (1) at no cost to users but the cost is recovered through advertisements, (2) based on a fee and free of advertisements and (3) based on a fee with negotiable terms (Jansen & Grance 2011:6). A Public cloud in this study provides affordable cloud computing resources over the Internet where data and BI applications are readily available for use by any organisation (Watson 2012b:5). This is shown in figure 4-2 below. With public cloud computing resources, business users are not required to purchase hardware, software or supporting technological resources that is managed and owned by the cloud service provider (IBM 2015:4). Furthermore, business users are enabled to use innovative SaaS BI applications with flexible and scalable IaaS for storage, and powerful PaaS for cloud application development and deployments.

Public cloud provided through the IaaS service model enable business users to have full control and management of the BI applications and data storage whenever they access these computing resources through the Internet. However, this kind of open deployment is accompanied by some challenges, such as, data protection, security as well as data integration between public clouds (TEC 2015:12; Halper 2014). It is, therefore, recommended that this approach be carefully considered due to the sensitivity of data to be stored in an open computing environment (Jansen & Grance 2011:25).

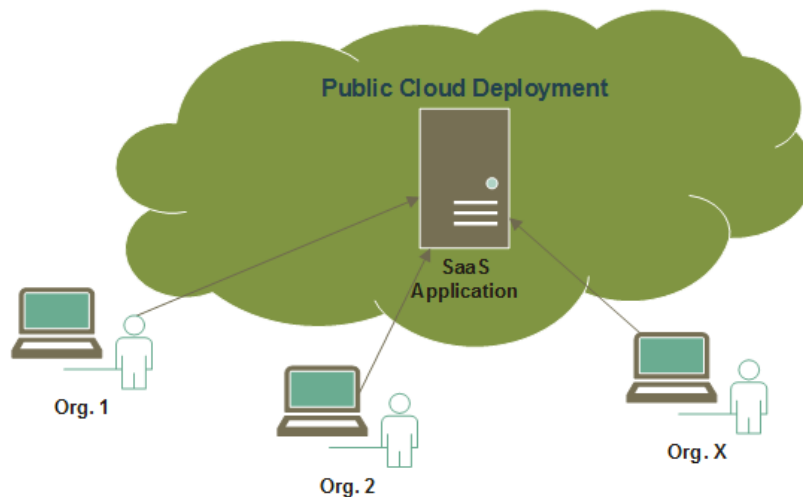


Figure 4-2: Public Cloud Deployment

a. Advantages of public cloud deployments

The following are some of the advantages of utilising public cloud computing (IBM 2015:5; Jansen & Grance 2011:9; Menon & Rehani 2012:7):

- There is often increased staff specialisation as staff becomes more focused on certain key areas of high concern to the organisation, such as, privacy and security rather than on executing several duties.
- There is greater resource availability due to built-in redundancy and disaster recovery capabilities.
- There are superior and robust backup and recovery policies and procedures compared to those that organisations implement on-premise.
- The risks of managing data in a public cloud are less when compared to distributing data to business users who use mobile computing devices.
- Powerful, flexible and scalable computing resources for application development, storage and computing services at the moment of request.

b. Limitations of public cloud deployments

The following are some of the limitations of utilising public cloud computing (Jansen & Grance 2011:6-11; Zimmerman 2014:89):

- When mission critical applications or data are moved from the organisation's secured resources to the cloud, there are risks that these computing resources may be interfered.
- Since public cloud computing target general users and has a variety of computing components involved, it has a large attack surface.
- The shared multi-tenant environment depends on logical separation which places organisations utilising public cloud services at high risk since they must share the infrastructure with other unknown consumers.
- Organisations lose control of the computing resources residing in the cloud as they transfer the responsibilities and control to the cloud provider.

4.2.3.2 *Private cloud deployment*

Private cloud deployment refers to computing resources operated entirely for one organisation and are often managed either internally or by a third party (IBM 2015:5), as illustrated in figure 4-3 below. In private cloud deployment, computing resources, such as, data storage and BI applications can either be hosted internally or externally, and users can use cloud-based technology to access these resources (Watson 2012b:5). Furthermore, private cloud computing provides SaaS BI deployment options where organisations can place their data warehouse or data mart and BI applications in the cloud that is only accessible to them.

PaaS provides the platform for business users in a private cloud to develop customised BI applications that are only accessible to them (MicroStrategy 2012:330). For example, Oracle Cloud PaaS which is powered by the industry's number one database and cloud application server, that is, Oracle Database and Oracle WebLogic Server, respectively allow Oracle engineers to use the platform to build Oracle's SaaS applications (Oracle 2014:14). Furthermore, the same cloud platform is used by business users to experience the unique advantages of application development which includes faster development of new applications and the ability to move application codes between the cloud and on-premises without implementing any code changes.

In a private cloud offering, organisations have the ability to design sophisticated security and governance features alongside a faster self-service interface delivering on-demand IT resources (IBM 2015:5). Private cloud overcomes the challenge in respect of data security faced by organisations using public cloud as it has the capability of implementing several layers of data protection, such as, password encryption and advanced security features that guarantee the requirements of the organisations being served (Rahman et al 2013:545).

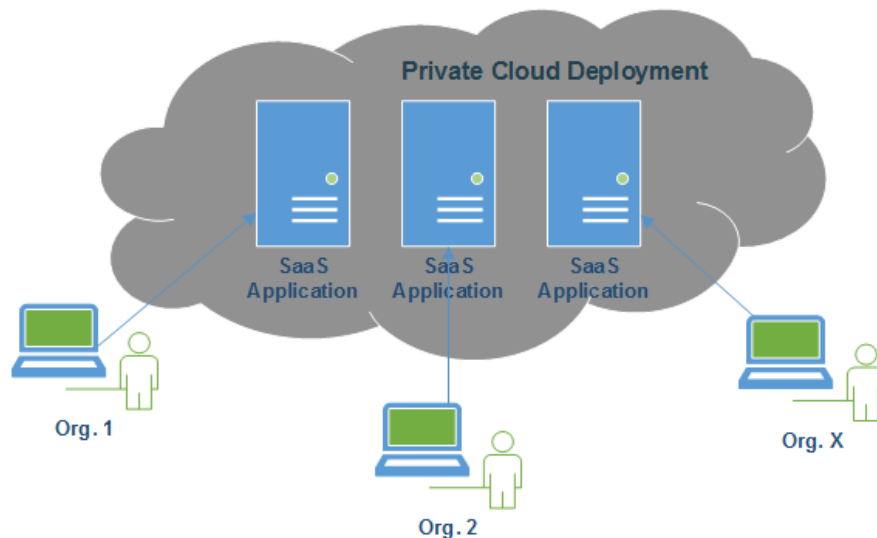


Figure 4-3: Private Cloud Deployment

a. Advantages of private cloud deployments

The following are some of the advantages of utilising private cloud computing:

- Organisations are guaranteed with additional levels of refined security and governance (IBM 2015:5).
- It offers superior inter-organisational relationships with the cloud service provider which often lowers the cost of negotiations (Garrison et al 2012:68).
- Organisations can host the BI applications internally or externally and the applications can be deployed to many business users within a short period of time, that is, in weeks without compromising BI functionality (MicroStrategy 2012:330).
- IT ensures there is higher consolidation densities and rapid self-provisioning of application and, or database services (Oracle 2014:20).

- Organisations are provided with more control over cloud computing resources as they are given dedicated resources (IBM 2015:5).

b. Limitations of private cloud deployments

The following are some of the limitations of utilising private cloud computing:

- Although most of the security concerns can be addressed, cost savings are limited (Garrison et al 2012:63).
- Despite most organisations outsourcing cloud based resources to manage sensitive information, safekeeping of such information is still one of the main objects of concern (Armbrust et al 2010:55).
- Some software licenses comes with limitation of the number of computers on which they can be installed, which results in organisations paying for the software and then having to keep paying on a yearly basis for maintenance (Armbrust et al 2010:58).

4.2.3.3 Community and hybrid cloud deployments

In this section, the co-existence of community and hybrid clouds has been defined, but they are less relevant in the context of this study. A community cloud is like a private cloud, but computing resources are allocated to only two or more organisations that have mutual agreement with regards to privacy, security, and regulatory concerns. On the other hand hybrid clouds are more sophisticated than the other deployment models because they comprise of a structure of two or more clouds, that is, private, community, or public clouds (Jansen & Grance 2011:3). The hybrid cloud combines the private and public cloud delivery modes in cases where organisations need the additional computing resources available through a public cloud although they mainly using the private cloud (Garrison et al 2012:63). In essence, the hybrid cloud, with its increased computing capacity and services, assists organisations to realise value by reducing IT costs.

a. Advantages of community/hybrid cloud deployments

The following are some of the advantages of utilising community or hybrid cloud computing (Garrison et al 2012:63; Paul et al 2013:24; Al-Roomi, Al-Ebrhim, Sabika, & Ahmad 2013:96):

- Organisations benefit from wider flexibility and efficiency from the combinations of two or more cloud service deployments.
- It allows organisations to keep critical applications and sensitive information in different cloud deployments.
- Cost and efficiency is much reduced.
- Multi-tenancy creates better sharing of computing resources.
- It provides better device and location independence.
- Organisations with similar objectives are able to share the same computing environment.

b. Limitations of community and hybrid cloud deployments

The following are some of the limitations of utilising community and hybrid cloud computing (Paul et al 2013:24):

- Security risks and threats, such as, loss of governance and data protection still remain a challenge.
- A single network failure can disrupt the entire community and hybrid cloud operations.
- The evaluation of a cloud solution presents a difficult process of analysing risks, associated costs, and benefits (Marinela et al 2011:10).

4.2.4 Cloud pricing models

Cloud computing offers a variety of cost structures that are designed to fit a number of business requirements. Cloud BI service providers offer several options that make it affordable to acquire BI solutions on a monthly or annual basis rather than organisations buying, configuring and maintaining BI solutions on-premise (Birst 2014:2). It is, therefore, important to be aware of what is involved in the current cloud

BI pricing or licensing models and some models that are currently offered are described below.

4.2.4.1 *Pay-as-you-go*

The pay-as-you-go pricing model allows organisations to pay only for computational resources that are deployed and used which often results in reduced cost (Chaudhuri et al 2011:98). This pricing models allows organisations to choose the best-of-breed BI solution and shut the solution down at any time if the business does not go as planned (Iyer & Henderson 2012:55). This also means that if, for some reason, the business is not performing well and the cost of the BI solution is not justifiable, an organisation can stop paying for the service and only resume later as, and when it becomes necessary.

4.2.4.2 *Pay-per-user*

Organisations in this pricing model only pay for active users each month and extra charges are applicable whenever new BI users are needed and added (Pentaho 2013:4). This pricing model is ideal for small-to-medium sized organisations that have a small handful of BI users in the team.

4.2.4.3 *Pay-per-use*

In this pricing model, organisations pay only for volumes of services used each month, such as, the number of transactions, and the size of storage and this prevents them from being locked into fixed long-term commitments (Kalyvas et al 2013b:25; Iyer & Henderson 2012:54). This pricing model is ideal for organisations that do not have high volumes of operational data.

4.2.4.4 *Server-based*

In a server-based licensing model, organisations only pay for the servers and there are no additional fees whenever they add more users. Thus, organisations have the capacity to allocate many BI users as they need depending on their server limitation (Pentaho 2013:4).

4.2.4.5 Subscription-based

Subscription-based pricing model allows users (organisations) to only pay for the services they require as and when needed, and based on a period of time. Thus, the organisations are in full control of the financial management of the service acquired and are able to scale up when the business needs changes or scale down when certain resources are not necessary anymore (Birst 2014:3; Kalyvas et al 2013a:11). In this pricing model, organisations have the opportunity of deciding on the duration of use of a BI service, and they only pay for that period.

4.3 Cloud-based business intelligence model

Based on the discussion in the previous section on the various cloud models, that is, service, deployment and pricing involved in the delivery of cloud-based BI solution, cloud-based BI deployment can be said to be largely reliant on the size and needs of the organisation. It is understood that Cloud BI can be provided as a service through the SaaS, PaaS and IaaS service models that can be deployed as public, private or hybrid clouds (Al-Roomi et al 2013:1; Garrison et al 2012:63; Jansen & Grance 2011:6). Despite the numerous benefits that come with the various cloud deployment models, organisations are still facing many challenges, as discussed in detail later in this chapter. The cloud-based BI model shown in figure 4-4 below demonstrates cloud-based solutions that enable BI and analytics to be accessed by different organisations through the Internet. The cloud-based BI model demonstrates the utilisation of various cloud-based computation resources. These include IaaS offering a variety of source systems and data stores for an unlimited number of organisations, and SaaS BI applications that are supported with the PaaS platform that allows business users to customise the BI services offered.

Furthermore, the cloud-based BI model illustrates how a number of organisations are using BI solutions. Some organisations have shown a reasonable level of cloud-based BI adoption as the solution is no longer considered risky and unmanageable. Furthermore they are cheaper to implement and operate (Liyang et al 2011:1025). An environment that has internet connectivity is all that is required to provide an organisation with access to cloud-based BI. Neither are there location restrictions.

The cloud-based BI model in figure 4-4 below is a simple demonstration of how different organisations are able to utilise cloud-based solutions that are related to BI optimally as discussed earlier.

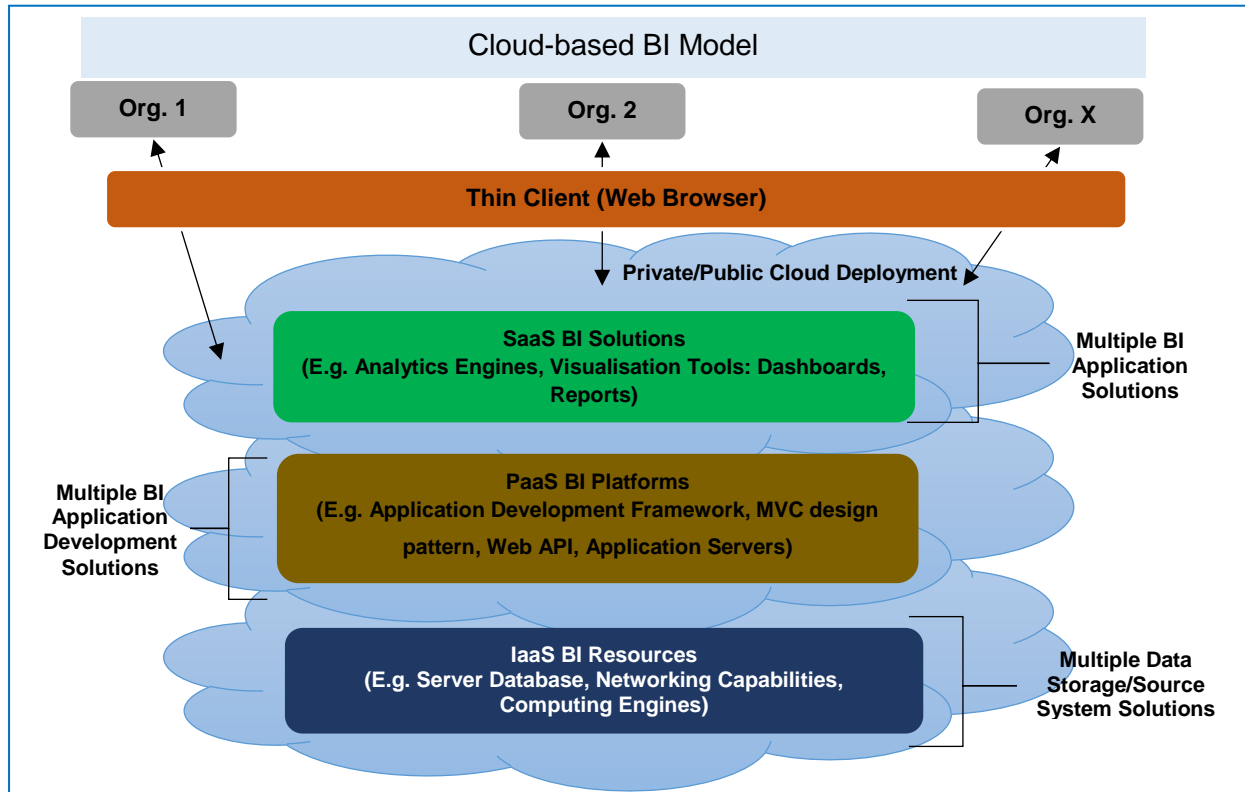


Figure 4-4: Cloud-based BI Model

Key:

Abbreviation	Description
Org.	An organisation (or business user within the organisation)
MVC	Model-View-Controller (MVC) is a fundamental design pattern for relating user interfaces to underlying data models such as ASP.NET MVC, the latest manifestation of Microsoft's web strategy released under Apache 2.0 license.
API	Application Programming Interface

The following sections discuss the components of the cloud-based BI model in detail.

4.3.1 Thin Client (Web browser)

The thin client is one of the crucial components of the cloud-based BI model. A thin client, also known as the web browser, allows organisations from different geographical locations to access BI application services provided by cloud vendors (Garrison et al 2012:63). The use of a thin client interface ensures that security considerations with firewalls, such as, the use of cookies are not an issue as storing of cookies on user devices is not permitted in most security policies (MicroStrategy 2012:276). For example, MicroStrategy Web Professional offers information consumers and power users a thin client Web interface that allows them to access PaaS development tools which enable them to modify reports and, or create new reports from scratch.

In the case of an organisation that has signed up for the use of SaaS and PaaS BI services, business users will only be able to access SaaS BI applications to conduct analysis, generate reports, and view dashboards through the use of a thin client (Liyang et al 2011:1026). At the same time, business users would be able to customise their SaaS BI application in order to meet changing business requirements through the use of PaaS application development and deployment tools. Business users will not have access to IaaS services that the cloud service provider may have on its service offerings.

In the case of an organisation that has signed up for all three cloud services, namely, SaaS, PaaS and IaaS business users will not only have access to reports, dashboards, and development tools, etc. found in SaaS and PaaS services, but will have access to IaaS BI resources that may include faster computing analytical engines and storage servers for storing all their business data (Kaur et al 2013:24). Therefore, depending on the level of cloud services signed for, business users can have single or multiple access to the three types of cloud-based BI solutions demonstrated in the model.

There are other comprehensive cloud-based BI platforms that are open-source, such as, the BI platform from Pentaho found in the Niche Players of the Magic Quadrant of BI and Analytics Platforms (Sallam et al 2015:6). Pentaho offers cloud-based solutions that comprise ETL, BI applications and analytics applications that are accessible

through a thin client and managed through a public or private cloud (Hagerty et al 2012:22). To make the use of a thin client much easier, cloud-based BI vendors, such as, Pentaho, SAP, and IBM, continue to invest and introduce user friendly applications such as, Business Analytics, BusinessObjects Cloud, and Cognos Analytics respectively, which make it easier to customise reports and dashboards.

In the cloud-based BI model, the thin client facilitates the utilisation of various BI components from different or the same vendors. For example, organisations 1 and 2 are both able to use the same SaaS BI solution, such as, SAP BI or Oracle BI solutions. However, each organisation will be configured differently according to business requirements. SaaS BI solutions are designed in such a way that an unlimited number of organisations and their business users can have access to the BI solution through the thin client at anytime and from anywhere as long as there is internet connection.

4.3.2 Multiple business intelligence solutions

As it was understood from the previous sections, cloud-based solutions allow the implementation of various Web services that come in forms, such as, SaaS, PaaS or IaaS solutions. When making the right business decisions, organisations in the cloud-based model often need to consider and include information from both their customers and suppliers. This can be achieved by using SaaS BI applications, such as, Tableau Online, SAP Business Objects, and BI On Demand, which are easy to set up, easy to integrate, and easy to use (TEC 2011:74). The use of such BI solutions is made possible through the internet as the cloud provides the best platform to collaborate easily with customers and suppliers since collaboration cannot be prevented by the organisation's firewalls.

In the cloud-based model, PaaS provides the environment from which business users are able to develop in-house, customise and deploy their signed up SaaS BI applications, such as, Power BI for Office 365, SAP Business Objects BI On Demand, IBM Cognos Express, Pentaho Business Analytics, Tableau Online, and so forth (TEC 2015:23). These applications have been proved to provide quality information, ease of use, are quick to implement, and are readily accessible from anywhere. Although

some research shows a slight decline in the interest in cloud-based BI, plans are in place to deploy some form of private, public or hybrid cloud BI with particular emphasis on the types of business of which data for analysis is already in the cloud (Sallam et al 2015:3).

The cloud-based BI model utilises the cloud infrastructure best known as IaaS that allows BI applications to be installed and deployed in the cloud easily at lower cost (Ouf & Nasr 2012:652). This enables a large number of organisations (Org. 1, Org. 2, etc.) to have access to cheaper public, private, or hybrid cloud-based BI applications with ease of use. The model further demonstrates that various BI solutions from different vendors can be provided, to satisfy the needs of different organisations from small to large scale.

In the model, organisations that might not be able to afford the cloud-based BI services offered by IBM, SAP, and Oracle, have the opportunity to use other BI applications, such as, Power BI, and Pentaho Business Analytics (an open-source cloud-based BI platform). These applications can provide BI services that meet organisations' business requirements and they are charged based on their preferred pricing model, such as, pay-per-use or subscription basis (Plattener & Zeier 2012:220). The main challenge that organisations face is when the ideal BI application, cloud deployment model and service provider have to be selected (Madaleno & Marinheiro 2013:1). An extensive assessment process needs to be instituted to establish the correct information from which decisions can be made because some SaaS BI applications cannot be integrated seamlessly.

4.3.3 Multiple data source or storage systems

The use of cloud-based BI has additional benefits related to the optimal utilisation of IaaS storage resources which organisations can choose to purchase or rent and can be either capped and, limited storage, or unlimited storage (Mavodza 2013:134; Suci et al 2013:114). As organisation data grows in volume, most on-premise BI solutions have problems of managing large scale data, as such, a cloud-based BI solution becomes the only ideal option since organisations are able to see Return-*Before*-Investment. According to Rosslyn Analytics (2014:4), Return-*Before*-Investment is a

new business metric for measuring value from cloud-based solutions, meaning that organisations are able to calculate the total cost of upgrading on-premise BI solutions to cloud-based BI and determine within the organisation where value will be created by whom, and when.

In addition to the provision of network components, storage, servers, and so on, IaaS supports the management of organisations' data sources, that is, operation applications that may include ERP, SCM, HR, etc. by providing the infrastructure in the cloud (Yang 2012:4). Most organisations use IaaS because of its features that support data and BI applications integration in the cloud (Han 2013:88). Organisations literally have full control of the environment as if they own the underlying infrastructure.

In order to further understand how the cloud-based BI model works, the next section will discuss the architecture that underpins the deployment of a cloud-based BI solution.

4.4 Cloud-based business intelligence architecture

The cloud-based BI architecture in this section demonstrates how the utilisation of various cloud service models (SaaS, PaaS and IaaS) can be made possible for any organisations residing anywhere in the world.

With the discussion on the cloud-based BI model in mind, and in order to achieve functional and technical understanding, an architecture was suggested consisting of four main layers, namely, the presentation layer, the network layer, the security layer, and the service layer. In other words, the architecture accommodates technology (computers, servers, network devices, operating systems, etc.), cloud-based business applications, and availability, security, consistency and reliability of information (TEC 2015:9). Figure 4-5 below depicts the overall architecture of which each layer (and components therein) is discussed in the following subsections.

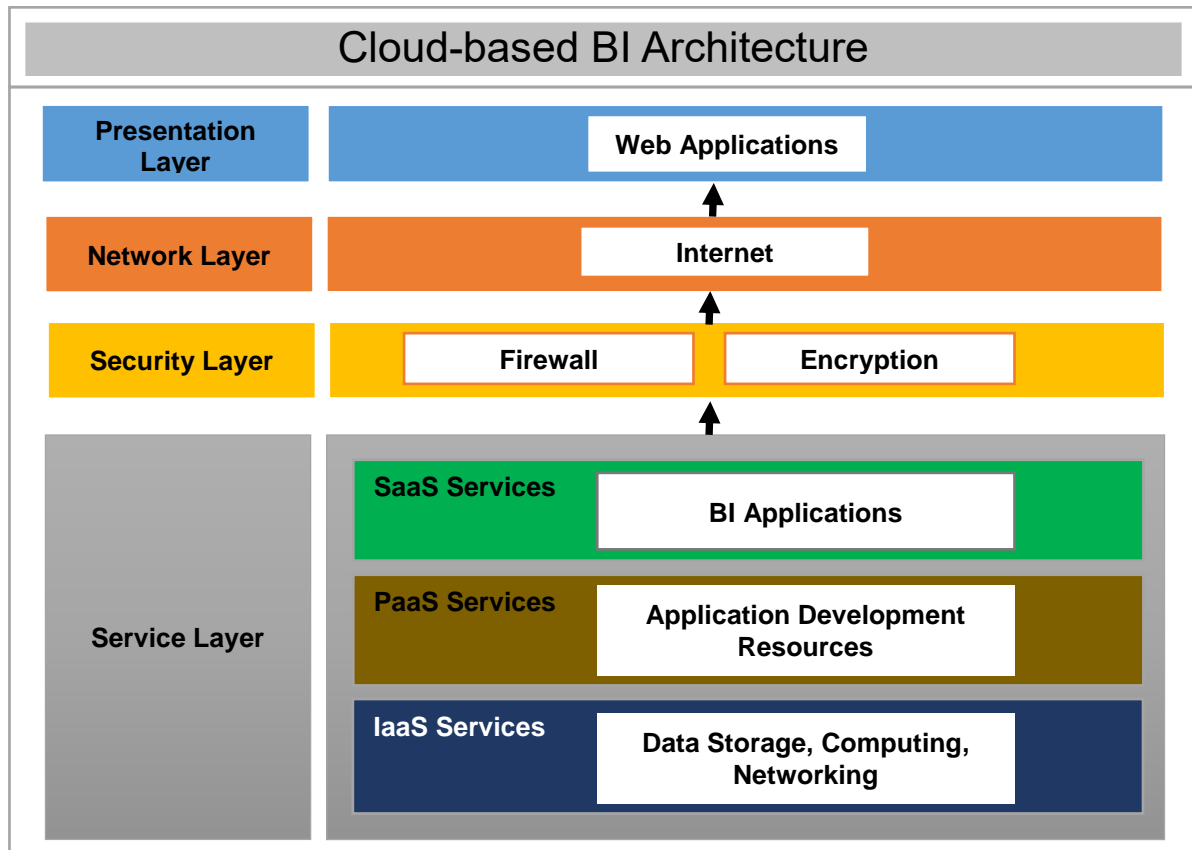


Figure 4-5: Cloud-based BI Architecture

4.4.1 Presentation layer

The presentation layer provides the organisations in the cloud-based BI model with the capabilities for web applications, such as, Windows Edge (formerly known as Internet Explorer), Google Chrome, Mozilla Firefox, and Apple’s Safari. to be used as thin clients to access a vendor’s online portal (Ali, Khan & Vasilakos 2015:360; Garrison et al 2012:63). The vendor’s online portal acts as the entry point for clients to access cloud-based services through the internet (Mavodza 2013:136). An online BI portal serves as a secure platform for applications where data emanating from various source systems is consolidated for the purpose of delivering different types of information that meet user needs (Ong et al 2011:7). When organisations in the cloud-based BI model have computing infrastructure that is connected to the internet, they can use electronic devices, such as, laptops, personal computers, mobile and smart phones, etc. to run thin client applications and gain access to specific cloud-based BI applications or portals.

Once business users in an organisation log into the vendor's online portal, depending on the user profile (power user, information consumers, etc.), they are able to use cloud-based BI services, such as, MicroStrategy Web, to conduct all necessary BI tasks which may include configuring or setting up reports, manipulating data, analysing data and, thereafter, generating reports for monitoring purposes (MicroStrategy 2012:87). It is through the same thin clients that business users are able to share information generated from the cloud-based BI applications. For example, once power users configure the reports, dashboards, etc. information consumers can access these features and share the outputs with other users within the organisations that do not have access but need to use the resulting information. Depending on usage requirements, organisations in the cloud-based BI model can have as many users that they need to access the vendor's portal through a thin client and can get charged on a pay-per-user or subscription basis (Kalyvas, Overly & Karlyn 2013b:25; Birst 2014:3).

4.4.2 Network layer

The network layer is one of the most critical layers for the deployment of cloud-based BI. The network layer facilitates the communication between the organisation and the cloud-based BI services from any place with internet connection (Watson & Wixom 2007:98). For organisations to fully realise the benefits that accompany cloud-based BI, the infrastructure for the network must provide the best internet connection (Siller 2013:34). An internet connection that is uninterrupted and is available constantly is one of the key components in the success of a cloud-based BI deployment.

The network layer provides the functions as defined in the Open Systems Interconnection network model or networking framework (Minoli 2008:107; Techopedia 2015). These functions include, the logical addressing of the Internet Protocol (IP) address; switching and routing technologies for transmitting data on the internet; error handling and diagnostics which involves exchanging hosting status information on the network and; selecting and managing the best logical path for data transfers, etc.

The network layer facilitates the utilisation of various interconnection networks, such as, Local-Area Network (LAN), Wireless Local-Area Network (WLAN), Wide-Area Network (WAN), etc. for the purpose of performing cloud-based computing tasks (TEC 2015:6). A LAN or Ethernet network is a communication technology based on physical cables which has a geographically limited communication capability (Minoli 2008:106). Organisations in the cloud-based BI model that have LAN networks connected to the internet usually experience the best internet connection with minimum signal interference. Wireless Local-Area Network (WLAN) is a wireless network based on radio waves instead of Ethernet cables, that is, physical cables, that connect two or more devices (Gartner 2015c). Most organisations use this type of connectivity as it allows flexibility, among other benefits, within the workplace.

As an alternative, the network could be outsourced as an IaaS solution where the organisations (business users) in the cloud-based BI model manage the network resources as though they own it (Han 2013:88). In addition to this, Jansen and Grance (2011:4) assert that a IaaS computing infrastructure can be provided as a service that can enable business users to develop and execute SaaS BI applications as and when needed.

4.4.3 Security layer

The security layer addresses the other critical elements, that is, securing the transmission of data through the Internet, Intranet, Extranet, LAN, or WLAN. Since cloud-based BI is an internet-based solution, the possibility exists that when data is being transmitted between the organisation and a cloud-based BI vendor, it may get intercepted (Suciu et al 2013:110). The security layer ensures that the correct security considerations, such as, firewalls and encryptions are implemented to facilitate the smooth transmission of encrypted data and prevent unauthorised access to data.

Once data is in the cloud, the onus is on the cloud-based BI vendor to protect and manage all the expectations of the organisations. Most cloud-based BI vendors have ensured that access to their web portal is implemented through a secure connection using Hypertext Transfer Protocol Secure (HTTPS) (MicroStrategy 2012:140). HTTPS is an Internet protocol that ensures that all transactions performed over the Internet

through the use of any thin client whatsoever, are meant for the intended purpose and do not allow caching (e.g. storing user name and password) on web browsers (TEC 2014:40; Gastermann et al 2015:581). Liu, Guo, Zhao & Chou (2010:407) assert that HTTPS ensures that an end-to-end connection is established between the cloud-based BI vendor and the organisation requesting the service. This essentially means that user authentication and authorisation processes take place between business users with the right privileges and cloud-based BI provider.

It is one thing to have a secure connection to the cloud-based BI service provider and another to prevent unauthorised access to the organisation when on the internet. Firewalls and antivirus software play a vital role in preventing unauthorised access to business information in an organisation that is using cloud-based solutions (MicroStrategy 2012:94). The Internet, which is a public network, presents a number of threats that can affect the organisation during internal and external communication with the vendor. Thus, the use of firewalls has become an essential feature.

However it must be noted that the use of secure cloud-based BI solutions will need more than just firewall and antivirus software protection. Other security controls should be considered. These include “malware protection, vulnerability assessment, network monitoring, data security protection, user authentication and authorisation” (Rahman & Choo 2014:50; MicroStrategy 2012:125). Security controls are a key consideration for successful implementation and operation of a cloud-based BI solution.

4.4.4 Service layer

The service layer in figure 4-5 is the main feature of a cloud-based solution about which organisations in the cloud-based model do not have to be concerned with regards to managing, as long as the other features mentioned above are addressed within the organisation. Cloud-based BI service providers have state of the art technology to ensure that their clients, be it individuals or organisations, get the best services whenever they access the service (Rahman et al 2013:542). The proposed architecture shown in figure 4-5 demonstrates a full package of SaaS, PaaS and IaaS services to which an organisation can subscribe. In this deployment an organisation is able to access SaaS BI applications for analysis and reporting, PaaS tools for

customising and developing their SaaS BI applications, and IaaS resources for networking, storage, and faster computing resources (Kaur et al 2013:24; Jansen & Grance 2011:4). In the following sub sections, the different types of options which organisations can choose to implement, depending on their BI business requirements, are discussed.

Firstly, an organisation that only needs SaaS BI applications may sign up for such on-demand services only and will not have access to development tools and, or platforms, storage, computing or networking resources. In such a SaaS BI deployment, an organisation would have an architecture as depicted in figure 4-6 below, where SaaS BI applications can be accessed through the use of a thin client to conduct analysis, generate reports, and view dashboards (Liyang et al 2011:1026). Other SaaS applications, such as, CRM from Salesforce.com can also be acquired and integrated with SaaS BI applications for easy access to business data for analysis and reporting.

The advantage of an organisation choosing SaaS services is that they are given the opportunity to use faster BI applications and they do not have to be concerned about the maintenance and operations costs of hardware and software (Jansen & Grance 2011:4). However, since the data still reside on-premise, they are often faced with the challenge of database performance issues as the SaaS BI application will always have to rely on the internal internet connective when accessing data. In the event that SaaS BI applications do not meet certain BI requirements and need to be customised, then the organisation would consider the option of acquiring development or customisation tools and that would involve PaaS BI services.

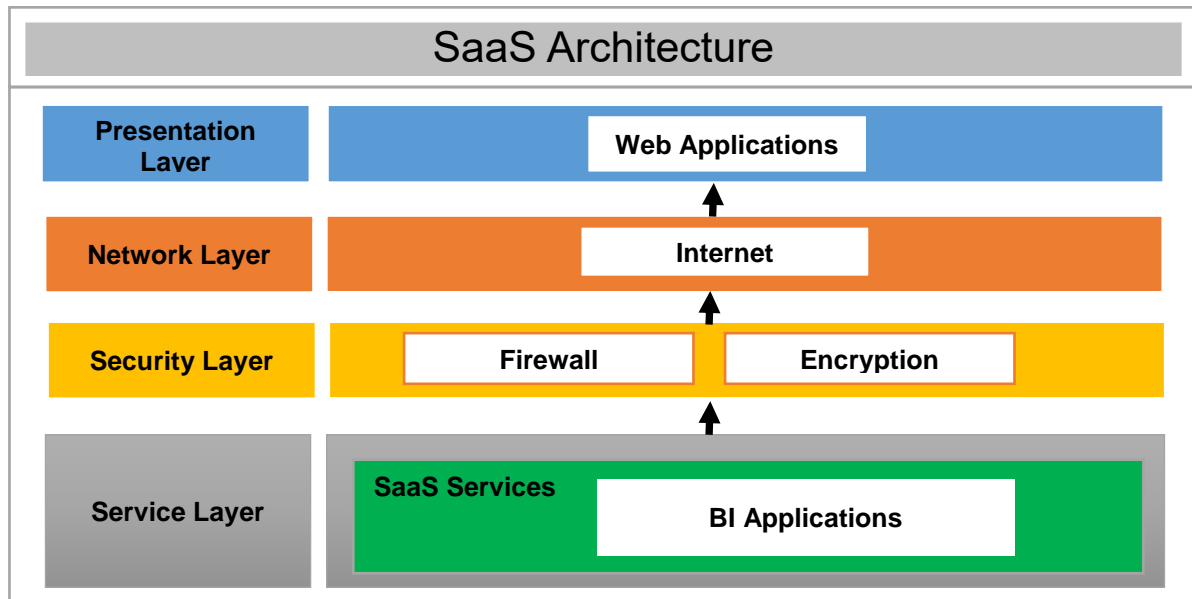


Figure 4-6: SaaS Architecture

Secondly, the service layer can offer organisations in the cloud-based BI model the opportunity to build new SaaS BI applications or customise them through a PaaS platform (Bibi et al 2012:86). The computation resources used in the cloud, such as, operating systems, integrated development environments, and the run-time engine, are provided as PaaS services and the organisations in the model do not manage the computation resources (Ali et al 2015:360). The SaaS and PaaS architecture shown in figure 4-7 below proposes how such a BI deployment could be implemented.

The advantage of having both SaaS and PaaS services is that business users are able to customise their SaaS BI application in order to meet changing business requirements through the use of PaaS application development and deployment tools (Mavodza 2013:135). However, business users will not have access to IaaS services hindering them from the opportunity of using faster networking resources, unlimited storage facilities and powerful computing resources that could be offered by the cloud service provider.

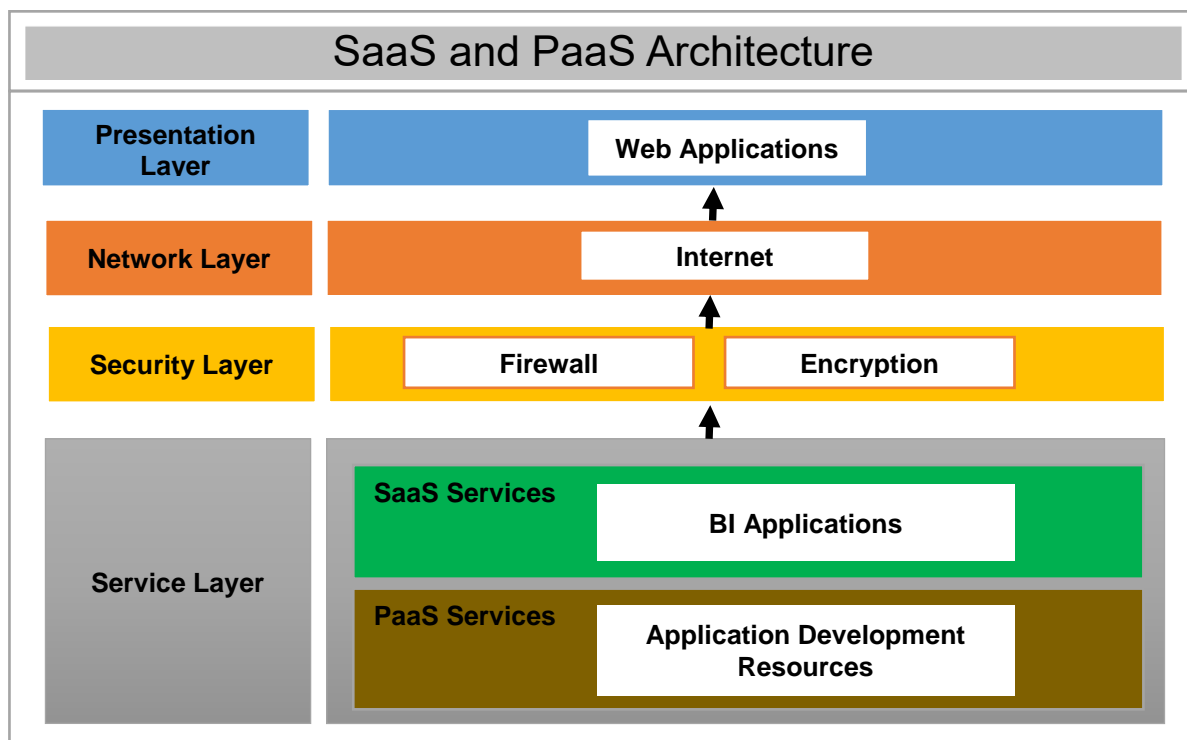


Figure 4-7: SaaS and PaaS Architecture

Thirdly, the service layer can offer robust IaaS computation resources that allow organisations to use servers or storage remotely in their data centres from where their SaaS BI applications data can be securely stored, accessed and paid for, based on usage (Mvelase et al 2012:215). The advantage of using such a deployment is that business users have full control of the SaaS BI applications, the cloud environment, and the remote electronic resources. In essence this cloud-based service offers computing resources or storage services that make it possible for BI applications to be provided as a cloud-based solution.

The only problem with such a deployment is that business users in the organisations are not able to customise or build new SaaS BI applications although they have full control of the applications and hosting environment (Chen et al 2012:1177). The lack of PaaS services hinders the organisation from adapting to a dynamic IT environment as they will have to wait for the cloud service provider to make changes that may have a positive impact on the business.

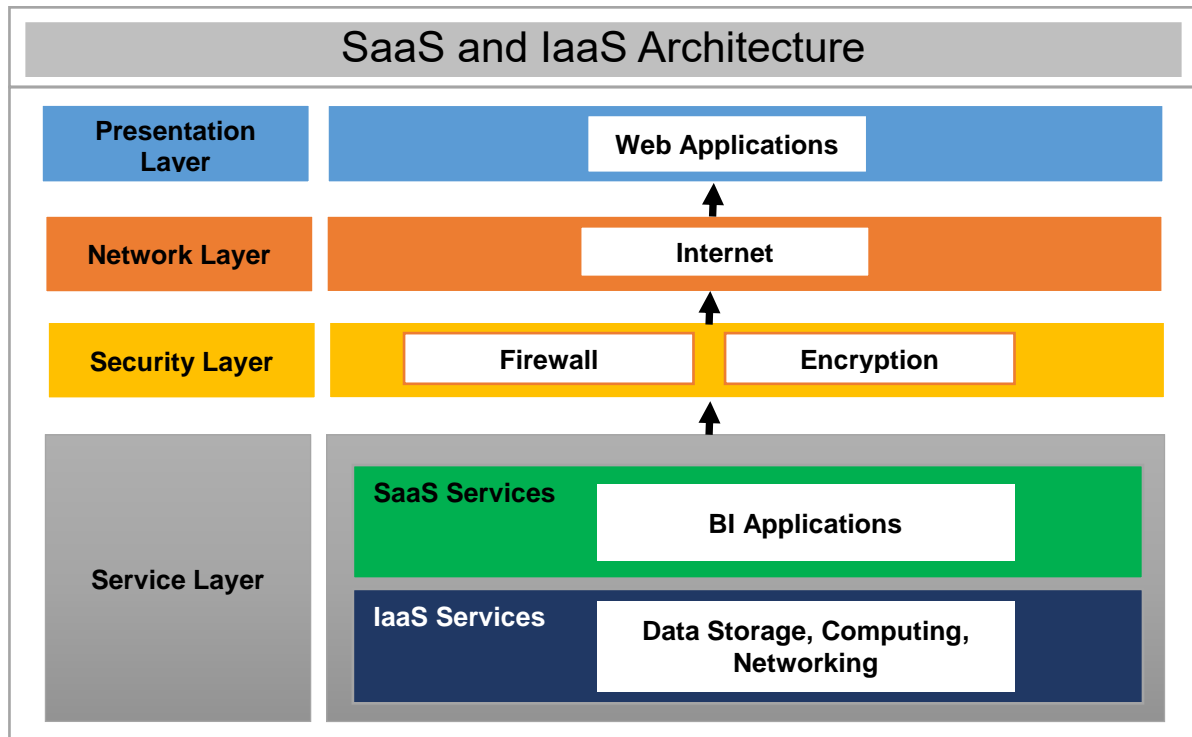


Figure 4-8: SaaS and IaaS Architecture

In summary, the organisations in the cloud-based model need a robust Internet network infrastructure, reliable network devices and a secured connection to be able to have unlimited access to cloud-based BI applications. This will ensure that an organisation’s data is kept safe, secure, and ready and is easily available to the appropriate business users. This is critical in any BI implementation as it will ensure that there is no unauthorised access (MicroStrategy 2012:141). Depending on the business requirements, organisations can sign up for a full package comprising SaaS, PaaS and IaaS services or select one or two services that can meet their BI requirements.

The next section discusses the advantages of deploying a cloud-based BI solution.

4.5 Advantages of cloud-based business intelligence

Cloud-based BI offerings provide a number of benefits irrespective of which cloud solution vendor is chosen. However, to find the best solution for the organisation, it is the responsibility of the organisation to consider a cloud vendor offering solutions that

meet their unique business needs, is cost-effective and maximises performance when it comes to supporting fact-based business decision-making. Some of the substantial business benefits from the use of cloud-based BI solutions are described below.

- ***Easy integration of cloud-based BI solutions:*** Organisations considering the integration of a cloud and BI solution must have a clearly defined strategy addressing all cloud computing capabilities (Marinela, Bogdan & Marian, 2011:4). The advantage, among other benefits, is that expanding in the cloud is accompanied by an unlimited computing capacity. This means that when planning for the initial capital investments, hardware expansions must be carefully considered (Gash et al 2011:264). There are many computing resources that make up a cloud-based BI solution, and the hardware and software that make up these resources can quickly add up to millions of Rands. However, thanks to the PaaS, IaaS, and SaaS service models, organisations using the cloud-based model do not have to be concerned about acquiring computation resources. The speed at which cloud-based BI solutions are implemented and deployed comes as an immediate benefit as there is no reliance on the time-consuming implementation phase related to on-premise BI solution procurement, and deployment (Menon & Rehani 2012:6).
- ***Advanced cloud technology resources:*** Organisations considering deploying cloud-based BI solutions can save not only on the costs of running their own data centre but can reap the benefits of the vast advanced technology resources available, such as, on-demand computing, and backup or disaster recovery compared to traditional on-premise solutions (Mvelase et al 2012:214). The advantages of utilising on-demand cloud services include faster processing power, reduced time to deploy, and realisation of the BI value proposition which is not the case with on-premise deployments (Ouf & Nasr 2012:653). Cloud on-demand computing provides functionality in various forms such as remote and, or mobile user browser-based access for performing all necessary BI tasks, that is, database management and analytics. This approach eliminates the need for organisations in the cloud-based model to hire experienced IT professionals to manage BI services (Kaur et al 2013:25), as the cloud-based solutions are looked after by the relevant vendors' well-trained IT and business professionals.

Most organisations believe that cloud-based BI solutions would lead to improved business flexibility, elasticity, parallel processing, accelerate response to market changes, exposure to massive computing power and concentration on the core capabilities of the business (Marinela et al 2011:2, Abadi 2009:2). This effectively results in better utilisation, performance and availability of the ever growing business data, as the burden of managing BI applications is outsourced and, or managed by the cloud vendors who have the capabilities of managing the such applications in the cloud (Abadi 2009:3).

- **Data an organisation asset in the cloud:** Business data has become an asset that organisations can use to strengthen their competitive advantage or intelligence (Karim 2011:198; Hewlett-Parkard 2012:4). With the right business data and the use of appropriate cloud-based BI, organisations in the model are able to stay a step ahead of their competitors. As a result, organisations are constantly looking for cloud-based tools that can assemble the right mix of data in order to produce quality data for reporting and decision-making processes (Mantha, 2014:10). In most cases, traditional on-premise BI tools are often upgraded or slowly integrated into cloud-based BI applications while considering TCO.
- **Lower Total Cost of Ownership:** In order to benefit from the lower TCO that comes with cloud-based BI solutions, most organisations convert some part of their capital expenditure to operational expenditure, and are largely utilising available cost-effective pricing options (as discussed earlier), such as, the pay-per-use model, which is provided by most public cloud BI vendors (Menon & Rehani 2012:6). The outcome of such an investment is that the TCO is reduced as no hardware and, or software to set up is required and, in most cases, a very modest use of internal IT resources is required. In addition to the initial capital investment of bringing the on-premise BI systems to the cloud, in some cases hardware expansions are required as BI becomes adopted throughout the organisation (Gash et al 2011). However, once the system is in the cloud, there is no limitation when it comes to expansion capabilities and computation resources.
- **Access to unlimited cloud resources:** In the cloud, organisations are no longer limited to discovering new information from existing cloud resources which is often

experienced in an on-premise BI deployment (Ali et al 2015:359). Organisations ensure that BI value is created through integrating multiple cloud-based data sources with BI applications that can provide analysis and reporting to support specific business processes, strategies and better business decision-making (Elbashir et al 2013). As far as access to, and utilisation of cloud-based software and hardware is concerned, organisations in the cloud-based BI model do not face the challenge of on-going upgrades or architectural changes after every three years as this is handled automatically by the cloud BI vendors as a value added service for the IaaS or SaaS BI solution.

- **Limited software and hardware investments required:** Organisations with inadequate budgets, and physical and virtual space opt for digital computation resources as it is practically impossible to procure every technological requirement that users need (Mavodza 2013:135). The use of SaaS, PaaS, or IaaS BI solutions in this dynamic business environment supports various methods to using the emerging cloud technology at no additional software and hardware costs. Although cloud-based BI solutions are still considered as high-risk projects that requires working together with the stakeholders involved, no up-front costs and investment in additional IT infrastructure are needed to realise business value from the investment (Wagner & Weitzel 2012).
- **Limited use of special expertise to operate:** Cloud-based BI are accessible to non-technical business users which ensures that BI applications are managed by users without relying on experts for insights (Birst 2014:3). In addition, multiple business users in the model with access to cloud-based BI applications across the organisation can discover prospects for performance and functional enhancements that are often not easily seen in the business data.
- **Easier and quicker expansion on demand:** Any organisation in the cloud-based BI model have the ability to scale up simply by adding more users, ask for more server space, or change the account type without having to buy additional computation resources as is the case in an on-premise deployment (TEC 2015:12; Birst 2014:4). Cloud-based BI applications are faster to implement on schedule

than on-premise BI deployment which leads to enhanced decision-making (Dimensional Research 2013:3).

- ***Least likely to have implementation cost overruns:*** Since the implementation of cloud-based BI occurs immediately upon registration with the cloud solution provider, implementation time is reduced significantly as it eliminates the days or weeks of preparing for an implementation project (Dimensional Research 2013:5; Menon & Rehani 2012:14). Cloud-based BI offers faster implementation compared to on-premise BI and is often done in less than three months (Birst 2013:2).
- ***Robust multi-tenant architecture model:*** Organisations in the cloud-based model take advantage of the resource pooling of the cloud-based BI computing platform as they are able to use or share the same BI application (TEC 2015:7). Cloud BI resource pooling leads to substantial cost savings as well as savings on maintenance services (Ouf & Nasr 2012:653). Business users sharing a BI computing platform may be part of the same organisation or from totally different organisations.
- ***Optimised cloud resource usage:*** The fact that cloud BI solutions are offered on a measured service basis allows organisations in the cloud-based BI model to use the acquired BI applications and other computing resources optimally (Ali et al 2015:359). In addition, this allows for cloud computing resources to be scaled up and down while being charged according to the pricing model that is appropriate to the organisation's needs.

All the above-mentioned benefits would be realised by the organisations in the cloud-based BI model if business users have the appropriate minimal IT skills. Garrison et al (2012:66) assert that business users' IT skills are crucial if the organisation is to achieve the benefits of cloud-based BI applications as they symbolise training, experience, and knowledge behind the operational and managerial capabilities (business and technical skills) needed for a successful cloud BI service deployment.

The next section discusses the limitations associated with implementing a cloud-based BI solution.

4.6 Limitations of cloud-based business intelligence

The various cloud computing models discussed in the earlier sections provide improved, optimised and cheaper services to the organisations in the cloud-based BI model. However, as with any good emerging technology cloud-based BI has its own disadvantages as far as the business of an organisation is concerned. A number of disadvantages worth noting are listed in no particular order of importance below.

- **Unclear service level agreements (SLAs):** Organisations in the cloud-based model often face the difficulty of understanding the full services they expect to receive from cloud solutions service providers (Pombriant 2012:8). This creates a major problem as misinterpretation of SLAs from cloud BI vendors may lead to an organisation losing data control and ownership (Yang 2012:6). Hence, all the advantages that come with SaaS, PaaS and IaaS computation resources will not be fully realised if comprehensive IT governance structures are not put in place and service delivery standards are not fully understood by the organisations concerned.
- **Vendor lock-in:** Vendor lock-in often occurs when organisations in the cloud-based BI model have misunderstood their SLAs which have resulted in them being locked into a set of cloud-based BI technologies dictated by a chosen vendor (Gash et al 2011:264). This is an issue which most organisations face as far as cloud-based BI solutions are concerned. Some current cloud-based BI applications and, or architecture do not support nor ensure data and application interoperability as well as portability (Phatak & Kamalesh 2010:12). This poses concerns for cloud-based BI software control especially if an organisation is locked into a long term engagement.
- **Data security concerns:** Most organisations are concerned with data security regarding the privacy, reliability and availability of their data utilised in the cloud (Kalyvas et al 2013b). It has emerged that for some organisations, the uncertainty of data security is something difficult to resolve, hindering the migration to the optimal use of cloud-based BI computation resources. This is because of the

limited availability of the correct information or the perception that data is stored at an untrustworthy cloud vendor's data centres, which might lead to the violation of the confidentiality of its users and others gaining unauthorised access to data (Abadi 2009:3; Menon & Rehani 2012:6). This requires that stringent and appropriate business rules should be in place when on-premise based data is moved to the cloud because there is a possibility of data security breach.

- **Higher inherent risk in cloud integration and configuration:** The integration and configuration of cloud-based BI indicates higher inherent risk than many information systems, and this is largely because of the extensiveness and complexity of the BI system (Elbashir et al 2013:4). In a cloud-based BI environment, there is higher inherent risk when it comes to consolidating mixed data sources, creating data warehouses, managing data quality, mining applications, working with BI applications on unstructured or semi-structured data, etc. (Bucur 2012:106). Therefore, organisations in the cloud-based BI model do not need to consider only BI technologies to realise business value but should rather view it as a base that will support the decision-making process.
- **Security concerns:** Gash et al (2011:264) assert that security is the major concern faced by organisations using cloud-based BI solutions. These security issues may be concerns about open source systems, data that travels from inside the organisation's walls to a third party cloud vendor, or even whether the organisation should put its trust in a third party vendor to manage a secure cloud environment. It is, therefore, imperative that there are no security misconfigurations, as a minor mistake can breach and compromise the entire cloud-based BI solution (Ali et al 2015:362). When organisations subcontract their BI solution from the cloud vendors, they hope and expect that their data will be kept safe in the cloud, the security settings need to be robust both during implementation and business operations.
- **Data stores performance issues:** Depending on the vendor being used, organisations could experience performance issues related to the data storage or warehouses (Boateng et al 2012:230). There could be substantial inactivity if cloud BI applications are utilising the data that resides on-premise, that is, mainly when

managing big data (Menon & Rehani 2012:7; Abadi 2009:4). Since organisations will always be extracting, transforming and loading data in the cloud-based data warehouse, data latency will always be experienced.

- **Lack of standardised pricing models:** Organisations are facing a major challenge as far as existing pricing models are concerned as there are no standardised pricing models (Menon & Rehani 2012:7). This affords cloud vendors with the opportunity of taking advantage of their customers (organisations or individuals) and it is often difficult for customers to select the right vendor who will provide value for money solutions.
- **Data ownership concerns:** Organisations have been found to overlook the implications of data ownership before moving their BI solution to the cloud (Kalyvas et al 2013a:11; Gash et al 2011:264). In an on-premise BI environment, data ownership concerns, especially highly sensitive data are managed in-house often by IT management within the organisation. In addition to this, it is often a challenge to deploy mission-critical applications in the cloud that will handle highly sensitive data if organisations must comply with security or regulatory laws or if data must not be lost (Watson 2012b:5). Organisations in the cloud-based model will need to run through some sort of risk mitigating process before deploying their BI applications in the cloud.
- **Data volume concerns:** Business data is growing rapidly and this comes with the challenge of transforming, storing and analysing large data volumes (Gash et al 2011:264). Furthermore, it should be emphasised that BI application performance within a cloud-based environment remains a work in progress as far as performance improvements are concerned as there will always be data loading and refreshing. During the process of integrating business data into the cloud from various sources (internal or external), organisations are often faced with the challenge of identifying and properly managing large volumes of data (Chaudhuri et al 2011:97).

- **Reliability concerns:** The non-availability of cloud-based BI services over a particular period of time poses a considerable challenge in organisations operating in a cloud environment (Phatak & Kamalesh 2010:11). The Microsoft Sidekick outage which led to millions of users losing their data permanently is a good example in this regard. It is, therefore, critical when choosing a cloud vendor to fully assess the availability, or to obtain a guarantee, of cloud-based BI services.
- **Multi-tenant attacks on communication infrastructure:** Organisations utilising cloud-based BI are exposed to attacks on the communication infrastructure due to the resource pooling capabilities of cloud-based BI services and this, in particular, affects the IaaS service model (Ali et al 2015:362). This is often because most users are granted with power-user privileges in order for them to have full control of the PaaS platform and this exposes the services to malicious use.
- **Repetitive tasks:** One of the common challenges faced by organisations with on-demand cloud services, is the task of repeating certain BI activities (application development and, or customisation, storing data) that were performed previously because organisations tend to release on-demand cloud services as soon as they are no longer using them because of the nature of the pay-per-use pricing model.

From the above-mentioned discussion on the challenges related to the use of cloud-based BI it is clear that organisations in the model to some extent still face similar challenges as in the case of on-premise BI deployment. It is important that organisations or individuals embarking on the utilisation of cloud-based BI are made aware of these challenges as it will help them to carry out operational and cost benefit analysis, and ultimately encourage them to deploy cloud-based BI solutions.

4.7 Total cost of ownership of cloud-based BI deployment

TCO in the cloud-based deployment model is believed to be largely dependent on the organisation's strategy, organisational needs and size. It has become clear that Cloud BI is provided as a service through SaaS, PaaS and IaaS service models that can be deployed as a public, private or hybrid cloud (Al-Roomi et al 2013:1; Garrison et al 2012:63; Jansen & Grance 2011:6). Similar to the on-premise model, there are a

number of strategies that an organisation can adopt in order to keep the TCO cost down in a cloud-based deployment. For example, the TCO will be reduced in cloud-based BI deployment due to the fact that most of the direct cost, indirect cost or related cost individual elements will not be applicable. Another example is that the costs related to hardware, consulting, storage, upgrades, etc. do not have to be considered because all these costs will be managed by the cloud service provider. In the case of the naïve model the above-mentioned costs would have been unaffordable. In comparison to the naïve and on-premise categories of users, the table below summarises the TCO associated with cloud BI deployment.

Table 4-1: TCO Cloud-based Deployment

Key metrics	Naïve	On-premise	Cloud
Direct costs (DC)			
• Cost of hardware (new servers, etc.)	X	✓	X
• Cost of middleware (operating system, database servers, Web servers)	X	✓	✓
• Cost of software licenses	X	✓	✓
• Cost of consulting services for implementation	X	✓	X
• Cost of staff hours to implement the project	X	✓	X
• Cost of software quality	X	✓	X
• Costs of changes needed in the network	X	✓	X
• Cost of additional storage	X	✓	✓
Indirect costs (IC)			
• Cost of upgrading related software	X	✓	X
• Cost of upgrading related hardware/middleware	X	✓	X
• Cost of customisation/continued application evolution (e.g. new end-user reports/dashboard)	X	✓	✓
• Impact on use of existing servers	X	✓	X
• Impact on existing network traffic	X	✓	X
• Impact on existing storage	X	✓	X
• Increase in support calls to IT	X	✓	X
• Increase in maintenance of hardware/middleware	X	✓	X
• Downtime cost (loss on productivity costs)	X	✓	✓
Related costs (RC)			
• Staffing costs	X	✓	X
• Training end-users	X	✓	✓
• Training BI professionals, e.g. business analysts, data scientists	X	✓	✓
• Training IT staff	X	✓	X
• Creating governance rules	X	✓ / X	X

Key metrics	Naïve	On-premise	Cloud
• Implementing governance rules	X	✓ / X	X
• Maintaining governance rules	X	✓ / X	X

Depending on the organisation's size, most of the cost elements for the cloud model are not applicable to TCO contribution because the cloud service provider manages the required services, and the costs are usually based on the selected subscription option and agreement. For example, the applicable staffing costs in the on-premise model are not applicable in the cloud as there will not be any need to employ more users as opposed to training the current employees on how to use the new cloud BI services. Therefore, cloud-based and on-premise deployment models have different needs and can be used for cross-comparison as they are both centred on the organisation as opposed to a naïve adoption of BI deployment.

4.8 Conclusion

The purpose of this chapter was to provide an understanding of BI deployment from a cloud-based BI deployment perspective. It discussed the various cloud-based BI deployment solutions as well as the associated business value.

In this chapter the term 'cloud computing' was firstly defined in order to clarify basic concepts. The literature review has shown that there are many sources which deal with cloud computing in terms of benefits, implications, potential uses, costs savings, etc. through the utilisation of cloud service models, such as, SaaS, PaaS and IaaS. This was followed by a discussion of the various cloud deployment models (private, public, hybrid, and community) in which BI applications could be implemented to best suit the business requirements of the organisations based on the various pricing models.

A discussion on a cloud-based BI deployment model which illustrated the utilisation of the key cloud services model, namely, PaaS, IaaS and SaaS followed. The model later informed the development of an ideal cloud-based BI architecture consisting of four (4) layers namely: the presentation, network, security and service layers. Finally, the

advantages, challenges and related TCO for implementing the cloud-based BI deployment were explained.

Having achieved the purpose of this chapter, it has been observed that organisations have the opportunity to choose an on-premise or cloud-based BI implementation depending on what model will best meet the organisations' business needs. Furthermore, considering the perception of cloud BI features, adopting a BI solution at a business level may indicate the beginning of a superior form of information civilisation for organisations. This means that organisations will be able to adapt to present economic and technological challenges when they utilise available cloud computing resources by means of sharing resources, improved (or increased speed of) business processes and use of integrated BI solutions.

This chapter addressed some of the challenges faced in the on-premise BI deployment, and also emphasised that it also has its own unique challenges that still need to be addressed. Therefore, the next chapter seeks to address the solution to challenges raised in this chapter and focuses on the discussion in the previous chapters on any BI deployment, be it on-premise or cloud-based deployment.

CHAPTER 5 HYBRID DEPLOYMENT OF BUSINESS INTELLIGENCE

5.1 Introduction

The previous chapters discussed in detail the commonly used BI deployments, namely, on-premise BI deployment which covered isolated, overlapping and centralised or shared deployment modes, and cloud-based BI deployment covering cloud service models (SaaS, IaaS, and PaaS), private and public deployment models, as well as their associated business values and challenges. These discussions provided a comprehensive picture of how BI can be deployed to support decision-making processes in an organisation.

The purpose of this chapter is to introduce a solution that will address all the challenges identified in previous chapters by suggesting a hybrid BI deployment approach (a combination of on-premise and cloud deployment) as well as identifying its associated business value. In doing so, an answer to the third sub-research question on an improved model of BI deployment that could resolve the limitations of existing models is provided. The suggested hybrid BI deployment model represents the main contribution of this research. The chapter discusses the various technology solutions that support the deployment of a hybrid BI solution, such as, in-memory data management, in-memory analytical and transactional processing, big data management, and hybrid data warehouses that are readily available to implement in order to manage the identified challenges.

The chapter is structured in the following manner: section 5.2 provides a general discussion on preliminary elements and the fundamental computing technologies for a successful hybrid BI deployment. Then section 5.3 discusses the overview of a hybrid BI solution, followed by section 5.4 which proposes and discusses a potential hybrid BI model which is followed by an ideal architecture for implementing it in section 5.5. Also covered in this chapter are the discussions on the advantages, challenges, and TCO for a hybrid BI deployment in sections 5.6, 5.7, and 5.8. This chapter will be concluded with a summary discussion in section 5.9.

5.2 Preliminaries

This section discusses the various technological foundations that have an influence on the performance of a BI deployment. It covers the technology elements such as in-memory data management, in-memory analytical and transactional processing, big data management and hybrid data management. Investigations in this regard could be relied upon to form the basis of the suggested model, and could help to demonstrate how limitations enumerated on existing models could be overcome.

5.2.1 In-memory data management

Although these days in-memory technology is often referred to as modern technology, but it has not been in full scale use on the market for a number of years (SAP 2013:1). Nevertheless, it has just lately become possible to implement in-memory technology on large scale (Sikka 2013:1; IBM 2012:3). In-memory technology can be defined as the process of pre-loading the entire datasets into a computer's random access memory, relieving the requirement for transferring data from storage to memory whenever a query is executed (Read 2014:2; SAP 2013:1). This advancement in real-time processing technology, eliminates the traditional way of processing business data, that is, batch-oriented and workload-specific to real-time processing without boundaries within a specified window of opportunity (Sikka 2013:5). Acker et al (2011:129) believe that this virtual real-time technology allows operational data to be held in a single database and has capabilities of managing all the customer transactions, updates, and analytical requests.

For the purpose of this study, in-memory technology is associated with optimising input and output access for database management systems in the hybrid BI deployment, by placing data nearer to the central processing unit, that is, by pre-storing frequently accessed data into the main memory, and leaving disk search operations unnecessary whenever a query is executed. IBM (2012:2) confirms that in-memory processing technology reduces or eliminates the need for disk input and output which ultimately improves performance and is believed to have capabilities of powering interactive visualisations of large multidimensional datasets without the need to move data into specialised databases. In addition, in-memory technology's application development platform offers the ability to customise applications and make them work faster, and

allow easier processing and reporting on data in real-time (Forrester 2015:1). Figure 5-1 below is an example of a traditional disk-based query operation versus an in-memory query operation.

Read (2013:47) illustrates with the example in figure 5-1 that the traditional query approach is triggered by a data request from the user or BI application. The database assesses if the in-memory cache has data. If not found, the database generates a retrieve data from disk request. “The results are aggregated and sent back to the user. If the information is not exactly what the user wants, they re-issue another request and start the process again. This can be a very slow process and makes *ad hoc* analysis and scenario modelling very difficult”. By contrast, with in-memory query approach, the data needed for analysis is readily available and cached the memory. When the user issues requests for data, they get an immediate response interaction and the results. This ensures that users are able to analyse business situations in real-time.

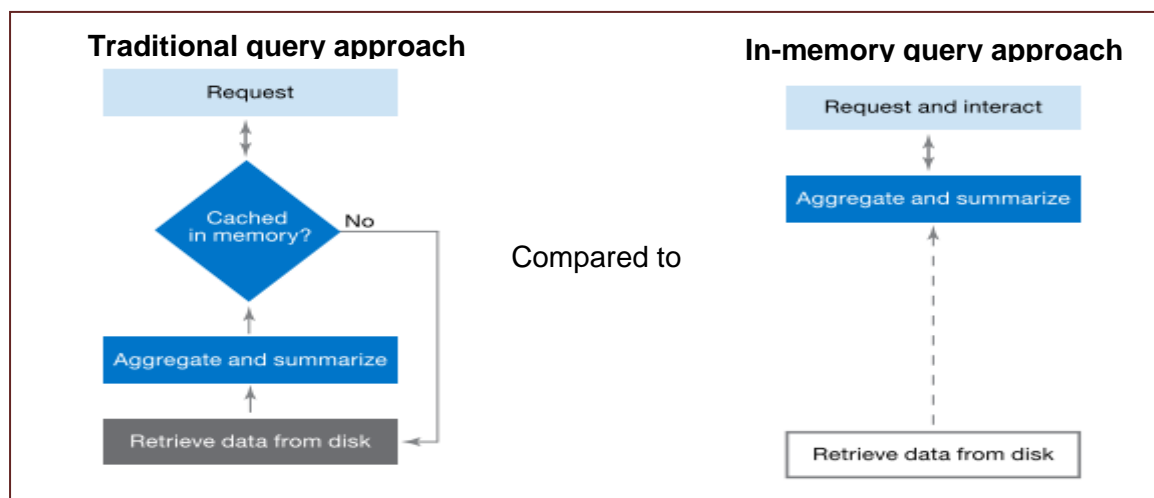


Figure 5-1: Traditional Query vs In-memory Query Approach (sourced from Read 2013:47)

From this query approach example, it can be seen that in-memory multidimensional platforms retrieve data from the disk once, and store the information in memory permanently. This provides and promotes superior analytical performance compared to traditional query approach which requires temporary storing of data in memory whenever needed.

Another consideration of using in-memory technology is one based on open source technology with analytic tools and execution engines powered by Apache Hadoop,

MapReduce, HBase and Hive, such as, Pentaho and Cloudera. Open source MapReduce is a programming framework or compiler that can give multi-threaded parallel processing to hand-coded routines (that is, procedures, functions or subroutines) that are coded in various programming languages (Russom 2013b:30). For example, a typical analytic application can be a written logic code in Java, Pig, or R routines. Then MapReduce execution engine can execute the routines and retrieve the data from the database which is stored in a Hadoop Distributed File System (HDFS) server cluster (Rahman et al 2013:544). When open source MapReduce is deployed as an execution engine layer on top of the HDFS in this way, the result is a high-performance analytic application with lower cost processing of massive data.

In addition, the commonly used parallel DBMS demonstrates clear advantages in well-organised query execution, whereas MapReduce excels in the execution of the ETL processes and analysis of unstructured data. The current innovative Hadoop and MapReduce-based systems, have grown into possible alternative solutions for analytics besides the commonly used systems, such as, Relational DBMS, column-based DBMS, in-memory DBMS, and parallel DBMS (Chaudhuri et al 2011:11; Chen et al 2012:1175).

In-memory data management must not only be viewed as a technology but as an alternative consideration in software development by addressing issues related to reading and writing data stored in memory versus disk, and the possibility of using parallel processing (Plattener & Zeier 2012:21). On the other hand, with the option of open source solutions, organisations that cannot afford the proprietary in-memory data management solutions from vendors, such as, SAS or Oracle can also go for an open source implementation with a Pentaho solution that uses MapReduce or Hadoop. Open source solutions are a cheaper viable option to increase performance, create unique, custom-built applications to meet the distinctive needs of organisations, and enable integration across the enterprise (Forrester 2015:1).

5.2.2 In-memory analytical and transactional processing

The other capabilities that have come with in-memory technology is that of having OLAP and OLTP in the same application component (Plattener & Zeier 2012:235).

While centralised data warehouses has capabilities of consolidating data from various sources, the benefit of having both OLTP and OLAP capabilities in one system can add more value to users (Plattner 2009:1). However, this is not the case anymore as some BI solutions have the capability of housing both OLAP and OLTP in one solution with various deployments models (SAP 2015:11). These solutions are based on a unified data model that eliminates the redundant layers often required in traditional RDMS and this reduces TCO. These types of solutions have eliminated the challenges of data getting locked during updates, and possible deadlocks during parallel transactions.

Plattner & Zeier (2012:235) assert that since speedy reporting can be accomplished instantly by placing an in-memory database system near the OLAP and OLTP, organisations can easily switch off commonly used OLAP and OLTP capabilities, leaving a high-performance in-memory database system. In addition, this consolidated in-memory system allows for new applications to be added either in the cloud or on-premise to accommodate any new functions that the business may need.

Figure 5-2 illustrates an integrated approach of using in-memory analytics that speeds up the handling of operational data. Acker et al (2011:130) have demonstrated in figure 5-2 below that major advances have been achieved “in the speed, sophistication of storage, memory technology, and in the power of processors”. Most of the current in-memory applications make is feasible to have access to the entire operational data and instantly perform analytical functions for reporting purpose.

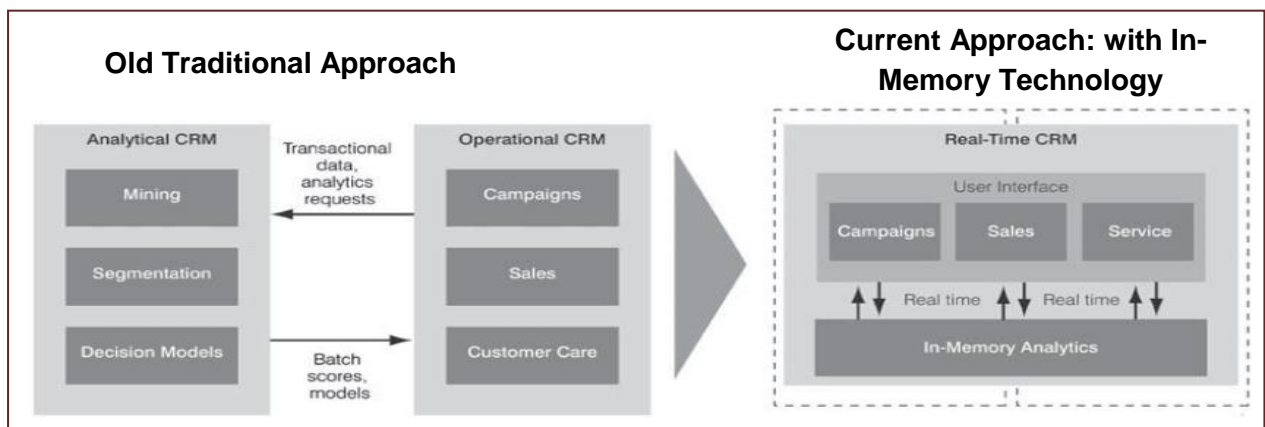


Figure 5-2: Integrated Analytical Approach (sourced from Acker et al 2011:130)

In-memory technology has transformed the way organisations are analysing and using their data, through providing instant access to real-time operational data. By itself, in-memory analytical and transactional processing comes with three important benefits compared to traditional data warehouses:

- **Performance improvements:** There is a dramatic improvement in response time and calculation performance because users work with, and retrieve data in memory according to Acker et al (2011:130) and Sikka (2013:10). This increase in performance allows business users to execute difficult queries and get superior modelling capabilities, resulting in better business benefits.
- **Customer value creation:** Acker et al (2011:130) and Plattner and Zeier (2012:208) assert that in-memory analytics when applied effectively has the capability of bringing information to users' fingertips through self-service features, and provides an innovative way of understanding customers which can be used to improve primary business processes, such as, when handling a query in a customer-facing process.
- **Lower costs:** When compared to traditional data warehouses, there is lower TCO and cost for managing storage resources as more data can now be stored in one place, that is in memory which, in return, reduces the need to create data marts (Acker et al 2011:131). As a result of the merging of OLTP and OLAP solutions, with their associated hardware and software, organisations are able to significantly reduce the TCO of managing client data resources. It is also easier to implement and manage the process of releasing IT resources that were previously dedicated to respond to report generation.

5.2.3 Big data management

As already shown in earlier chapters organisations are faced with the challenge of managing large volumes (terabytes) of unstructured, semi-structured and structured data also well-known as big data that is flowing into their data stores. Big data refers to a collection data the size of which is beyond the management capability of traditional relational DBMS (Manyika et al 2011:1). It is also anticipated that the advancement in technology in future will affect and increase the type of data that is classified as big data. Big data needs to be managed properly so that its full potential can be realised.

Big data management involves the application of data management disciplines, applications, and platforms in order to process and deliver collected and stored multi-structured data, that is, old data, new big data, or both (Russom 2013:5). Figure 5-3 below illustrates a traditional data integration and analytics architecture for data management (Henry & Nys 2013:7).

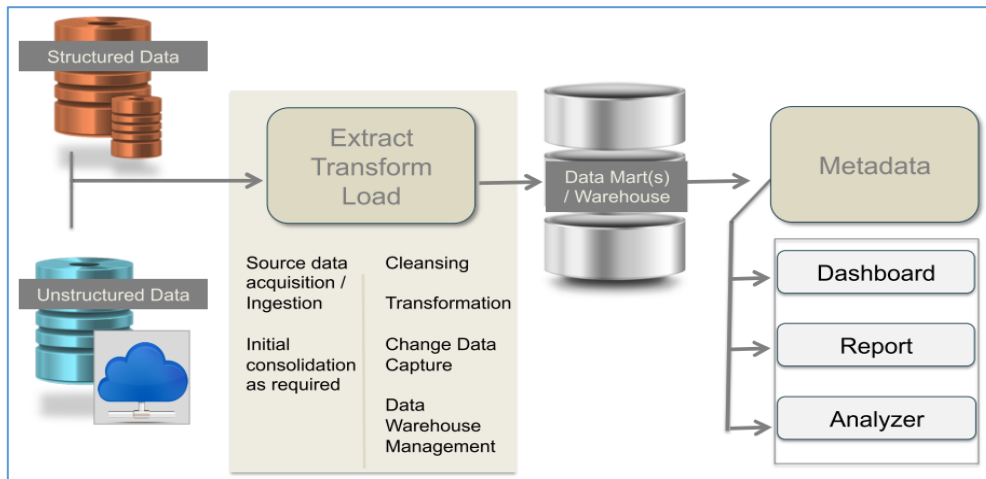


Figure 5-3: Traditional Data Integration and Analytics (sourced from Henry & Nys 2013:7)

Figure 5-3 illustrates that structured and unstructured data coming from various sources is consolidated and cleansed using ETL tools and loaded into a warehouse or data marts in preparations for reporting, analysing and preparing dashboards.

For the purpose of this study, big data is referred to as the massive business data that organisations are accumulating and face the challenge of managing and analysing with their current BI applications. As an example, such an organisation could be a consulting firm, operating with different business units that are collecting data with various data formats such as pictures, text, documents, graphs and so forth, from their various business units and social media platforms. A superb data management approach is crucial for BI and analytics as well as achieving business value as without the utilisation of old and new data organisations will not be able to make informed decisions. Organisations have been able to achieve better and faster decision-making due to available technologies, such as, OLTP, OLAP, in-memory database systems, in-memory analytical and transactional processing, parallel processing, and hybrid technology.

The capability of aggregating big data at granular level with in-memory technology has allowed organisations to analyse their businesses dynamically, in real-time, and to be consistently flexible and responsive to customer needs and changes in the market place (Forrester 2015:6). In addition, in-memory data management supports data mining across multiple different data sources and enables real-time access and integration of very large and complex data.

Big data is identified by five characteristics that do not only refer to volume or size (Rahman et al 2013:544; Hazen et al 2014:72). These characteristics are the following:

- *Volume*: huge amount of multi-structured data;
- *Velocity*: capable of receiving, processing and storing data faster;
- *Variety*: data emanating from various sources with a variety of formats, such as, unstructured, semi structured and structured;
- *Veracity*: inconsistency of data quality and provenance of received data, as data is mostly unstructured or semi-structure; and
- *Value*: data is traceable and processed using best practice processes.

The advancement in technology currently as resulted in efficient management of these five characteristics of big data, as discussed below (Henry & Nys 2013:16).

According to Rahman et al (2013:544) HBase, Teradata and in-memory database systems have emerged as column-oriented and a big data capable database system scaling to billions of rows. In addition, Teradata is ideal for managing big data as it comes with a parallel processing capabilities. Open source Hadoop provides a HDFS that handles large volumes of data in which MapReduce is used for parallel processing on server clusters. This ensures that big data is distributed and only a subset of data is stored on each node. "Each node holds its own independent unit of CPU, RAM and storage resulting in less data movement as the processing occurs where data resides. MapReduce allows lower cost processing of massive data. Hive provides the capability of a data warehouse with SQL-like access. On the data mining front, Mahout provides a library of machine learning and data mining algorithms. Sqoop is used to import data from relational databases. Zookeeper is used as a configuration management and coordination" tool. A combination use of such big data technologies has proven to reduce the processing time from days to hours to minutes down to seconds. These technologies have proved to be capable of addressing big data management

challenges with regards to the capturing and managing of real-time data (Russom 2013a:6).

5.2.4 Hybrid data warehouse

Hybrid data warehouse is a modern information management platform that provides all-in-one access to assorted data both from the perspective of its source, access mode and representation pattern, such as, public, cloud, unstructured, structured, semi-structured, and multi-structured (Hortonworks & Denodo 2015:9). Hybrid data warehouses use lower cost storage, such as, Apache Hadoop 2, an open source distributed processing technology with capabilities of processing large amounts of unstructured data and acting as a staging area for incoming information. These data warehouses have optimised data processing workloads, that is, data transformation and integration as well as access to all the data of an organisation through multiple use and workload data processing from batch to real-time. In addition, highly adjustable workload data processing calls for an interesting cloud computing solution, such as, PaaS basically Hadoop-as-a-service and IaaS (Lopez & Antoni 2014:14).

Hybrid data warehouses allow OLTP systems to store tuples to be organised in rows and warehoused in “blocks on disk and cached in main memory in the database server” (Plattner 2009:2). On the other hand, OLAP systems organise data in star schemas with an optimised column data structure arrangement. Column data storage is ideal for the modern CPUs that come with multi-core architecture which provides enormous computing power. SAP HANA uses “hybrid in-memory databases that combine row-based, column-based and object-based database technology” aimed at providing optimised data analytics (Oracle 2014:2).

In order to have an optimised, powerful and cost-effective data warehousing solution, vendors such as Netezza (now an IBM company), Greenplum (now part of EMC), DATAlegro (now part of Microsoft), and Teradata have integrated hardware, operating systems, databases and analytic applications (Watson 2011:7). In addition to this, these appliances are being used to off-load specific analytic applications or host an entire data warehouse, and some employ a columnar database to increase query processing speed. Hybrid data warehouses are employing analytics inside the

warehouse database. For example, SAS/ACCESS interfaces with Teradata and Oracle to have an in-database data mining and predictive analytics feature that eliminates the need for having the ETL tools on a separate server to perform the data cleansing and analytics (Webb 2008:8). The SAS/ACCESS Interface to Teradata enables access to more sources on more platforms and direct, easy, secure access to native interfaces. Figure 5-5 below illustrates a conventional (traditional) SAS analytical modelling compared to an embedded analytical modelling in respect of scoring functions and data transformations.

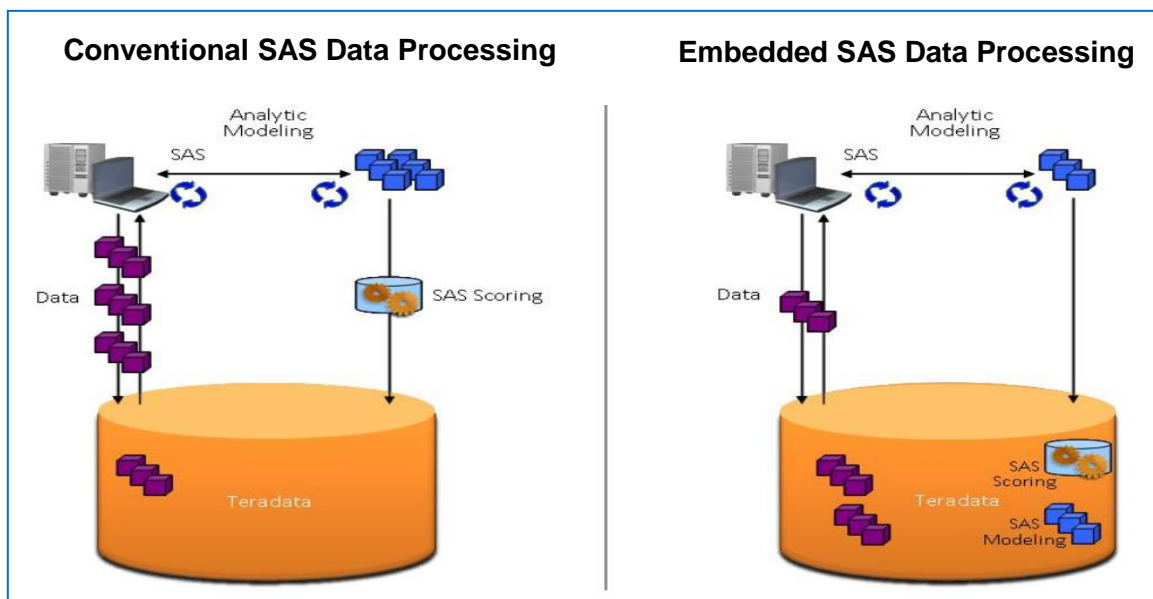


Figure 5-4: Conventional Data Processing vs Embedded Processing (sourced from Webb 2008:9)

Figure 5-4, illustrates the transfer of work and related reduction in data movement between the conventional data processing environment, which relies on the SAS/ACCESS Interface to Teradata, and the environment that is created by SAS In-Database processing (Webb 2008:9). A conventional data processing model uses a Teradata function to deploy the SAS scoring function from outside the DBMS. Whereas, in the embedded data processing model, advanced SAS analytic procedures push enhanced scoring SQL functions directly to the DBMS. The benefits of deploying this hybrid data warehousing technique are that the unique and powerful SAS analytic algorithms are executed inside a Teradata DBMS, and enhanced performance, reduced data movement, exploitation of Teradata's parallel processing, and optimised processes of data discovery and analytic model execution are ensured.

In summary, all these advancements in technology can be deployed in the hybrid BI deployment, that is, on-premise which may include isolated, overlapping and shared and, or centralised deployments (MicroStrategy 2012:8; Schlegel et al 2014:2) and cloud-based which comprises PaaS, IaaS and SaaS service models that can be deployed in a private or public cloud (Watson 2012b:5; IBM 2015:5). The utilisation of on-premise BI and analytics technologies has proved to support faster decision-making as well as reduced latency in data warehouse performance and reduced computation resources consumption, among other benefits (Rahman et al 2013:543). In addition, the utilisation of cloud-based BI and analytics has proved to be quicker and cheaper to implement and operate (Liyang et al 2011:1025).

However, one solution cannot be used in all circumstances, and there are various possible alternatives for in-memory processing, big data management and hybrid data warehousing. The best solution that meets the business requirements may be determined by factors, such as, “business objectives, workloads, data sources and volumes of data” (IBM 2012:2). The combination of the technologies mentioned above finally enables an iterative link between the instant analysis of data, the prediction of business trends, and the execution of business decisions without delays (Plattener & Zeier 2012:27).

5.3 Overview of hybrid business intelligence

In the previous chapter, a hybrid cloud was discussed in brief and it was found that it is a more complex type of deployment than the other private and public deployment models as it involves a structure of two or more clouds (Jansen & Grance 2011:3). It was also found that while primarily using a private cloud, organisations may combine the private and public cloud delivery models in order to use the extra computing capacity of a public cloud (Garrison et al 2012:63). As a result of this deployment, organisations are able to realise value from reduced IT costs by having an increased computing capacity and services. Although, there are various benefits associated with a hybrid cloud deployment, organisations still have a number of problems as previously discussed. Nonetheless, this deployment model is less relevant to the context of this study. The good news for organisations is that there are many different options (vendor tool types, features, techniques and methodologies) for managing

high volumes of business data as well as for addressing the problems identified in on-premise and cloud-based BI Deployments (Russom 2013:27).

For the purpose of this study, a hybrid BI deployment is a combination of an on-premise BI deployment and a cloud-based BI deployment. Whereas an on-premise BI deployment focuses on hardware, software and application systems deployed inside the organisation's premises, and cloud-based BI deployment focuses on the standard cloud service models, namely, SaaS, PaaS and IaaS. A blend of on-premise and cloud-based BI deployment options is believed to be an ideal way of BI deployment as the benefits from such a deployment outweigh the individual deployment (Srinivasan 2011:344). A Hybrid BI deployment deploys BI solutions in such a way that some applications are housed on-premise and some in the cloud through a third party (Menon & Rehani, 2012:8). Most organisations choose this option of using a hybrid BI deployment to ensure that data ownership and security remains with the organisation and to be able to use faster cloud-based BI services.

5.3.1 Purpose of a hybrid business intelligence deployment

The purpose of using a hybrid BI deployment is to ensure that organisations are able to utilise the state-of-the-art (robust and faster) computing resources found in both on-premise and cloud-based BI deployments. Organisations are able to benefit from the best of both deployments as it allows for BI solutions that manage sensitive business data to be kept on-premise in order to retain data ownership and security, as well as outsource the services from cloud BI vendors BI applications for reporting (Russom 2013:12; Menon & Rehani 2012:8). This approach does not only provide peace-of mind to the business users in an organisation where data management is concerned, but it offers a perfect opportunity for them to effectively and efficiently manage their data and BI computing resources.

5.3.2 Value of a hybrid business intelligence

The use of integrated data management systems and BI solutions have made some organisations realise the full business value of utilising BI applications and have achieved technical efficiency (Russom, 2013:33). One of the key elements to

achieving a level that is operational efficient, logical and supports the decision-making process throughout the organisation, is largely based on the models and approaches used when evaluating the BI capabilities and competencies of the needed BI application deployment (Ghazanfari, Jafari, & Rouhani 2011:1581). It is the broad intention of this study to discover and produce an end-to-end BI solution that can help organisations reach this operational level. A detailed discussion on the advantages of a hybrid BI deployment is provided later in this chapter.

5.3.3 Potential users of a hybrid business intelligence solution

The use of a hybrid BI deployment is largely dependent on the size and needs of the organisation. It is suitable for organisations that need to keep certain applications and business data on-premise due to the insecurity of the cloud environment, but at the same time need to use cloud-based BI services to achieve other business objectives (Cruz et al 2015:162). Hybrid BI solutions are becoming common deployment modes due to the fact that organisations are able to utilise SaaS, PaaS, IaaS computing resources that come with cloud-based BI solutions optimally (Yang 2012:4, Garrison et al 2012:63). Some solutions offer cheaper data integration and BI tools in one solution. For example, Pentaho has a Visual MapReduce engine that removes the need for writing sophisticated codes which is usually required to support multiple layers of analytical appliances (Russom 2013:33). For organisations that cannot afford expensive BI solutions, a hybrid BI solution offers the opportunity to optimally utilise a combination of proprietary traditional on-premise BI applications, and cloud-based open source BI solutions, such as, Pentaho and Birst (Dimensional Research 2013:11).

The provided hybrid BI solution suggests the utilisation of various advances in technologies, such as, in-memory technology, big data management, and hybrid data warehouse to address any database performance and analytical challenges. It also promotes the use of technology aimed at turning the ever growing volumes of data into intelligent information that may be used for operational and strategic decision-making.

To sum up, some of the challenges emanating from on-premise and cloud-based BI deployments as discussed in the previous chapters which the proposed hybrid BI solution will try and address are presented in table 5-1 below.

Table 5-1: On-premise and Cloud-based Summary Challenges

On-premise	Cloud-based
<ul style="list-style-type: none"> • Managing and maintaining large number of BI applications • Costly IT infrastructure investment • Inconsistent reporting leading to manual interventions • Managing big data challenges • High maintenance and transaction costs • Lack of real-time access to critical business information • Lack of application scalability • Lack of geographical scalability 	<ul style="list-style-type: none"> • Unclear Service Level Agreements • Vendor lock-in • Data security concerns • Higher inherent risk in cloud integration and configuration • Security concerns • Data stores performance issues • Lack of standardised pricing models • Data ownership concerns • Data volume concerns • Reliability concerns • Multi-tenant attacks on communication infrastructure

The following section demonstrates a potential hybrid BI model that can incorporate the above-mentioned various technologies.

5.4 Hybrid business intelligence model

The hybrid BI deployment is largely dependent on all factors mentioned in the on-premise and cloud-based deployments. Figure 5-5 below provide a model for the suggested hybrid BI solution that enables blended BI and analytics solutions to be accessed through on-premise and cloud-based deployments.

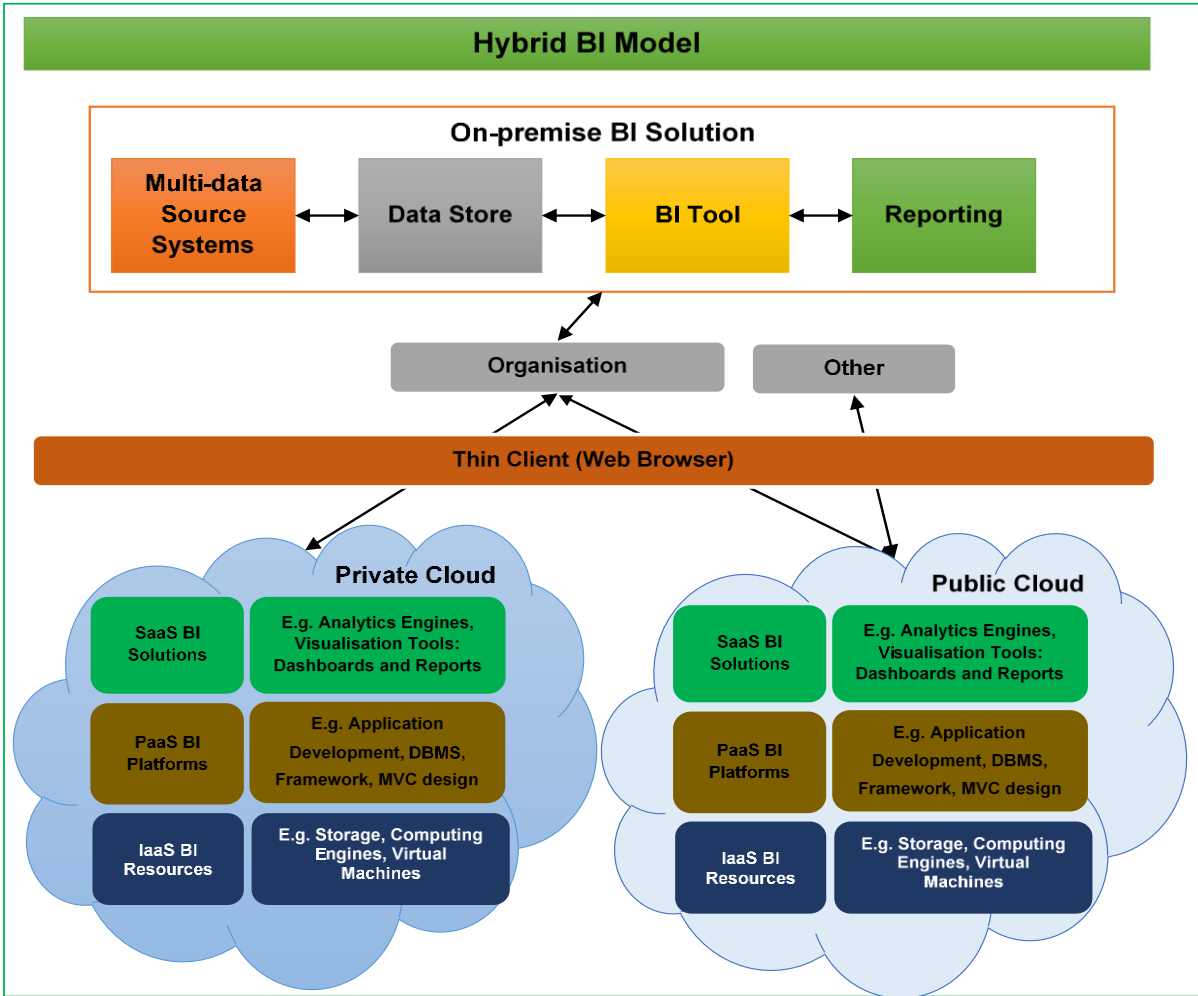


Figure 5-5: Hybrid Business Intelligence Model

The following sections discuss the components of the hybrid BI model in detail.

5.4.1 On-premise BI solution

The on-premise BI solution consists of multiple source systems, data stores, BI tools and reporting types. This deployment is normally achieved through the use of dedicated and specialised BI departments that act as an official department with expertise and responsibilities for ensuring that BI supports business decision-making processes (Dyché et al 2011:44). Organisations in the model have the opportunity to utilise cloud-based BI solutions in the event that the on-premise BI deployment experiences limitations in regard to expansion. The hybrid BI solution offers an integrated solution that can manage data maintained in different enterprise

applications and has the ability to scale up or down to meet future organisational needs (Birst 2015:2). This is normally achieved through the use of a thin client interface.

5.4.2 Thin client interface

The thin client is one of the key components of a cloud-based BI deployment. In the hybrid BI solution, it serves the same purpose of allowing the organisation(s) in the model to access cloud-based BI application services (Garrison et al 2012:63). The use of a thin client interface alongside the firewalls will ensure that security issues that might hinder performing the required BI tasks are properly managed (MicroStrategy 2012:276). The thin client interface will allow the organisation in the model to access cloud-based BI solutions that are not available in the deployed on-premise BI solution.

In the hybrid BI model with thin client interface, organisations are able to utilise the various SaaS, PaaS, and IaaS services from different or the same vendors and each organisation will be configured according to the needs of the business. The cloud-based BI solutions are designed in such a way that an unlimited number of business users from various organisations are able to access BI solutions through the thin client at any given time from any geographical region as long as there is Internet connection.

5.4.3 Private and public business intelligence solutions

Based on the findings from the previous chapters that on-premise and cloud-based BI solutions allow the implementation of various services in the form of isolated, overlapping, centralised and, or shared SaaS, PaaS or IaaS BI solutions (MicroStrategy 2012:8; Schlegel et al 2014:2; IBM 2015:5). In order to avoid some identified challenges associated with the above-mentioned services, the hybrid BI solution ensures that different BI solutions are deployed based on the nature of information they will be handling. On the other hand, less sensitive data can be stored and managed using cloud-based solutions and the necessary BI and analytical solutions can be deployed in order to get the much-needed information that will support better and informed decision-making.

Our hybrid BI solution provides a blended environment where BI applications can be developed or customised either on-premise or in the cloud. For example, cloud service providers provide a PaaS environment for developers to be able to cheaply develop, customise and deploy SaaS BI applications such as Power BI, Pentaho Business Analytics, and Birst (TEC 2015:23). The same can be done in an on-premise deployment, but, the cost is always higher as it will involve various circumstances, such as, license costs, consultants and developers (as illustrated in section 5.7) to come up with an ideal BI solution. Thus, in most cases, organisations in the hybrid model will prefer those options that ensure ease of use, quick implementation, and are readily accessible from anywhere, that is, cloud-based BI solutions that will produce quality information to support strategic, tactical and operational decision-making (Davis et al 2009:3).

This study's on-premise and cloud-based BI models demonstrate that an organisation can make use of various source or storage systems for managing business data residing in their respective departments. The advantage of using various source systems in the hybrid BI model, is that organisations are able to allocate specific source systems to specific BI applications (Zhang & Tu 2009:182). Since most OLTP-based source systems have the ability to process large amounts of data, organisations can then decide either to utilise cloud-based storage and thereafter make use of cheaper cloud-based BI and analytical applications or store data on-premise and utilise applicable costly tools (Ong et al 2011:4; Williams & Williams 2007:201).

The hybrid BI solution overcomes the challenging tasks faced by multiple source systems deployment related to multiple data sources. Managing multiple data sources is often costly to an organisation if it only relies on on-premise or cloud-only sources. For example, if only on-premise is used, there will not be enough capacity, or even tools to handle the ETL process. Thus, an organisation is obliged to invest more to have specialised in-house tools to handle multiple data sources. The same applies in the case of cloud-based solution may be that the solution in the cloud is a proprietary solution that has been adopted and it is possible that the tool in the cloud could not handle data from disparate sources. The hybrid model overcomes this because organisations may have access to on-demand services, that is, specific specialised SaaS BI tools for dealing with the ETL process for disparate tools. These tools allow

the positioning of data in an aggregated fashion within an appropriate IaaS data repository for further processing either on-premise or in the cloud. Organisations have full control when they utilise IaaS storage resources that support data and BI application integration so that information emanating from various departments can be easily incorporated for the purpose of BI (Han 2013:88).

5.5 Architecture for a hybrid business intelligence solution

The proposed hybrid BI architecture in this study demonstrates how to utilise a combination of on-premise and cloud-based BI solutions, including service models, such as, SaaS, PaaS and IaaS. The hybrid BI architecture complements the discussion on the hybrid BI model earlier in order to elaborate more on how a hybrid BI solution can best deliver and meet an organisation's BI needs. The architecture in figure 5-6 below demonstrates the ideal set-up for delivering enterprise wide on-premise and cloud-based BI deployment.

In view of the discussion on the proposed hybrid BI model, and in order to achieve functional and technical understanding, an architecture has been suggested consisting of five main layers, namely, the presentation layer, the application layer, the network layer, the service layer, and the data layer. In other words, the architecture will illustrate or show front-end BI and technology tools, on-premise/cloud services, back-end data management and processing applications, on-premise and cloud infrastructure services.

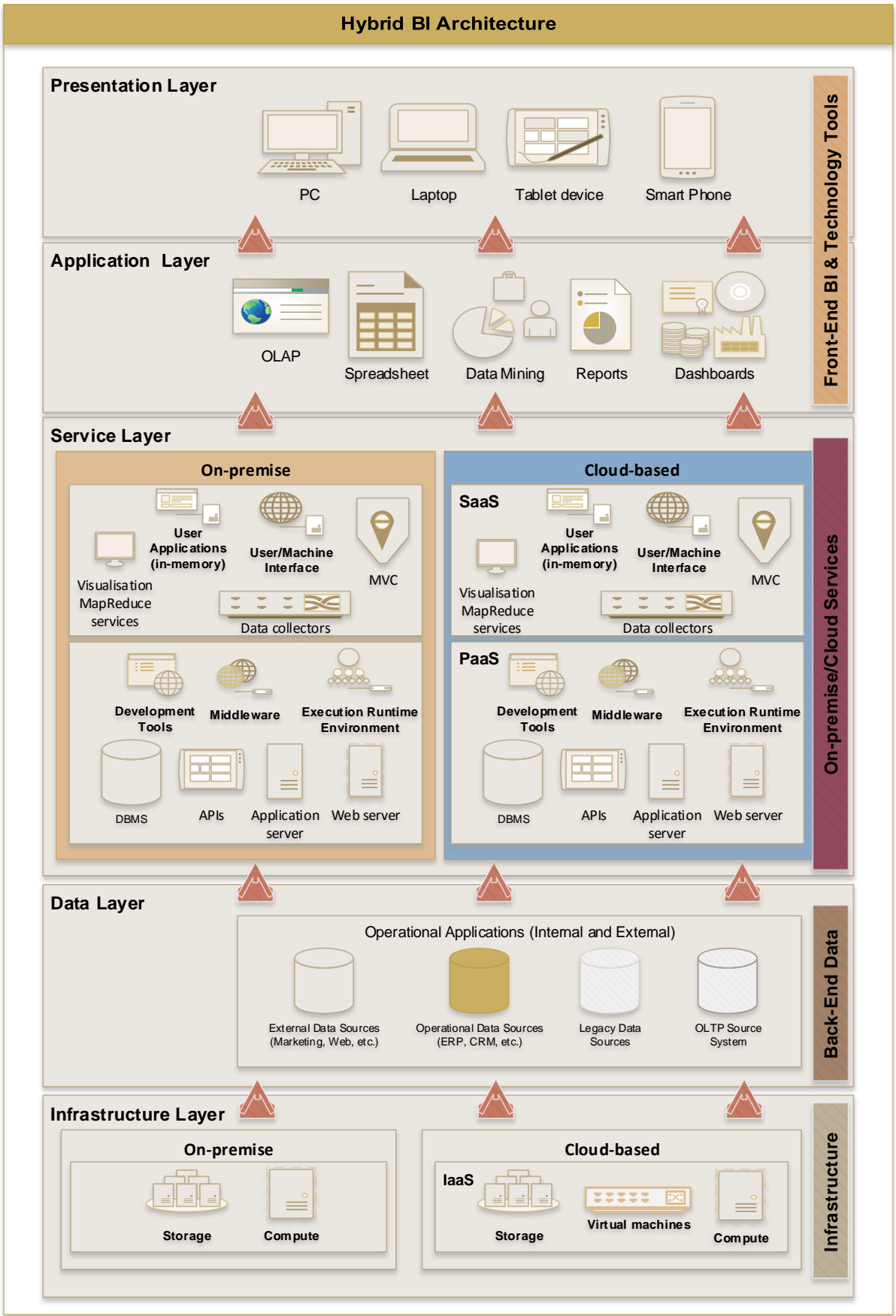


Figure 5-6: Hybrid BI Architecture

The following subsections discuss each layer in more detail.

5.5.1 Presentation layer

As in the cloud-based BI deployment, the presentation layer is the front-end for consuming BI applications and services either residing on-premise or in the cloud. The presentation layer allows organisations in the hybrid BI model to use their technological devices (computers, laptops, tablets, smartphones, etc.) that have the capability of running web applications through the thin clients, such as, Internet Explorer, Chrome, Firefox, Safari, etc. in order to access a vendor's online portal (Ali et al 2015:360; Garrison et al 2012:63). The vendor's on-premise or online portal facilitates the entry for organisations to access BI services through the Intranet or Internet (Mavodza 2013:136). Once organisations have access to the BI services, they are able to run specific BI applications, such as, data mining, dashboards, and reports that can yield information to be used in making strategic, tactical, and operational decisions.

The pricing model utilised will determine how many business users will access the online vendor's portal (Birst 2014:3). Organisations are able to structure who has access to what type of BI applications to be accessed depending also on the number of users they have. On the other hand, on-premise BI applications are accessible depending on the licensing model which organisations are using and based on certain access levels granted to business user on LAN or WAN networks. Business users are not limited to what they can do with their technological devices once their devices are configured to access specific BI applications. This enables easier access to information required for making informed decisions and information sharing at the same time.

5.5.2 Application layer

The presentation layer supports the use of both on-premise and cloud-based BI applications deployment. The observation from the previous discussions shows that BI tools that are commonly used in the presentation layer can range from Pentaho Business Analytics, SAP BusinessObjects Dashboards and, or Mobile app, Cognos, Oracle BI Publisher, etc. These tools allow power users to analyse business data and

publish reports to any technological devices according to information requirements by business users (Russom 2013:34). Furthermore, the self-service BI features that come with most BI solutions, allow business users with non-IT backgrounds to query or generate BI reports that address particular business needs (Weber 2013:19).

The applications layer enables organisations to create personalised, interactive, high-impact and dynamic dashboards that can be securely published instantly to any technological devices which ensures access to the right information at anytime and anywhere (TEC 2014:46). Furthermore, since business users are easily able to interact visually with their BI tools, either on-premise or in the cloud, organisational wide BI adoptions become easy and organisations are able to see increases in return on their investments in their enterprise data. With the use of devices, such as, smartphones and tablets, interactive dashboards can be securely published precisely to meet business user specifications and help them make better decisions faster.

5.5.3 Service layer

The service layer in the hybrid BI model hosts the on-premise and cloud service models. The service layer provides various analytical engines that ensures BI and analytics services are delivered to the business users when, and as needed (Watson 2012:5). When organisations in the hybrid BI model choose to use the on-premise BI deployment, all their BI services will be securely hosted locally in their local servers. However, if organisations choose to use cloud-based BI services, the cloud services vendor will host and manage all needed BI services through the provision of SaaS and PaaS services. Cloud-based BI service vendors use state of the art technology to ensure that their clients get the best service whenever they access the service (Rahman et al 2013:542). This includes the use of technology discussed earlier, such as, in-memory data management, in-memory analytics and transactional processing, big data management and hybrid data warehouse.

When organisations in the hybrid BI model, choose to use cloud-based BI services, they benefit from the advantages that come with SaaS and PaaS services. As discussed in chapter 4, a variety of combinations of cloud service models can be selected depending on what the organisations are aiming to achieve. There are

several opportunities associated with cloud-based BI, such as, the ability for business users to customise or build their SaaS BI applications (for example visualisation MapReduce services, MVC, Data collectors) through the PaaS platform (Bibi et al 2012:86). PaaS provides all the necessary application development resources, such as, Application Servers, APIs, and Execution Runtime Environment needed to customise or build SaaS BI applications and there is no need for business users to worry about the underlying cloud infrastructure (Ali et al 2015:360). This ensures that organisations are adaptive to these constantly changing computing and business environments.

All these on-premise or cloud-based analytical capabilities, integrated development environments and, or platforms, storage, computing or networking resources are accessible through the use of technological devices demonstrated in the presentation layer. The applications that offer OLAP, Spreadsheet, data mining, reports and dashboards in the application layer are available through desktop applications, web portal, mobile clients, and external application programming interfaces. A robust Internet network infrastructure, reliable network devices and secure connections, applicable to both on-premise and mobile computing devices, are required to fully realise the value of using a hybrid BI deployment.

5.5.4 Data layer

In the hybrid BI model, the data layer hosts multiple data source systems which provide data stored in a data warehouse that can either reside on-premise or in the cloud (Dyché et al 2011b:15; Chaudhuri et al 2011:90). Multiple data sources, such as, ERP, SCM, Web, Legacy systems can be housed on-premise or in the cloud depending on the nature of the information which they need to handle (Ong et al 2011:4; Zhang & Tu 2009:182). Organisations that handle very sensitive data can use the on-premise BI deployment so that information and BI services are managed locally. This will ensure that those organisations that still do not trust cloud-based BI solutions are able to use the necessary BI applications on-premise (Armbrust et al 2010:55).

Since data originates from multiple data sources, the use of hybrid warehouse or ETL tools (e.g. InfoSphere, DataFlux, or DataCleaner) will extract data from multiple

internal or external source operational systems, then transform it by cleansing and reformatting it to fit operational needs and eventually loading the data into a targeted data warehouse (Manyika et al 2011:32). Upon completion of the ETL processes, the data layer will be ready and enable users to use BI tools and, or analytical engines to perform extensive analysis on large volumes of data in real- or near real-time through three (3) commonly used components, that is, data warehouse, operational data store and data mart (Boateng et al 2012:223).

5.5.5 Infrastructure layer

The infrastructure layer offers BI resources, either as cloud-based (IaaS) services or on-premise resources (Mvelase et al 2012:215). Organisations can benefit from the utilisation of IaaS services when it comes to data storage for the multiple source systems and managing of big data. Since business data will continue to grow and organisations face the issue of storage, IaaS services offer unlimited storage services at cheaper pricing models compared to on-premise deployment which often require purchasing additional infrastructure resources, such as, storage hard drives whenever the need arises (Cruz et al 2015:167). The use of IaaS services eliminates the strain often developed in on-premise BI deployment when isolated BI applications and databases have to be managed and maintained (MicroStrategy 2012:7). All in all, Birst (2014:2) asserts that cloud-based BI service providers offer several options that make it affordable to pay per month or annually for storage, BI and analytics solutions, rather than organisations buying, integrating and maintaining costly BI solutions on-premise.

The next section discusses the advantages of deploying the hybrid BI solution.

5.6 Advantages of hybrid business intelligence

In view of the discussion on the hybrid BI deployment, it is clear that this type of deployment is associated with considering the use of various technological advances that can enable successful implementation. The advantage of this option is that business users can experience the best of both BI deployments, that is, on-premise and cloud-based deployments as discussed in previous sections. Table 5-2 highlights some key points from previous chapters.

Table 5-2: Hybrid BI Related Advantages

Common themes	On-premise	Cloud-based	Hybrid
Investments	High	Low	Low
Costs	High	Low	Low
Technology	Old / Legacy	New	New
Storage	Limited	Unlimited	Unlimited
Maintenance	High	Low	Low
Scalability	Limited	Highly scalable	Highly scalable
Upgrades	Slow / costly	Fast / no cost	Fast
Disaster recovery	High cost, often complex	Low cost, vendor handles the process	High/Low cost
Deployment	Longer	Immediate	Immediate
IT needs	Complex, costly	Simple at no additional cost	Simple at no additional cost
Agility	No	Yes	Yes

The following advantages are associated with hybrid BI deployment with the consideration of the various technologies discussed:

- **Guaranteed data ownership:** The ownership of business data still remains with the organisation, and they are responsible for managing their multiple data source systems that reside on-premise (Mavodza 2013:139). Doubts in regard to data ownership only exist when it stored in the cloud, but this is often clarified when clear SLAs are understood and implemented.
- **Business model transformation:** Not only does hybrid BI deployment offer cost savings, it drives BI business model transformation by bridging the gap between legacy applications residing in traditional on-premise and cloud-based BI applications as well as mobile computing applications (Zimmerman 2014:87). More and more applications are being built for cloud-based delivery at very low cost and are quicker to implement, such as, the SaaS market place for high value consumable business and BI applications for various business sectors.
- **Balance workload:** A hybrid cloud strategy helps organisations to balance the workload between business and IT by enabling them to identify areas that can be used to manage a certain workload, that is, on-premise or in the cloud (Zimmerman 2014:93). Workload classification (e.g. processing sales data on-

premise and producing and sharing reports in the cloud) is important in the sense that not all critical workloads are suitable for operating in the cloud as it may have financial and operational cost implications for the organisation.

- **Parallel Processing:** Cloud-based BI applications in hybrid BI deployment offer a massively parallel processing mechanism that handles big data when loading and retrieving data from data warehouses which organisations can use to execute data analysis (SAP 2013:4; Sikka 2013:9; Rahman et al 2013). As an example for this processing mechanism, SAP HANA uses parallel multicore processors to manage the SQL processing instructions. This essentially means that no locking of data in a data warehouse is experienced when executing any analytical tasks that need to support decision-making.
- **Performance improvements:** There is dramatic improvement with regard to response time and calculation performance as users can now manipulate, and request data in memory (Acker et al 2011:130). For example, in an in-memory vs traditional data processing environment, the former has instant access to the entire operational data and analytical functions are executed simultaneously. As a result, a greater business value is achieved as business users are able to execute difficult requests or queries within the shortest possible time, and have improved modelling capabilities.
- **Customer value creation:** In-memory analytics provides an innovative way of understanding customers which can be used to improve primary business processes, such as, call centre management or service request handling (Acebo et al 2013:45; Acker et al 2011:130). The scalability that comes with self-service allows ordinary business users to adopt BI applications (IBM Cognos Insight, Power Pivot, Power View, SAP Lumira, etc.) quicker, utilise their expertise throughout the organisation and generate the necessary information needed to make informed decisions. Evelson (2013:3) asserts that with self-service, business users are able to execute simple tasks, such as, consolidating, pairing and removing duplicated data without having to wait for ETL and data cleansing cycles.

- **Real-time analytical processing of big data:** There is an improved analytical speed from a technical perspective, when executing data analysis tasks in memory, it can be up to 1,000 times faster than the ordinary disk-to-processor execution (SAP 2013:1). In-memory technology has made this possible with its ability to instantaneously respond to users' requests for ad hoc analysis by analysing and evaluating business scenarios at the speed of thought (Read 2013:2). TEC (2014:68) asserts that using mobile computing devices, such as, iPads, and smartphones, business users possess the ability to instantly access and analyse data, whether on-premise or remotely.
- **High performance data-processing systems:** Data-processing systems that use in-memory technology have the capability to allow unrestricted query complexity, data volume and variety; to benefit from insertion of real-time data into the analysis and response time; and to eliminate pre-processing of data (Sikka 2013:10). Furthermore, organisations that utilise BI applications that are supported by in-memory technology achieve high performance, which results in very high value when it comes to decision-making capabilities.
- **Lower costs:** Compared to traditional data warehouses, the TCO, this is, direct, indirect and related costs, is considerably lower as the entire data is now stored in memory, reducing the cost of producing data mart (Acker et al 2011:130). It is, therefore, essential for organisations in the hybrid model to consider cloud-based BI solutions that use in-memory technology as they have the capabilities of analysing big data in real-time and are easier and cheaper to set up and maintain.
- **In-memory BI engines:** Chaudhuri et al (2011:93) assert that the advancement in technology is providing the possibility of a new type of OLAP systems which can manipulate the data stored in main memory, and instantly and interactively respond to *ad hoc* queries. The following few trends have emerged: the average time it takes to access data on disk compared to data in memory is increasing; "very large addressable memory sizes (for example, 1TB) are possible; and the

cost of memory has dropped significantly”, making cloud servers or data centre storage with huge main memory capacity affordable. Therefore, as opposed to ordinary OLAP servers, in-memory BI engines depend on a variety of methods for achieving the desired high-performance by using data structures that work well with in-memory.

- **Cloud-based DBMS:** Russom (2013:19) emphasises that most of an organisation’s big data is stored on legacy disk-based DBMS (traditional spinning hard drives) due to their popularity and they will remain as such for years to come. But Solid-State Drives (SSDs) and in-memory technology are gaining popularity too. Few organisations continue to use disk drives as their primary medium for storing big data although they are costly but reliable, while others use speedy SSDs as an important secondary medium although they are expensive. Other organisations have completely opted for a cloud-based DBMS as their primary medium for big data management because they are very much faster as they avoid most I/O operations, and are also faster in respect of queries and data processing, as well as highly available and purposefully redundant in regard to storage (Russom 2013:19).
- **Analytical data management applications:** Organisations in the hybrid BI model can use on-premise or cloud-based BI applications that have the ability to run analytical data management systems that can read from, and write to data store in real-time (Abadi 2009:5). In addition, these BI applications are developed and engineered based on shared-nothing architecture for analytical data management. Furthermore, transactional isolation and recovery, that is, Atomicity²², Consistency²³, Isolation²⁴ and Durability²⁵ (ACID) guarantees are normally not required (Plattener & Zeier 2012:), and, in most cases, particularly when dealing with sensitive data, this can be omitted during analysis.

²² **Atomicity Database** concept that demands that all actions of a transaction are executed or none of them.

²³ **Consistency Database** concept that demands that only correct database states are visible to the user despite the execution of transactions.

²⁴ **Isolation Database** concept demanding that any two concurrently executed transactions have the illusion that they are executed alone. The effect of such an isolated execution must not differ from executing the respective transactions one after the other.

²⁵ **Durability Database** concept that demands that all changes made by a transaction become permanent after this transaction has been committed.

- **Specialised cloud analytic data management convergence:** This technology has contributed various opportunities to the business (Ouf & Nasr 2012:650). Firstly, it immediately creates analytical opportunities that are able to optimise response to current unexpected market changes although it is not necessary to employ infrastructure. Secondly, organisations are relieved from expensive maintenance associated with processing and managing data warehouses and upgrades. Thirdly, organisations have the ability to test new cloud applications without the need to purchase and implement their own data centre infrastructure. Last, but not least, organisations are able to run full BI and data warehousing applications exclusively in the cloud.
- **Cloud big data analytics:** Cloud-based BI applications have the capabilities of developing strategic and competitive advantage by processing heterogeneous and informal data emanating from externally based sources, such as, web sites, social networks and, or media, and so forth (Kwon, Lee & Shin 2014:390). Cloud-based applications supported with in-memory capabilities gives organisations a competitive advantage as they are able to use BI and analytics smartly to make data-driven business decisions and predict the needs of customers and markets (SAP 2013:3). In addition, the capabilities of aggregating big data at granular level in the cloud with in-memory technology has allowed organisations to analyse their businesses dynamically in real-time, and to be consistently flexible and responsive to customer needs and changes in the market place (Forrester 2015:6)

The above-mentioned advantages show that hybrid BI deployment is already solving many of current challenges faced by organisations demonstrated in the models.

5.7 Limitations of hybrid business intelligence

Besides the common limitations and disadvantages experienced in the on-premise and cloud-based BI deployments, some key points from the previous chapters have been highlighted below in table 5-3.

Table 5-3: Hybrid BI Related Disadvantages

Common themes	On-premise	Cloud-based	Hybrid
Security concerns	Low	High	High
Data location	Known	Unknown	Known/Unknown
Compliance issues	Secure	Insecure	Limited security
Data Control	Total control	No control	Limited control
Multi-tenant architecture	Not supported	Fully supported	Fully supported
Reliability	Yes	No, Internet connection dependant	Yes
Inflexible SLA	No	Yes	Yes
Vendor managed	No	Yes	Yes
Scalability	No	Yes	Yes
Manual interventions	Yes	None	Limited

In addition to the above, the following disadvantages are associated with hybrid BI deployment, taking into account the use of various technologies discussed earlier:

- **Technology adoption challenges:** As in the case of any new technology implementation effort, the advances in the use of in-memory or big data management technologies needs to be sold to business users first, accompanied by monitoring and properly managing its usage for the benefit of the users (Acker et al 2011:136). Most organisations have not yet adopted and taken advantage of advances in technology simply because they do not know the full benefits that come with such technologies.
- **High-priced up-front cost:** Organisations in the hybrid BI model that are trying to implement on-premise BI solutions face high-priced up-front costs on infrastructure requirements (Abadi 2009:2). This usually comes as a result of the organisation’s current traditional DBMS, which includes large up-front costs in hardware and software needed to get the BI project up and running. Organisations need to move away from the traditional on-premise way of managing and storing big data on their premises to cloud-based systems, such as, remote data centres in the cloud.
- **Latency and delays in data processing:** The use of the hybrid BI model has been reported to be accompanied by minor performance- related issues with data warehouses, as there are some hidden problems and delays when it

comes to the management of big data in the cloud which resides or is stored on-premise (Suciu et al 2013:118). The ideal deployment when using the proposed hybrid model is to have all data reside in the cloud and only keep a small amount of sensitive data on-premise if necessary. In addition, organisations may choose to run BI tools on-premise which may have less performance issues.

- **Integration and data loading issues:** Organisations in the hybrid model often face the difficult issues that come with trying to integrate ease of use of BI software (that is, MapReduce-like software) which is capable of loading data and creating performance enhancements that are linked to data structures (Abadi 2009:9). It is crucial that data is accessible and is ready to be used in order to achieve ease of use features that come with certain BI tools.
- **Additional cost for deployment infrastructure:** There is always additional investment cost for hardware related to the run-time environment running on-premise that facilitates the development and testing of programs, websites, software or applications for developers (Menon & Rehani, 2012:10). On the other hand, the cost of servers (e.g. production servers) that are used to deploy and host Web BI applications will cost nothing.
- **Slow disk-based DBMS:** Organisations that are using legacy disk-based DBMS can still benefit from the disk hardware that is getting larger and cheaper. However, as far as performance is concerned they are not getting any faster (Russom 2013:19). Although organisations in the hybrid model are utilising larger and cheaper disk-based DBMS, business users will not be satisfied if they are not able to run real-time analysis and reports that can assist in the decision-making process.
- **Limited ACID guarantees:** It is difficult to guarantee ACID properties when data replication is required on geographically distributed data (Abadi 2009:4). This means that BI applications will not be able to set a number of operations on the database so that they do not look as if it is a unique operation to a user,

nor will they ensure consistency of a transaction from start to finish. Furthermore, data access will neither be restricted per transaction, nor will it be free from common types of system failure (Plattener & Zeier 2012:8).

- **Data is stored at an untrustworthy host.** Organisations in the model still face trust issues as there is no guarantee that cloud-based BI vendors will not have unauthorised access to customer data, and the possibility of such an event occurring deter prospective customers (Abadi 2009:4). The cloud still remains vulnerable to a number of potential security risks Thus, it is highly recommended that appropriate precautions be taken.

5.8 Total cost of ownership of cloud-based BI deployment

TCO in the hybrid deployment is dependent on both factors associated with on-premise and cloud-based deployments. As it has been understood, hybrid involves both on-premise and Cloud BI deployments. As in the case of the previous models, in a hybrid deployment there are a number of strategies which an organisation can adopt in order to keep the TCO cost down. The nature and geographical location of the organisation is crucial for cloud-based deployment, it can choose to rely on more cloud adoption and less on-premise, or the other way around. For example, the contribution to TCO will be high when the organisation is using more of the on-premise resources whereas it would be reduced if it is cloud-based because most of the DC, IC or RC individual elements will not be applicable. In comparison to all the categories of users, table 5-4 below summarises the TCO associated with hybrid BI deployment.

Table 5-4: TCO Hybrid Deployment

Key metrics	Naive	On-premise	Cloud	Hybrid
Direct costs (DC)				
• Cost of hardware (e.g. new servers)	X	✓	X	✓ / X
• Cost of middleware (operating system, database servers, Web servers)	X	✓	✓	✓
• Cost of software licenses	X	✓	✓	✓
• Cost of consulting services for implementation	X	✓	X	✓ / X
• Cost of staff hours to implement the project	X	✓	X	✓ / X
• Cost of software quality	X	✓	X	✓ / X

Key metrics	Naive	On-premise	Cloud	Hybrid
• Costs of changes needed in the network	X	✓	X	✓ / X
• Cost of additional storage	X	✓	✓	✓
Indirect costs (IC)				
• Cost of upgrading related software	X	✓	X	✓ / X
• Cost of upgrading related hardware/middleware	X	✓	X	✓ / X
• Cost of customisation/continued application evolution (e.g. new end-user reports /dashboard)	X	✓	✓	✓
• Impact on use of existing servers	X	✓	X	✓ / X
• Impact on existing network traffic	X	✓	X	✓ / X
• Impact on existing storage	X	✓	X	✓ / X
• Increase in support calls to IT	X	✓	X	✓ / X
• Increase in maintenance of hardware/middleware	X	✓	X	✓ / X
• Downtime cost (loss on productivity costs)	X	✓	✓	✓
Related costs (RC)				
• Staffing costs	X	✓	X	✓ / X
• Training end-users	X	✓	✓	✓
• Training BI professionals, e.g. business analysts, data science	X	✓	✓	✓
• Training IT staff	X	✓	✓ / X	✓ / X
• Creating governance rules	X	✓	✓ / X	✓ / X
• Implementing governance rules	X	✓	✓ / X	✓ / X
• Maintaining governance rules	X	✓	✓ / X	✓ / X

As can be observed in the above table, the contribution to TCO is largely dependent on the organisation's characteristics, such as, strategy, industry, size, number of employees, and so forth. For example, there would be a considerable tendency for an international organisation to adopt the hybrid deployment and never adopt the on-premise option only because of the fact that it is international. Organisational data and process-sharing or data centralisation within a data centre for sharing purposes is necessary enabling regional offices across the globe to access such data and processes. Therefore, a hybrid deployment can be used for cross-comparison and is the ideal model for the organisation to keep TCO low.

5.9 Conclusion

The purpose of this chapter was to provide a comprehensive understanding of the proposed hybrid BI deployment solution. It discussed the various emerging

technologies that can support the hybrid BI deployment solutions as well as the associated business value propositions.

The chapter began with the definitions of hybrid BI deployment in order to understand the fundamental principles. The various emerging technologies were discussed, that is, in-memory data management, in-memory analytical and transactional processing, big data management, and hybrid data warehousing that can be used to augment the best hybrid BI implementation which meets the business requirements of organisations.

A theoretical hybrid BI deployment model was discussed which illustrated the utilisation of the key components, that is, thin client interface, multiple business intelligence solutions and multiple data sources and, or storage systems. The model later informed the development of an ideal hybrid BI architecture consisting of five (5) layers, namely, the presentation, application, network, service, and data layers. Finally, a summary of the advantages, challenges, and related TCO for implementing the on-premise and cloud-based BI deployments as well as hybrid BI deployment was presented.

Having achieved the purpose of this chapter, it has been observed that organisations have the opportunity to choose a blended BI implementation depending on what will best meet the organisation's business needs. Organisations can choose an on-premise, cloud-based or hybrid BI implementation depending on its needs and the solution required to achieve it. Proper utilisation of the emerging latest technologies in BI, big data management and analytics has not only helped support faster decision-making, they have decreased data latency when reading or writing to the data warehouse without affecting performance and resource utilisation.

Although the discussion on the proposed theoretical hybrid BI deployment addresses some of the challenges faced in the on-premise and cloud-based deployments, it has been noted that it also has its own unique challenges that still have to be addressed. Therefore, the next chapter seeks to address the challenges and validate the proposed solution by means of a case study. The resulting information will be used to improve the hybrid BI architecture.

CHAPTER 6 DATA COLLECTION

6.1 Introduction

The previous chapters discussed in detail a range of BI deployment solutions, namely, on-premise, and cloud-based resulting in this study's proposed hybrid solution. The theoretical investigations revealed that the above-mentioned BI deployment approaches are currently in use by various organisations depending on their needs and expectations. However, emphasis needs to be given to the hybrid architecture because the proposal is merely theoretical and a practical evaluation and validation ought to be performed in order to assess its potential at an organisational level.

The previous chapters focused mostly on theoretical constructs (literature, design and architectures), whereas this chapter will deal with the results of an empirical investigation which was conducted to determine at organisational level the kind of BI solutions mostly in use and also to ascertain whether the hybrid BI solution is of value to organisations in the South African context.

The purpose of this chapter is to discuss the data collection methods adopted for the survey conducted on specific industries in South Africa in order to achieve the above-mentioned objectives. The survey was divided into two parts. Part one focused on determining the state of BI utilisation in an organisation, and part two focused on validating the proposed Hybrid BI architecture by experts for plausible improvement. This chapter and the chapters to follow ultimately contribute to providing an answer to the fourth research question of this study on the validation of our proposed model and its improvement.

This chapter is structured as follows: sections 6.2 and 6.3 below briefly discuss the case study and data gathering respectively, followed by discussions on the research instrument and questionnaire in sections 6.4 and 6.5. Software tools utilised for data gathering and preparation are briefly discussed in section 6.6. In section 6.7, the research sample design is discussed, followed by the actual data preparation and cleaning in section 6.8. The study limitations, ethical considerations as well as the pilot study are covered in section 6.9, 6.10, and 6.11 respectively. The conclusion to the chapter is provided in section 6.12.

6.2 Case study

When conducting research, a case study method allows close consideration of the data within a real-life phenomenon with fewer subjects of the study, such as individuals (Rugg & Petre 2007:16; Zainal 2007:2). In this research, the case study method ensured that the focus was on respondents (i.e. BI experts and BI users) exposed to BI solutions in various industries. The following are some of the characteristics that were considered when choosing the case study method (Kothari 2004:114; Sukamolson 2007:8):

- A case study can be focused on a single social unit or more such units;
- A selected unit can be studied intensively for some time in order to obtain enough information for drawing an informed conclusion;
- A case study can be used to directly study the behaviour patterns of the relevant unit; and
- A case study can use one or more data collecting methods, such as, questionnaires, documents, etc. that would warrant the desired results.

Considering the above, this research was a two-pronged approach. Firstly, an extensive literature review was conducted on BI in general, and then specific BI deployments and their associated TCO were discussed. Secondly, in order to assess BI deployments and validate the resultant proposed BI architecture from an extensive literature review, a case study was conducted. The first case study focused on BI users in which the current state of BI utilisation in their organisations was assessed; and the other focused on BI experts which looked at validating the components of the proposed architecture so that it can be improved.

6.3 Data gathering

The primary data for this research was collected by administering questionnaires. The questionnaire had specific investigative questions aimed at getting sufficient information to respond to the identified research questions (Cooper & Schindler 2006:358). The questionnaire was designed with structured (closed-ended) questions

with minimum comments in the respondents' own words which ensured focus on the unit of study. It was time efficient, and cheaper to administer (Kothari 2004:118). Furthermore, closed-ended questions allowed responses linked to specified answers or stated alternatives to be coded into data ready for analysis. The provision of specific options helped the respondents to understand the meaning of the questions.

Although closed-ended questions were simple to administer and fairly cheap to analyse, a wide range of data was not obtainable. In addition to closed-ended questions, the main questions or statements related to BI were all designed with a five-point Likert-type scale which ensured that the questionnaire remained consistent and a choice was required for each statement provided (Likert-type scale is discussed in detail later). The questionnaires were pre-tested to ensure that all the questions were well formulated. The final questionnaire was first administered through a pilot study as discussed later in this chapter. The questionnaire ensured a quicker way of administering the survey, and maintained low cost and anonymity (Sukamolson 2007:15).

6.3.1 Survey research

The survey research considered a quantitative approach in order to achieve accurate and consistent measurements (Cooper & Schindler 2006:198-199). The quantitative approach is discussed in detail in the next chapter. For the purpose of this research, and in order to answer the research questions and achieve the research objectives, it was appropriate to quantify the opinions, attitudes and behaviour of BI experts and BI users (Sukamolson 2007:9). The use of the survey research method encouraged the logical collection of data from participants so that a scientific sampling and questionnaire design could be used to measure characteristics of the population with statistical precision (Sukamolson 2007:4,12).

6.3.2 Questionnaire administration

This evidence-based research was based on measurable data collected through clearly defined survey questionnaires that were easy to handle and quick to answer.

All research characteristics or questions were specifically designed before data collection to ensure collected data was relatively inexpensive to analyse, and so that the results could be used to generalise perceptions and, or predict future results (Alasuutari 2010:149, Sibanda 2009:3, Kothari 2004:103). The questionnaires were administered through an email with a link to an online survey, or an email with an attached soft copy (Word document) or hand-delivered (printed hard copy). The online survey questionnaire was developed using Google Forms and ensured administration at no cost and could be answered by any potential respondents within the sampling frame residing anywhere in South African. The hard copies were used where necessary and handed over to potential respondents who did not have internet access after working hours and wished to participate.

Each questionnaire that was sent out was accompanied by a consent letter requesting the respondent's permission to participate in the study. The questionnaires did not carry any identification information which ensured that all responses were completely anonymous and respondents still fell within the sampling frame.

6.4 Research instrument

This research study used questionnaires as the research instruments or tools for collecting data. The questionnaires were designed for two types of business users, namely:

- BI experts, that is, those who perform the much-needed implementation and, or advanced analysis, such as, Business Intelligence Consultant, Business Architect, Project Manager, Product Manager, Data Analytics expert, Business Analyst, and System Analyst; and
- BI users, that is, those that use the resultant quality information, such as, Executives, Business Analysts, and Operational Users.

Each questionnaire was split into sections, namely, a section that dealt with demographic information, and a section(s) that dealt with BI specific-related questions. The next section discusses the structures of the questionnaires.

6.4.1 BI users questionnaire structure

The questionnaire for BI users consisted of two sections which was intended to determine the state of BI utilisation in an organisation by business users. Section 1, covered demographic information, such as, the number of years of experience, the size of the organisation or industry, and the number of business users. Section 2 focused on assessing the current BI solutions and deployments in the organisations. Consult Appendix 3 for the detailed questionnaire.

6.4.2 BI Expert Questionnaire Structure

The Expert BI questionnaire was divided into 6 sections which aimed at validating the proposed hybrid BI architecture by experts in order to improve it. Section 1 covered demographic information, such as, years of experience, industry, deployment and geographical footprint of the organisation. Section 2 covered the respondent's familiarity with the current BI deployments and pricing models utilised. Section 3 to 6 focused on validating the proposed BI architecture consisting of front-end, service, back-end and infrastructure layers. Consult Appendix 2 for a copy of the BI expert questionnaire.

The questions in both the above-mentioned questionnaires were closed-ended questions, with a list of specific options from which to choose. Some questions offered an opportunity for the participant to provide additional answers or comment. The advantages of using a questionnaire structured in this manner are that the cost of administration is low, the resultant information is easy to quantify and maintain, and there is a fair chance of obtaining large volumes of data (Cooper & Schindler 2006:254). However, closed-ended questions have their own disadvantages. For example, the questions do not allow an open interaction, and in this way comprehensive responses are limited.

6.4.3 Choice of measuring scales

The BI specific-related questions were structured with a five-point Likert-type scale. A five-point Likert-type scale ensures that the respondents express their answers by agreeing or disagreeing to the statements that have been presented (Cooper & Schindler 2006:339). All BI specific-related questions required that the respondents selected one answer from the options presented on the scale, namely, *Strongly agree*, *Agree*, *Don't Know*, *Disagree*, and *Strongly Disagree*. In addition to this, optional questions were asked that allowed respondents to select answers from the list, or enter answers in their own words (Consult Appendices 2 and 3).

6.4.4 Validity and Reliability

Validity and reliability concerns were addressed to ensure that the measuring instruments were fit for their purpose, that is, that they measure precisely what is supposed to be measured with accuracy and precision (Cooper & Schindler 2006:318).

6.4.4.1 Validity

There are two forms of validity that are relevant, namely, internal and external validity. Internal validity is used to recognise face validity, content validity, criterion validity and concurrent validity (Cooper & Schindler 2006:318). In order to achieve face validity and content validity the content of the research instrument was assessed by the BI experts who agreed respectively that the questions were valid, and this ensured that the instrument was relevant to the respondents. In order to envisage the results and concurrency, criterion validity was established by analysing the results of the questionnaires completed during the pilot study. To assess the awareness of BI utilisation and the architectural components, concurrent validity was attained through analysing the responses obtained from the pilot study. Therefore, internal validity with regards to face, content, criterion-related and concurrent validity was addressed so that the research instrument measured what it was supposed to measure.

Furthermore, internal validity ensured that the research design carefully followed the principle of cause and effect (Khalid, Hilman & Kumar 2012:23).

Whereas external validity is used to indicate the degree to which an instrument measures what it is supposed to measure as already mentioned above. External validity is concerned with whether the research findings can be generalised to all similar cases across various characteristics of people, industries and organisations (Cooper & Schindler 2006:318; Khalid et al 2012:23). External validity was ensured by drawing a general conclusion on the findings that informed the enhancement of the proposed hybrid architecture.

6.4.4.2 Reliability

Reliability is used to define the degree to which a measure is free from random error and provides accurate and consistent results (Cooper & Schindler 2006:321; Khalid et al. 2012:22; Kothari 2004:75). Furthermore, in order to maintain consistency and accuracy of the research instruments, reliability is demonstrated in terms of “stability, equivalence and internal consistency”. Firstly, stability which is the degree to which the results achieved with the instrument can be repeated was achieved during the pilot study by means of testing the questionnaire with the same respondents at different times. Secondly, equivalence which shows the extent to which different methods of the instrument are used to achieve the same results was established by comparing the observations of the respondents who answered the questionnaire. Last but not least, internal consistency which deals with the extent to which items in the research instrument are standardised and reveal the same fundamental construct was achieved by using a five points Likert-type scale in the main questions as it provides an acceptable reliability measurement.

6.5 Objectives and categories of questions

There were two main objectives of the questionnaires together with their sub-objectives as discussed in detail below. These objectives were derived from the fourth research sub-objective.

6.5.1 BI user questionnaire

The main objective of the BI user questionnaire was to determine the state of BI utilisation in an organisation. In order to fulfil this objective, the questionnaire was structured with the following sub-objectives:

- *To determine the degree of BI practice in organisations.* This was achieved through statements that assessed the software capabilities that support BI solutions.
- *To determine the data management model used in organisations and its prevalence.* This was achieved through statements that assessed the management of the data sources or generating systems in the organisations.
- *To determine the level of understanding of the BI concept.* This was achieved through statements that assessed the location and management of the BI solution in the organisation.
- *To determine the pricing model of BI solutions currently used.* This was achieved through statements that assessed the buying or licensing options for BI solutions used in the organisation.

6.5.2 BI expert questionnaire

The main objective of the BI expert questionnaire was to validate the proposed Hybrid BI architecture by experts in order to suggest improvements. The questionnaire was structured in such a way that it addressed the following sub-objectives:

- *To determine the degree of BI knowledge and ideal procurement option.* This was achieved through statements that assessed the knowledge of BI deployments and the ideal pricing options when procuring a BI solution for the organisation.
- *To evaluate the devices and applications (Layer 1) for validation or improvement.* This was achieved through statements in a section that assessed the ideal front-end layer components of the BI architecture.
- *To evaluate the BI services (Layer 2) for validation or improvement.* This was achieved through statements that assessed the service layer components of the BI architecture.

- *To evaluate the data management systems (Layer 3) for validation or improvement.* This was achieved through statements in a section that assessed the service layer and data layer components for the BI architecture.
- *To evaluate the infrastructure resources (Layer 4) for validation or improvement.* This was achieved through statements that assessed the Infrastructure layer components of the BI architecture.

6.6 Research software tools

In order to avoid incurring extra costs for the research software applications, software applications that were already procured, readily available, and fit for purpose were utilised. Microsoft Word and Google Forms were used to design and develop the soft copy and online survey questionnaire, respectively. Microsoft Excel was used to store responses and analyse the data as it was readily available at no additional cost and was ideal for the sample size.

6.7 Sample design

The researcher had to plan how the sampling for the research study was to be carried out. The following sub-sections discuss the sample design in detail.

6.7.1 Survey population

A research population is the entire set of individuals or organisations, or total number of units from which data needed for research can be gathered (Lind et al 2008:6). This study identified 13 Industries in the South African context as a population. The industries purposely targeted in this study were: Automotive, Finance (Banking, Insurance), Government and Parastatals, Information Communication and Technology, Legal, Logistics and Transport, Manufacturing, Marketing and Media, Medical, Property, Retail, Travel and Tourism, and Other industries not specified (allowing respondents to indicate their specific industries).

6.7.2 Sample size

A sample size of the population is a fraction that is representative of a population from which a statistical analysis can be drawn (Dawson 2002:49). The sample size for the survey study was based on two assumptions.

Firstly, the assumption was that hybrid BI is the correct way to initiate BI deployment and, therefore, the targeted organisations in the above-mentioned specified industries are the ones likely to be using on-premise and, or cloud-based BI. To evaluate the proposed architecture, a sample of 50 BI experts from various industries was targeted.

Secondly, in respect of the assessment of the kind of BI solutions currently used in the industries, the population for this target group was organisations using BI in general. These organisations fall into various categories in terms of their size, that is, micro, small, medium, and large organisations across the targeted industry population. A sample of 50 BI users was targeted from various industries.

6.7.3 Sampling frame

A sampling frame is often used when a formal list of respondents can be utilised for the purpose of sampling the population (Kothari 2004:16). The majority of the sample respondents were selected from an organisation's recruitment database. The email list consisted of potential respondents with designations, such as, Business Intelligence Consultant, Business Architect, Project Manager, Product Manager, Data Analytics expert, Business Analyst, System Analyst, Management, etc. It was borne in mind that these potential respondents are exposed to a type of BI solution in their workplace.

6.7.4 Sampling technique

The survey sample was selected by means of a non-probability purposeful sampling technique. Non-probability sampling ensures that the selected sample of the population constitutes a representative sample of the whole that can be used to generalise the findings (Kothari 2004:59). This sampling technique includes purposive

sampling, convenience or haphazard sampling, judgement sampling, snowball sampling, quota sampling, etc. In order to maximise the potential responses, a snowballing approach was also adopted, whereby the invited respondents were asked to pass on the invitation to colleagues in their network who would meet the requirements. Snowball sampling, also referred to as 'network chain referral' is founded on the idea of a systematic rolling snowball, where a small sample of people is initially used, then a bigger sample is involved through a well-coordinated invitation which ensures that new participants meet the requirements (Rugg & Petre 2007:70).

6.8 Data preparation and cleaning

The survey data for this research was stored, prepared and cleaned in Microsoft Excel. The following are the steps followed during data preparation and the cleaning process:

Table 6-1: Data Preparation and Cleaning Process

Step	Description
1. Create tables	<ul style="list-style-type: none"> Two tables were created: one for coding or numerical value assignment; and the other for storing responses. All responses were stored in a table that was created based on the items questionnaire.
2. Assign numerical values	<ul style="list-style-type: none"> In the coding table, all questionnaire items were assigned numerical values for analysis purposes. Coding ensures that responses can be categorised into limited groups (Cooper & Schindler 2006:443).
3. Data entry	<ul style="list-style-type: none"> In the storage table, the columns represented individual respondents, and rows represented the questionnaire items. The columns were populated with data received from the questionnaires. Google Forms allowed data entries to be exported in an Excel format. This data was copied and pasted into the storage table.
4. Editing	<ul style="list-style-type: none"> Data cleaning was required to ensure data quality from manually completed questionnaires. This assisted in detecting errors and omissions, thus ensuring that data was accurate, consistent, uniformly entered, complete and arranged to simplify coding and tabulation (Cooper & Schindler 2006:441). Limited data cleaning was required as the majority of the responses were from the online link survey which was designed with questions assigned with 'required properties', ensuring the allocated answers were selected and eliminated the possibility of missing values.

6.9 Ethical clearance and considerations

Ethical clearance and considerations ensured that no respondent was abused or suffered adverse consequences from research activities (Cooper & Schindler 2006:116-130). The ethical considerations below applied.

- Participants had the right to participate voluntarily and be able to withdraw at any time without facing any penalties.
- A consent letter ensured that potential participants were fully aware of the procedure of the proposed survey, primarily the right to privacy, confidentiality, and anonymity.
- An ethical clearance was obtained from UNISA to ensure the legitimacy of the research being conducted.

All these ethical considerations ensured that the potential respondents did not suffer from physical harm, discomfort, pain, embarrassment, or loss of privacy when they participated in the research.

6.10 Pilot study

A pilot study is usually conducted to determine errors and weaknesses in the research instrument and to provide representative data for the identification of a possible sample (Cooper & Schindler 2006:76). Rugg and Petre (2007:82) recommend that it is wiser to conduct a pilot study carefully, which includes piloting the data analysis which would provide the researcher with a better idea of how much data will actually be needed and the best way to collect and analyse it.

Before engaging in the actual research, a pilot study was done to find any flaws in the research questionnaire and avoid mistakes that might have been overlooked. The research questionnaire was administered to check and ensure that it did not have any errors, such as, unclear terms, ambiguity, poor wording, and the repetition of questions. The pilot questionnaire was given to four BI Experts and ten BI users from various industries. However, after the analysis of the data gathered, it was realised that there was a high risk that the required results would not be obtained. Therefore, the questionnaires were altered to ensure that all identified concerns were addressed. This was achieved by structuring the majority of the questions in such a way that they

addressed specific components of the architecture under evaluation as discussed in section 6.4 and 6.5 above.

After making the necessary adjustments to the two questionnaires by using the results from the pilot study, a pre-test was again conducted. The outcome revealed satisfactory results which was evidence that the intended objectives would be achieved when the full survey was conducted.

6.11 Study limitations

A research's inadequacies are often not dismissible and it is imperative that the limitations for the study are clearly stated. The limitations for this study are discussed below:

- Due to the fact that BI is a broad subject, this research had limitations with regard to the research areas to be covered. The study only addressed issues related to the defined research questions and research objectives.
- Time was a major factor to ensure that only the elements covered in the scope of the study were handled. Although data collection took a long time to complete in order to receive enough responses, it was difficult to get responses from more than 50 participants. This could have been because of the nature of the BI profession which is a scarce skill as very few people have acquired a BI qualification and are performing other related roles.
- Quantitative research design often attracts a large number of respondents but does not guarantee sticking to the plan. For example, one of the questionnaires was supposed to be answered by respondents with the job title of BI expert, but the questionnaire was administered to participants who had job titles, such as, BI consultant, Business Architect, Business Analysts, etc. who were performing similar roles. If all the respondents were designated BI experts, this limitation could have been avoided.
- A small proportion of qualitative data was collected and analysed respectively. The qualitative research method in general, improves the efficiency of a research study (Kothari 2004:5).

6.12 Conclusion

The purpose of this chapter was to discuss approaches and research instruments used in conducting the survey. Emphasis was given to the purpose of the survey which was to collect data on the kind of BI solutions used in most South African organisations, as well as to validate the proposed BI architecture solution. The chapter commenced with a discussion of the two case studies conducted on BI users and experts, which was followed by a description of the data collection methods used in both cases. The research instruments used were described together with the research objectives. The software tools used were identified which was followed by a discussion of the sample design. Study limitations, ethical clearance considerations and the pilot study were also discussed.

Having achieved the purpose of this chapter of presenting and discussing how the research survey was conducted, the next chapter seeks to present and discuss the analysis of the data collected on the state of BI utilisation in an organisation by BI business users.

CHAPTER 7 STATE OF BI DEPLOYMENT IN ORGANISATIONS

7.1 Introduction

The previous chapter discussed in detail the data collection methods adopted for the survey that was conducted on specific industries in South Africa in order to achieve the research objectives. The survey that was conducted was subdivided into two parts. Part one focussed on the state of BI utilisation by BI users, and part two focused on validation of the hybrid BI architecture by BI experts.

The purpose of this chapter is to present and discuss the data analysis and the findings of part one of the survey conducted to determine the current state of BI utilisation in organisations within specific industries in the South African context. The chapter contributes partly to answering the fourth research question on the validation of the proposed hybrid model. Although this part of the survey was not specific to the study's proposed hybrid model, it nonetheless contributed to an understanding of the relevance and maturity of BI at organisational level. In the researcher's view, the validation of the proposed hybrid model was very important in ascertaining that industry-specific deployment of a BI solution is of significant value to businesses.

This chapter is structured as follows: section 7.2 discusses the quantitative and statistical data analysis. The results for part one of the research survey are discussed in section 7.3 and the summary discussion of the findings is presented in section 7.4. The chapter is concluded with a summary discussion in section 7.5.

7.2 Data analysis

Triangulation analysis ensures that a combination of quantitative and qualitative analysis complement each other to obtain two or three views about the aspects being studied (Olsen 2004:4). For data analysis, triangulation analysis was utilised to ensure that three kinds of information were covered, namely, the literature review, and quantitative and qualitative data obtained during the survey conducted. Triangulation analysis encourages the collaborating of data and, or research methods with the intention of using different views to validate the findings and to ensure that it is interdisciplinary and holistic. Triangulation of data sources and data types supports the norm in case study research that the subject area or item should be examined and

explored from various viewpoints (Baxter & Jack 2008:556). Thus, during the analysis process, data gleaned from the survey questionnaire responses, and information found in the various sources, that is, the literature review were integrated so that a thorough understanding can be reached. In the following subsections, the data analysis approaches, namely, the quantitative, qualitative and statistical data analysis that constituted the triangulation analysis are discussed in detail.

7.2.1 Quantitative data analysis

Quantitative data analysis is the measurement of numerical data and is expressed in terms of statistics (Kothari 2004:3). Quantitative analysis ensures that statistical analysis can be conducted to organise data in order to identify differences or similarities, and, or provide simple descriptive data. However, in some cases a research study may need to find out if there is a connection between a number of variables and that would require advanced and complex statistical analysis (Dawson 2002:126). Quantitative analysis was carried out for this part of the survey because most of the data collected was in numeric form, making it easier to perform statistical analysis. A descriptive quantitative approach was used to collect, analyse and describe the factors that contribute to assessing the state of BI utilisation in organisations.

7.2.1.1 Data preparation process

The steps that were followed during data analysis are shown in table 7-1.

Table 7-1: Data Preparation Process

Step	Description
1. Finalise data entry	<ul style="list-style-type: none"> Data was deemed ready for analysis after all data entries were included in the storage tables discussed in the previous chapters. This occurred after the responses that were obtained from the online survey, soft or hard copies were closed.
2. Analyse data	<ul style="list-style-type: none"> Once the data was ready for analysis after the preparation process discussed in the previous chapter, the pie charts and bar charts were drawn up to analyse and present the research findings. A small proportion of qualitative data was analysed contextually. This was the data in response to the few questions that required respondents to provide suggestions in their own words.

Step	Description
3. Present results	<ul style="list-style-type: none"> • For the purpose of analysis, the Likert scale items were grouped to obtain a majority score comprising responses in respect of the “Strongly agree and Agree”, and “Disagree and Strongly disagree” options. A majority score was used to represent the respondents’ selection which meant that they were in total agreement or disagreement with a particular statement. • All the results were presented quantitatively and were supported with a comprehensive qualitative explanation.

7.2.1.2 Statistical data analysis

Statistical data analysis can be performed in two forms, namely, descriptive and inferential statistics (Lind et al 2008:6; Wegner 2007:9). According to Lind et al (2008:6) descriptive statistics include categorising data using various methods, drawing conclusions, and displaying data or results in the manner that is self-explanatory. Inferential statistics are often used to generalise the sample findings to the broader population (Wegner 2007:9). This study used descriptive statistics to organise, summarise and present data in an informative manner and generalise the findings. The response data was statistically analysed and presented using Microsoft Excel. The flexibility of using exploratory data analysis was useful as it allowed presenting the results from a descriptive statistics analysis perspective (Cooper & Schindler 2006:472). This was achieved through the use of percentages to display the frequency of values in bar charts and pie charts and, thereafter, a confirmatory data analysis was conducted to explain the evidence found. The findings from the sample were later used in the findings section to generalise the results to the broader population (Wegner 2007:9).

7.2.2 Qualitative data analysis

Qualitative data analysis is an interpretative approach to data analysis that uses inductive reasoning aimed at obtaining an in-depth understanding of human behaviour and the reasons for occurrences (Khalid et al 2012:16; Cooper & Schindler 2006:472). Its primary aim is to deliver a full interpretation of the phenomena through understanding and interpreting social interactions. In this survey’s outcomes, a small proportion of data collected was qualitative which was required to be analysed qualitatively so that it could complement the quantitative data analysis. Content

analysis was applied to identify common patterns or features that were sought in the survey conducted (Dawson 2002:118).

7.3 Results of part 1: BI business users

This section presents and discusses the data analysis obtained from responses to the questionnaire directed at BI users. The following sub-sections discuss the analysis of the individual questions and present the results of the data analysis both quantitatively and qualitatively.

7.3.1 Response rate

A total of 50 questionnaires were administered for completion by BI users in the 13 industries listed in the previous chapter. Out of the 13 industries, responses were received from nine industries, achieving a 69% response rate. Thirty-three (33) participants responded to the 50 questionnaires, achieving a 66% response rate.

7.3.2 Section 1: Demographic information

All the targeted questionnaire respondents were BI business users which ensured that the best opinions on the questions were received. The demographic questions with their responses provide the necessary background information below.

7.3.2.1 *User's years of experience*

Question 1 concerned respondents' years of experience in their field. Figure 7-1 shows that 46% of the responses were obtained from respondents with two to four years' experience, 30% were obtained from respondents with between five and seven years' experience, and 9% of the respondents had between eight and 10 years' experience. Only 15% had more than 10 years' experience. Over one half (54%) of the responses were obtained from respondents with more than four years of experience. This means that the respondents had a fairly broad range of experience which ensured that they were able to provide the required feedback on the utilisation of BI solutions.

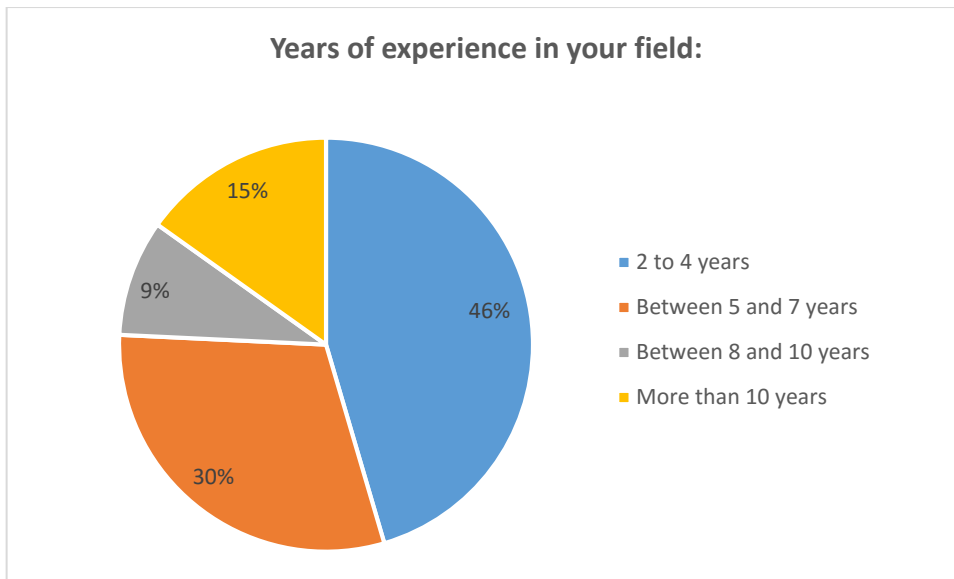


Figure 7-1: Years of experience

7.3.2.2 Size of organisation

Question 2 covered the size of the organisation in which the respondent was employed or to which the respondent was associated. Figure 7-2 shows that 43% of the respondents were employed by medium sized organisations. One-third (33%) indicated that they are employed by large organisations, 18% are employed by small and 6% by micro organisations. Responses were received from all categories of organisations giving this survey the opportunity of receiving a variety of opinions from different sizes of organisations. The results shows that all types of organisations use a form of BI application and the majority of them been from the medium and large organisations. Micro and small organisations had the minority responses which could suggest that they may be using simple BI applications in their organisations.

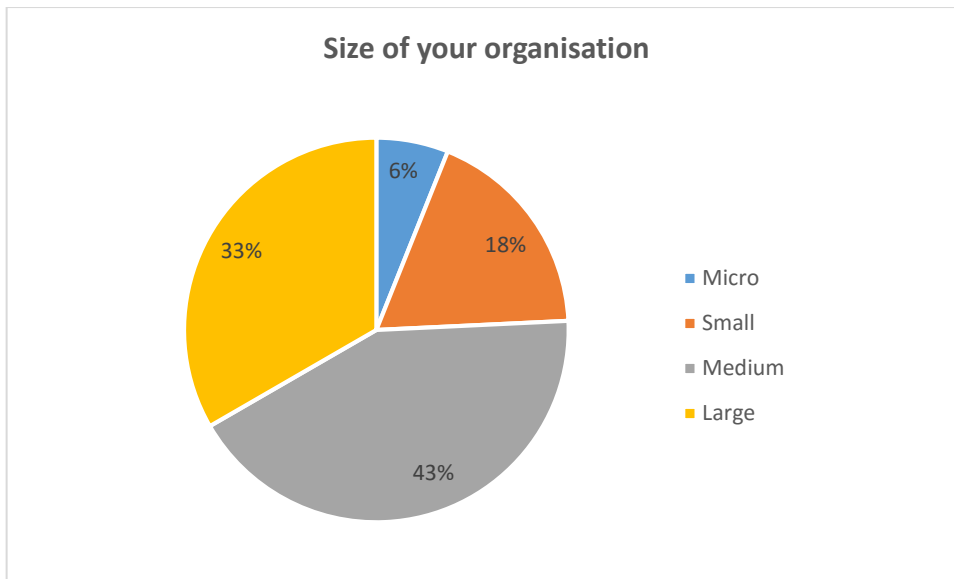


Figure 7-2: Organisation size

7.3.2.3 Kind of industry

Question 3 dealt with the kind of industry in which the organisation was involved. Figure 7-3 shows that 24% of the respondents were employed in the fields of Finance (Banking, Insurance) and the Information Communication and Technology industries respectively. Less than one-quarter (21%) of the responses were employed by Government and Parastatals. Only 6% of respondents were employed in the manufacturing, Medical, Retail and Other (i.e. "Business services", "IT and Business Consultancy", and "Medical services") industries, respectively. A combined percentage of 6% of respondents were employed in the Automotive, and Marketing and Media industries. The response rate for BI business users was 69%. This result allowed a fairly satisfactory representation of opinions from the respondents to be recorded from a number of different industries.

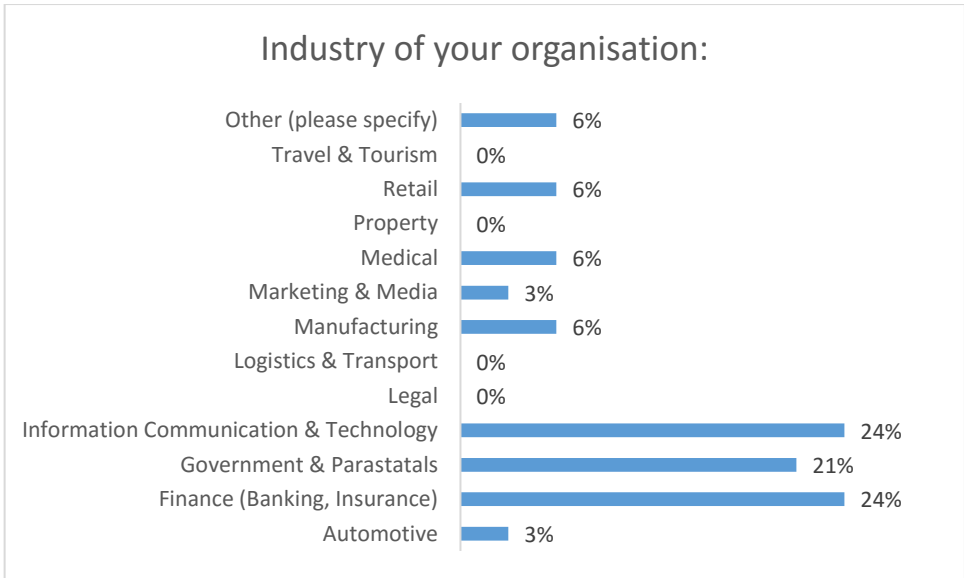


Figure 7-3: Industry of organisations

7.3.2.4 Number of business users with access to BI services

Question 4 dealt with the number of business users who had access to BI services. Figure 7-4 shows that 43% of the respondents were employed by organisations with two or less BI users. Thirty-nine (39%) of respondents were associated with organisations with three to five BI users, 6% were employed in organisations with six to ten BI users, and last but not least, 12% were associated with organisations with more than ten BI users. The majority of the respondents represented organisations with more than two users with access to BI services.

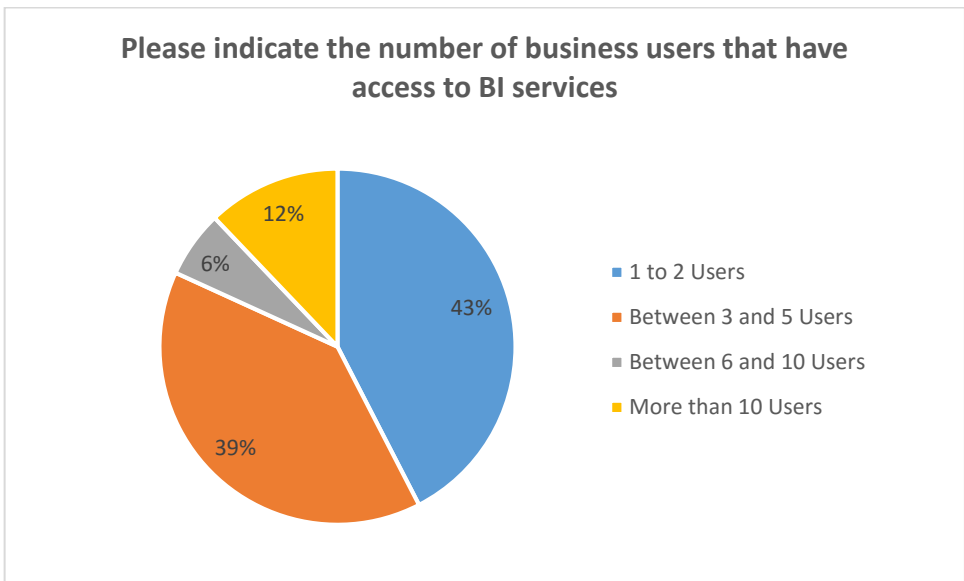


Figure 7-4: Users accessing BI services

7.3.3 Section 2: Business intelligence deployment

In order to shed light on the current BI solutions utilised by business users in their organisations, the respondents were required to assess their BI solutions by choosing particular options indicating whether they agreed or disagreed with certain statements.

7.3.3.1 *Assessment of Software Capabilities Supporting BI Solution*

Question 1 asked respondents to assess different kinds of software tools that support their BI solution. Figure 7-5 shows that nearly two-thirds (64%) of the respondents indicated that their BI solution could be used by non-BI professionals or experts. However, they also indicated that two-thirds (66%) of the current users are experienced BI professionals. Half (51%) of the respondents indicated that utilising self-service BI solutions ensures timely insight and decision-making. In terms of making informed decisions and performing analytical functions, the majority (85%) indicated that their reporting and analysis tools are able to meet their requirements.

As far as complexity is concerned, 63% of the respondents use complex BI tools, such as SAS BI, SAP BI and Oracle BI and 60% indicated that they use simple tools such as Microsoft Excel and Google Sheet to store data, retrieve, analyse and produce reports. More than two-thirds (70%) of the respondents indicated that they have multiple data generating systems and multiple users, and this was also supported by (55%) of the respondents who indicated that they did not have a single user and system. When it comes to access to BI solutions, 45% indicated that they have a single point of entry to their BI solution. These responses suggest that a certain type of BI solution is utilised by different types of organisations. This also means that organisations are not using simple BI solutions, but rather have structured and, or advanced BI solutions, such as, SAS BI, OLAP server, and so on.

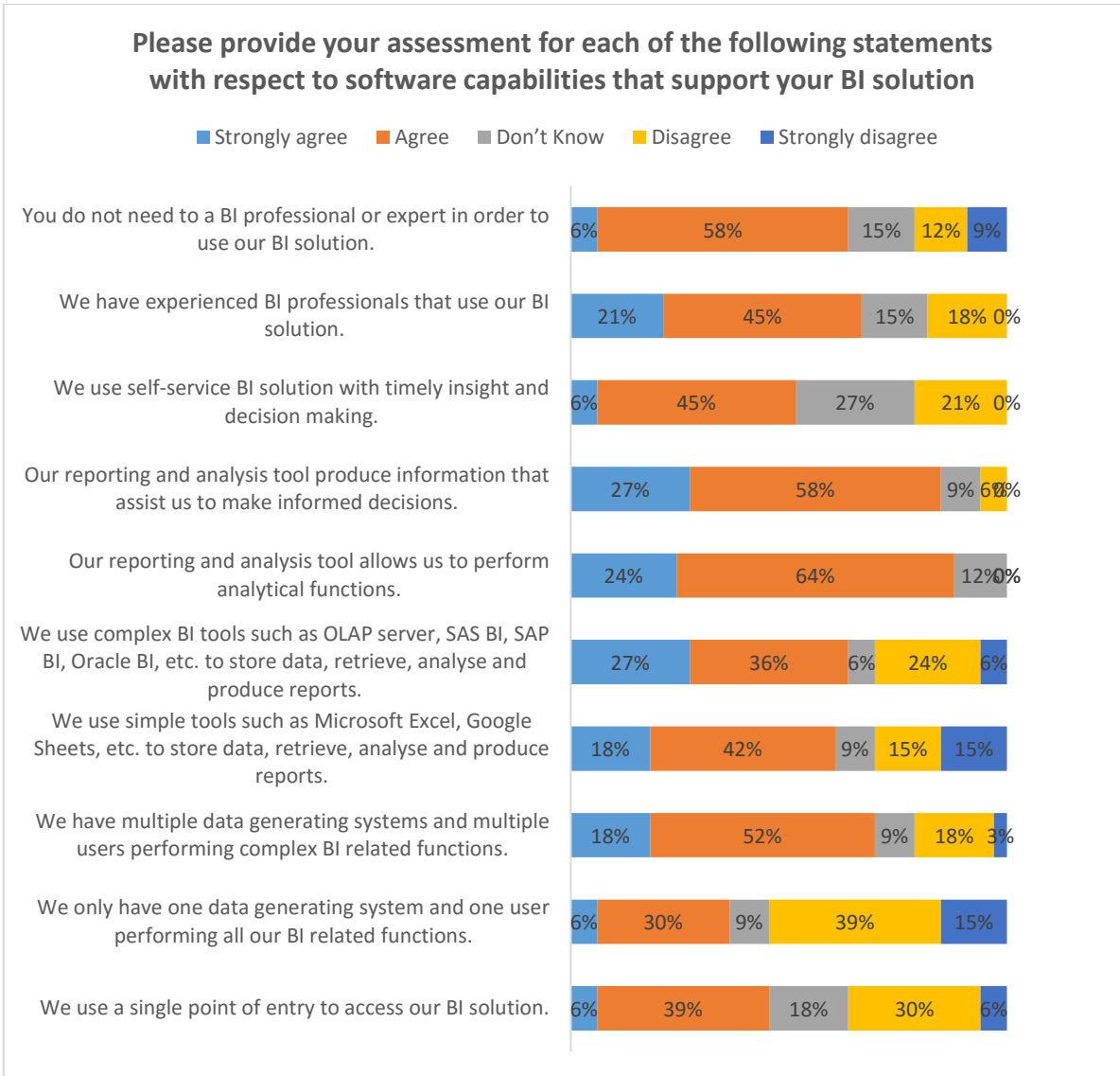


Figure 7-5: Software capabilities supporting BI solution

7.3.3.2 Assessment of Management of Data Sources/Generating Systems

In question 2, respondents were asked to assess the management of data source or generating systems for their BI solutions. Figure 7-6 shows that the majority (76%) of the respondents agreed that their data is managed through OLTP systems and their data source systems are managed internally. This was supported by half (51%) of the respondents who did not agree that their data source systems are managed by an external service provider. With regards to the type of data storages utilised, more than half (57%) indicated that they use a simple database while 48% indicated using complex data warehouse.

These data stores are mainly located at a central repository on-premise as indicated by more than two-thirds (70%) of the respondents. On the other hand, 42% of the respondents use cloud-based storage hosted by their service providers. As far as data accessibility is concerned, 42% indicated easy access to data repository hosted by the service providers compared to 39% that disagreed. Before this data is accessed, 69% indicated that it undergoes thorough data cleansing (ETL).

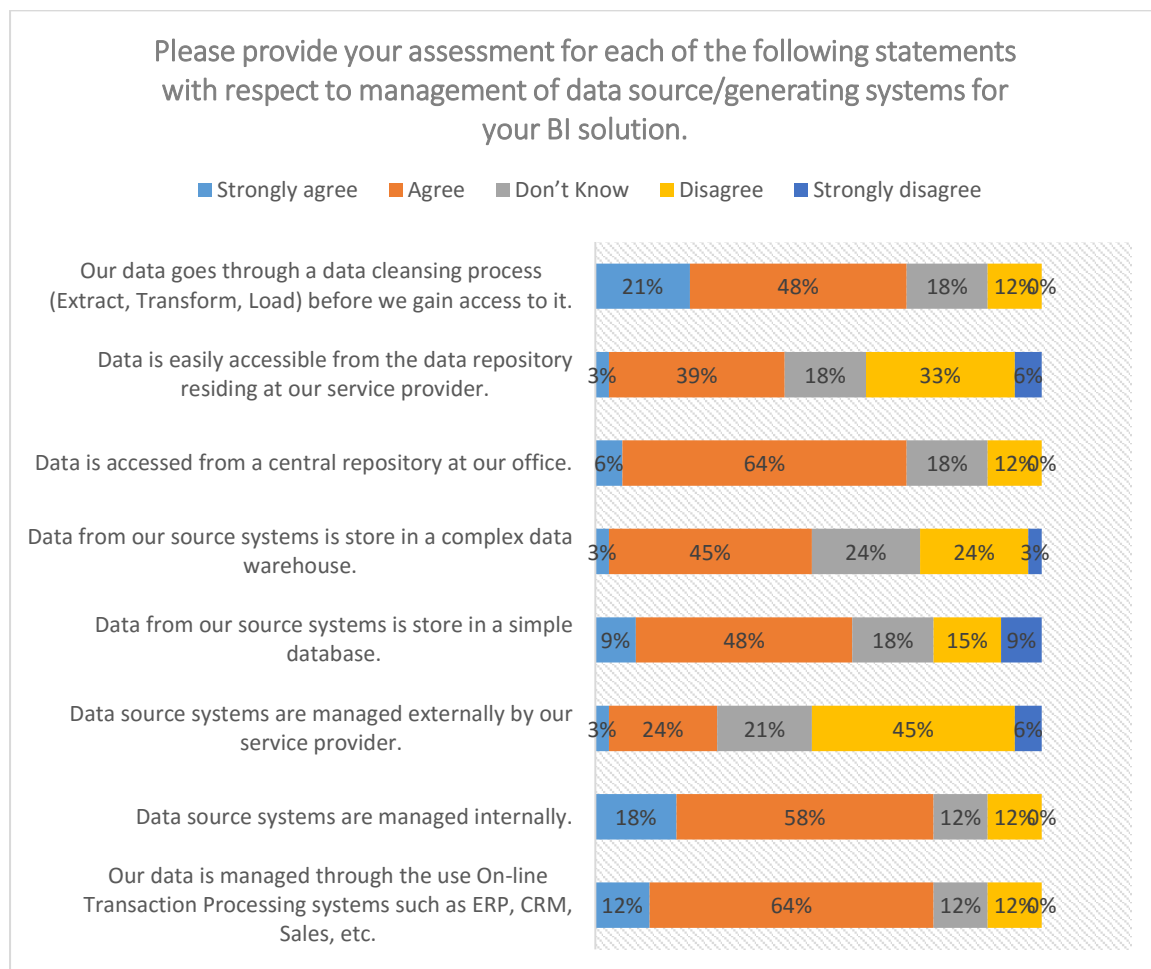


Figure 7-6: Management of data source or generating systems

7.3.3.3 Assessment of Location and Management of BI Solution

In question 3 respondents were asked to assess the location and management of their BI solutions. BI solutions are often installed on-premise and, or cloud-based. In this study figure 7-7 shows that 60% of the respondents indicated that their BI is located on-premise, whereas 45% indicated that they are hosted by the service provider. With regard to technology supporting BI solutions, 48% of the respondents indicated that they own the hardware and software compared to 30% that disagreed. More than two-

thirds (70%) of the respondents agreed that they have full control of the data processing applications and development environment, but nearly half (48%) also indicated having full control in their hosted environment.

With regard to the hosted storage and computing resources, 42% indicated that they do not have full control of their resources compared to one-third (33%) that have full control. On the other hand, the majority (76%) of the respondents indicated that they have full control of storage and computing resources residing on-premise.

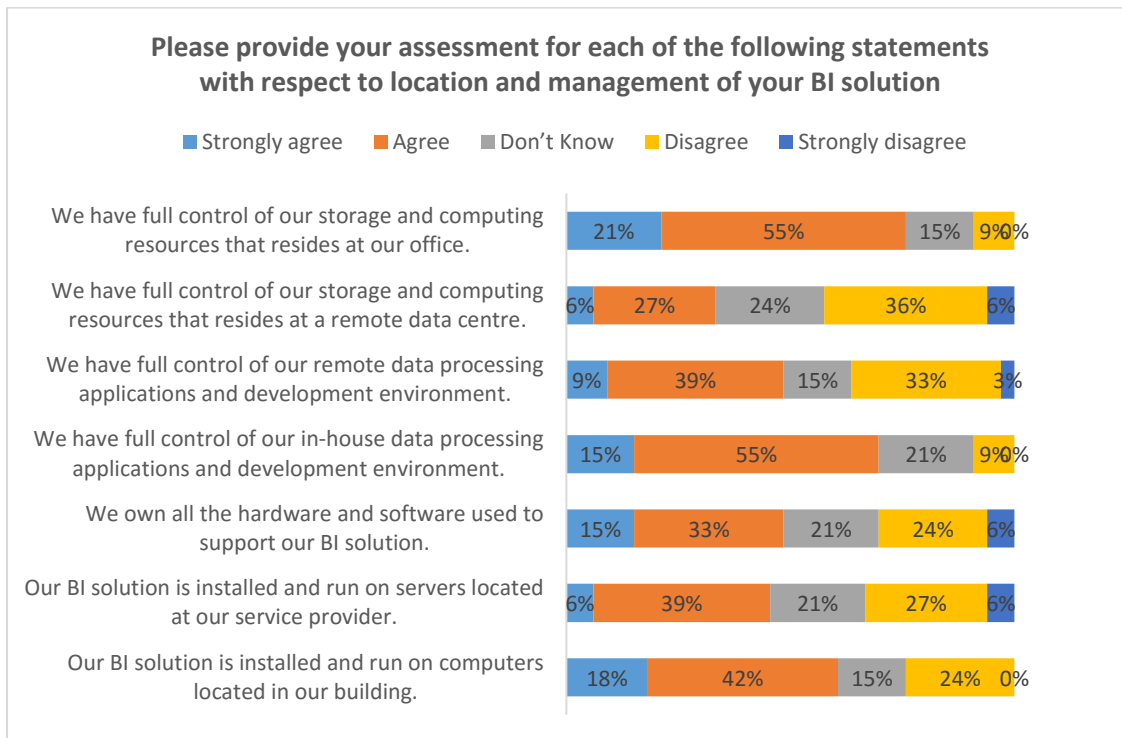


Figure 7-7: Location and management of BI solutions

7.3.3.4 Assessment of buying or licensing options for BI solution

In question 4 respondents were asked to assess their buying or licensing options. BI solutions are offered in respect of different pricing models. Figure 7-8 shows that 72% of the respondents indicated that they do not use a free version of any available BI applications. Nearly one-half (48%) indicated that they did not buy their BI solution from the service provider that offers it off-the-shelf. Less than one-third (30%) indicated using pay-as-you-go, whereas 42% did not agree. Forty-two (42%) percent indicated not using pay-per-user compared to 33% that agreed. Only 18% indicated using pay-per-use compared to 57% that did not use it. Over half (54%) indicated that they use

a subscription-based account, with 21% that did not. On average, 25% of the respondents did not know what pricing model was used by their organisation. These results indicate that there was certainly a fee that is payable for the BI users to use their BI solution.

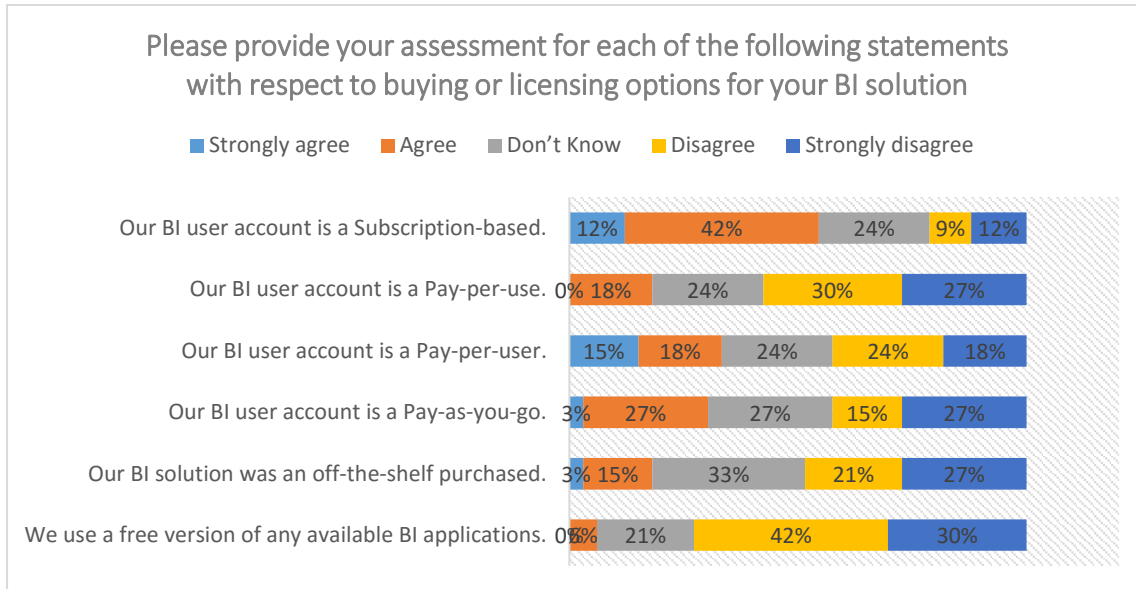


Figure 7-8: Buying or licensing options

7.4 Findings part 1: state of BI utilisation in organisations

The following is a summary of the key findings related to each sub-objective of the survey on assessing the state of BI utilisation in organisations in the South African context.

7.4.1 Degree of BI practice in organisations

In order to determine the degree of BI practice in organisations, the respondents were asked questions related to software capabilities that support their BI solution. The key findings are presented below.

- The majority of the respondents were associated with medium and large organisations. It was discovered that a variety of BI solutions are utilised in the different types of industries so that these organisations can meet their BI needs and expectations. Organisations have at least one user using a type of BI solution that is either simple or complex. The business user(s) did not have to be in the BI profession in order to use their BI solutions.
- It was found that some simple or complex BI solutions possess self-service capabilities making it easy for business users without BI skills to utilise the

applications to store data, retrieve, analyse and produce reports. The output from the BI solution was found to be useful as it provided insightful information for better decision-making. It was also found that multiple users utilised data for reporting that was generated from multiple data generating systems. This data was easily accessed through a single point of entry.

In summary, the utilisation of BI practice in organisations was found to be common in various industries. The findings also suggest that most of the organisations are not using a naïve BI model to set up the BI solutions, but they have a well-structured or advanced BI solution that resides either on-premise, in the cloud, or both. These BI solutions were found to be comprised of significant components related to data source systems, storage capabilities, data analysis services, reporting and information delivery (Popovi et al 2010:9).

7.4.2 Data management model used in organisations and its prevalence

In order to determine the data management model used in organisations, the respondents were asked questions related to the management of data sources or generating systems for their BI solutions. The key findings are presented below.

- The survey found that the majority of organisations utilise OLTP systems, such as, ERP, CRM, and Sales applications to manage their data sources internally. Furthermore, it was found that before data is used it undergoes a thorough data cleansing (ETL) process in preparation for analysis and reporting.
- It was also discovered that the data generated by the OLTP systems is either stored in a simple, or complex database. Databases were found to have been installed in a central repository that resides either on premise, or in the cloud.
- Access to this data repository residing on-premise or in the cloud was found to be easy, which means that organisations are able to manage data effectively in order to meet their BI requirements.

In summary, it was shown that for organisations to effectively manage their data sources, a variety of OLTP systems, ETLs, and simple or complex databases residing in a data repository were used. It is necessary that before data is stored in the relevant

DBMS, it undergoes a thorough ETL process so that it is integrated and transformed into the format that is required to meet BI needs (Manyika et al 2011:32). Although some of the data management systems were found to be stored on- premise or in the cloud, they were all managed internally.

7.4.3 Level of understanding of the BI concept

In order to determine the level of understanding of the BI concept in organisations, the respondents were asked questions related to the location and management of their BI solution. The key findings are presented below.

- All BI solutions are normally securely installed in a location that meets the requirements of the organisations. This survey revealed that most of the organisations have their BI solution installed on-premise, or in the cloud, or both.
- In an on-premise deployment, the technology resources that support a BI solution is owned by the organisations and they are aware of having full control of resources, such as, storage, computing, data processing applications, application development environment, and so forth.
- It was also discovered that in a cloud-based deployment, similar ownership benefits are encompassed, and organisations are able to exercise full control of their technology resources.

In summary, the level of understanding of the BI concepts, was found to be high because most organisations demonstrated the utilisation of on-premise and, or cloud-based BI solutions, and had full control of the technological resources associated with the solutions. Although the actual cost associated with on-premise or cloud-based BI solutions were not covered in detail, the utilisation of such solutions on-premise is believed to come with higher upfront investment and ongoing operational costs (Gastermann et al 2015:576).

7.4.4 Current pricing models of BI Solutions

In order to determine the pricing models of BI solutions currently used in organisations, the respondents were asked questions related to the buying or licensing options of their BI solutions. The key findings are presented below.

- As in the case of any other software applications, BI solutions are also procured through a certain pricing structure. It was discovered that some of the organisations have procured their BI solutions through the utilisation of a subscription-based pricing model which is not necessarily 'pay-as-you-go', 'pay-per-user', and, or 'pay-per-use'.
- It was found that most organisations do not subscribe to using free versions of any available BI applications that are either from 'off-the-shelf' or cloud-based solutions. Although the various pricing models presented are known, a very small percentage of organisations were found to be using BI applications procured through pay-as-you-go, pay-per-user or pay-per-use.
- It was discovered that a fair percentage of business users did not know the pricing model utilised by their organisations. However, they knew that there is a fee associated with the use of the BI application which is payable. This could also mean that these organisations are utilising leased pricing models that were not covered in this research.

In general, the conclusion could be drawn that organisations were using BI solutions that have been procured through a certain pricing model that meets their requirements. Although a certain fee was payable for the procured BI solution, most of the pricing models presented to the business users were not commonly used, except for the subscription-based model. This could also mean that organisations are procuring their BI solutions through in-house custom development or contracted, that is, rented or leased on a 'when and as needed' basis (Daneshgar et al 2013:1743).

7.5 Conclusion

The purpose of this chapter was to present and discuss the findings of the survey conducted within specific industries in South Africa which attempted to determine the

current state of BI utilisation. Firstly, the data analysis method covering quantitative and statistical data analysis was discussed. Secondly, the results of the analysis covering the response rate, demographic information and the state of BI deployment were discussed in detail. Finally, a broad discussion of the findings was presented.

Having dealt with the data analysis and findings of part one of the survey, the next chapter seeks to present and discuss the analysis of the results of the second part of the survey which focused on validating the proposed Hybrid BI architecture by questioning BI experts in order to improving the it. The improved architecture will also be presented.

CHAPTER 8 AN IMPROVED HYBRID BI ARCHITECTURE

8.1 Introduction

In the previous chapters the data that was collected during the first part of the survey were analysed and discussed in detail. This chapter continues with the analysis of the case studies by presenting and discussing the findings of the second part of the survey which focused on validating the proposed hybrid BI architecture by experts for plausible improvements. The proposed BI architecture is the main contribution in this research study.

The purpose of this chapter is to present and discuss the data analysis and the findings of the survey conducted on validating proposed BI architecture by experts within specific industries in South Africa as well as present an improved BI architecture which is based on the contributions from BI experts.

This chapter is structured as follows: section 8.2 discusses the data analysis briefly covering quantitative and statistical data analysis; the results for part two of the research survey are discussed in section 8.3; then the summary discussion of the findings is presented in section 8.4; followed by a discussion on an improved BI architecture and a practical demonstration in section 8.5 and section 8.6, respectively. The chapter will be concluded with a summary discussion in section 8.7.

8.2 Data analysis

In order to get the full benefits of the rich information in the literature review, quantitative, qualitative, and triangulation analyses were utilised (Dawson 2002:20). The weakness of quantitative and qualitative analysis can be neutralised using triangulation analysis. Quantitative analysis and qualitative analysis were applied due to the fact that each case study had specific objectives to be met through the collection of numerical and expressive data emanating from the questionnaire responses (Zainal 2007:4). Quantitative analysis was applied to analyse most of the data collected in numerical form. Qualitative data analysis was used to recognise and interpret the responses so that an in-depth understanding could be achieved (Khalid et al 2012:16).

Descriptive statistics was used to categorise, draw conclusion and display data or results which was later used to infer or generalise the sample findings to the broader population (Wegner 2007:9; Lind et al 2008:6). The statistical data was analysed and presented in percentages to display the frequency of values in bar charts and pie charts, and, thereafter, a confirmatory data analysis was conducted to explain the evidence found.

8.3 Results of part 2: BI experts responses

This section presents and discusses the data analysis of the responses to the questions in the questionnaire directed at BI experts. In the following sub-sections analyses the responses to the individual questions and presents the results of the data analysis both quantitatively and qualitatively.

8.3.1 Response rate

A total of 50 questionnaires were administered for completion by BI experts in the 13 industries listed in chapter 6. Of the 13 industries approached, eight industries responded so that a 61% response rate was achieved. Of the 50 questionnaires distributed, 31 participants responded achieving a 62% response rate.

8.3.2 Section 1: Demographic information

All the questionnaire respondents were BI expert professionals who were carefully selected to ensure that the best expert opinions about the issues raised in the questions posed were received. The responses to the demographic questions are presented below.

8.3.2.1 *User's Years of Experience*

Question 1 concerned respondents' years of experience in their field. The majority (84%) of the respondents had more than five years of experience in their fields as shown below in figure 8-1. These respondents possessed a reasonable average

number of years of experience in order to provide an expert opinion in the validation process of the proposed BI architecture. This also supports the assumption that the respondents had a good understanding of BI deployments in their organisations.

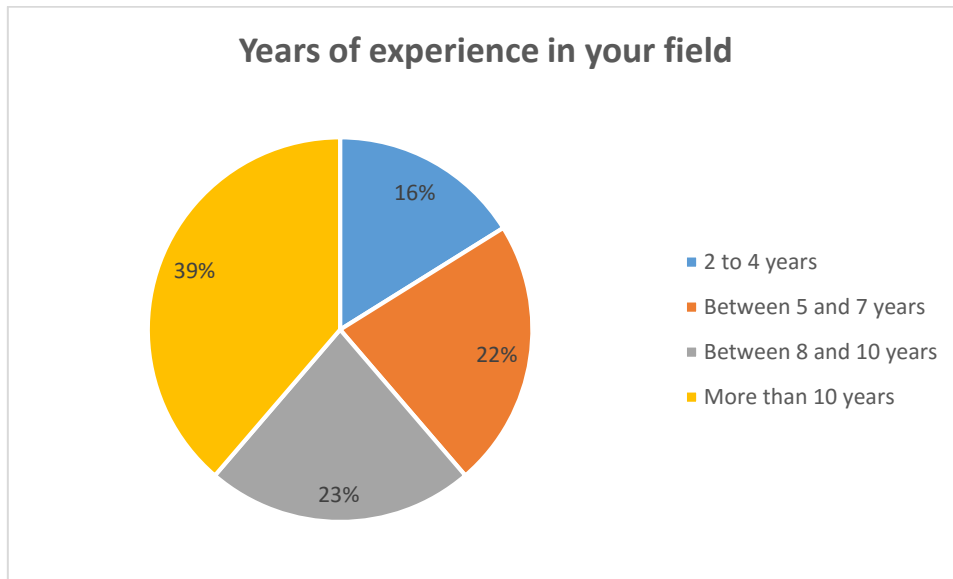


Figure 8-1: Years of Experience

8.3.2.2 *Kind of Industries*

Question 2 dealt with respondents' kind of industries they represented. Sixty-one (61%) of the industries surveyed were represented by at least one respondent. Nearly one-half (45%) of the respondents represented the Information Communication and Technology Industry, with only 16% employed in Finance (Banking, Insurance), 13% in Government and Parastatals industries and a combined total of 18% of the respondents were represented in the fields of Logistics and Transport, Manufacturing, Property and Retail as show in figure 8-2 below. The rest were employed in other industries not listed, namely, "Business Services", and "IT and Business Consultancy" industries. The fact that the industries represented were dispersed among various fields enabled the researcher to obtain a broad understanding of the views of BI experts from diverse industries as opposed to one.

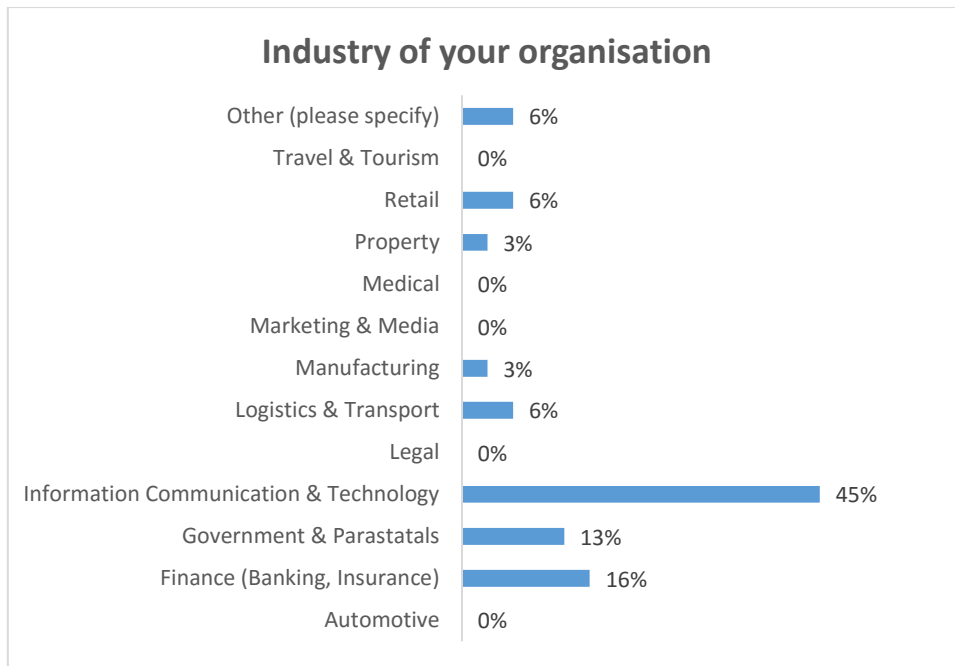


Figure 8-2: Industry of Organisation

8.3.2.3 *Level of Deployment of BI in Organisations*

Question 3 covered respondents' level of deployment of BI in organisations. Figure 8-3 shows that there are three commonly used levels of deployment of BI solutions. Nearly two-thirds (65%) of the respondents indicated that they have deployed a centralised BI solution. One-quarter (27%) of the respondents deployed distributed BI solutions, while the rest chose an isolated solution. A few respondents indicated a combination of centralised and distributed and, or isolated BI solutions. For instance, 16% of the respondents who indicated that they use centralised solutions also indicated that they use distributed BI solutions. From this analysis, it could be deduced that organisations are using a combination of centralised and distributed BI solutions in order to meet BI requirements and expectations.

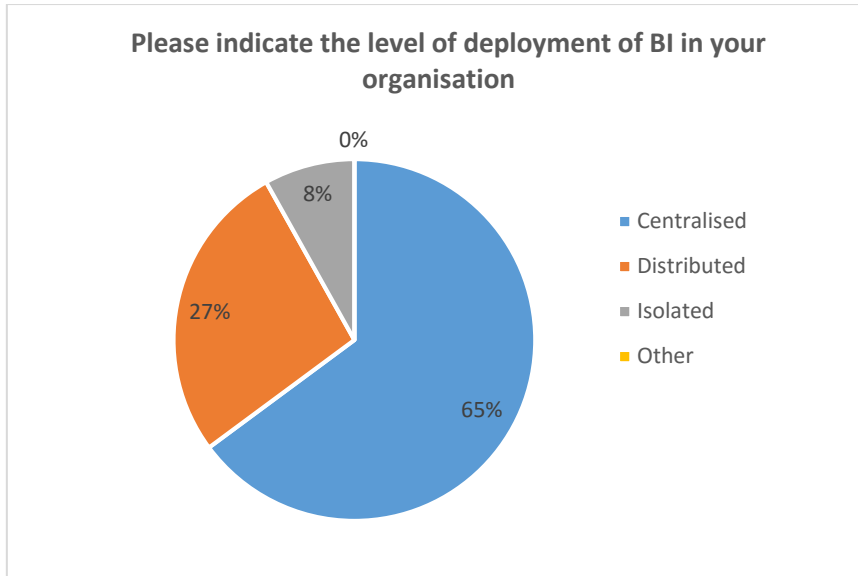


Figure 8-3: Level of Deployment of BI

8.3.2.4 Geographical distribution of organisations

Question 4 dealt with respondents' geographical distribution of organisations. Figure 8-4 shows that half (49%) of the respondents were located in the Gauteng province with at least 17% of them indicating that they have other offices across South Africa, or internationally. Only 7% indicated that they are an international organisation. These results show that most respondents were employed in organisations that are located within South Africa, and, therefore, may experience quite an amount of data integration and management across branch offices.

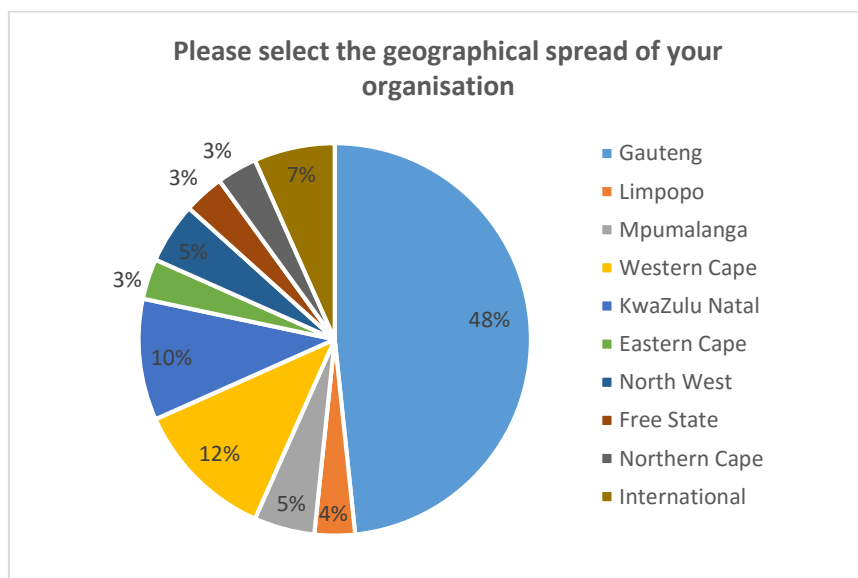


Figure 8-4: Geographical Distribution of Organisations

8.3.3 Section 2: Familiarity with BI deployment and pricing options

Most of the organisations these days are deploying a combination of BI solutions, such as, on-premise, cloud-based and, or hybrid solutions in order to manage their large volumes of data, critical information, and to make informed decisions, among other benefits. In order to get an overview of their BI deployment and pricing options, the respondents were asked to respond to various statements in regard to their expertise and knowledge of BI deployment options. The results are presented below.

8.3.3.1 Assessment of expertise and knowledge of BI deployment options

As shown in figure 5, nearly all (87%) of the respondents agreed that an on-premise BI solution has all the hardware, software and application systems located within the organisation and 87% agreed that the use of an on-premise BI solution ensures ownership of business data. More than two-thirds (71%) agreed that cloud-based BI deployment uses SaaS, PaaS and IaaS cloud computing models. The majority (81%) agreed that use of cloud-based BI solutions promotes the use of faster and cheaper BI services. Most (87%) respondents agreed that a hybrid BI is a combination of on-premise and cloud BI and 87% agreed that a hybrid BI solution allows for sensitive information to be kept on-premise and managed in the cloud. These results suggest that a combination of on-premise and cloud-based BI is ideal for organisations and benefits can be realised from both deployments.

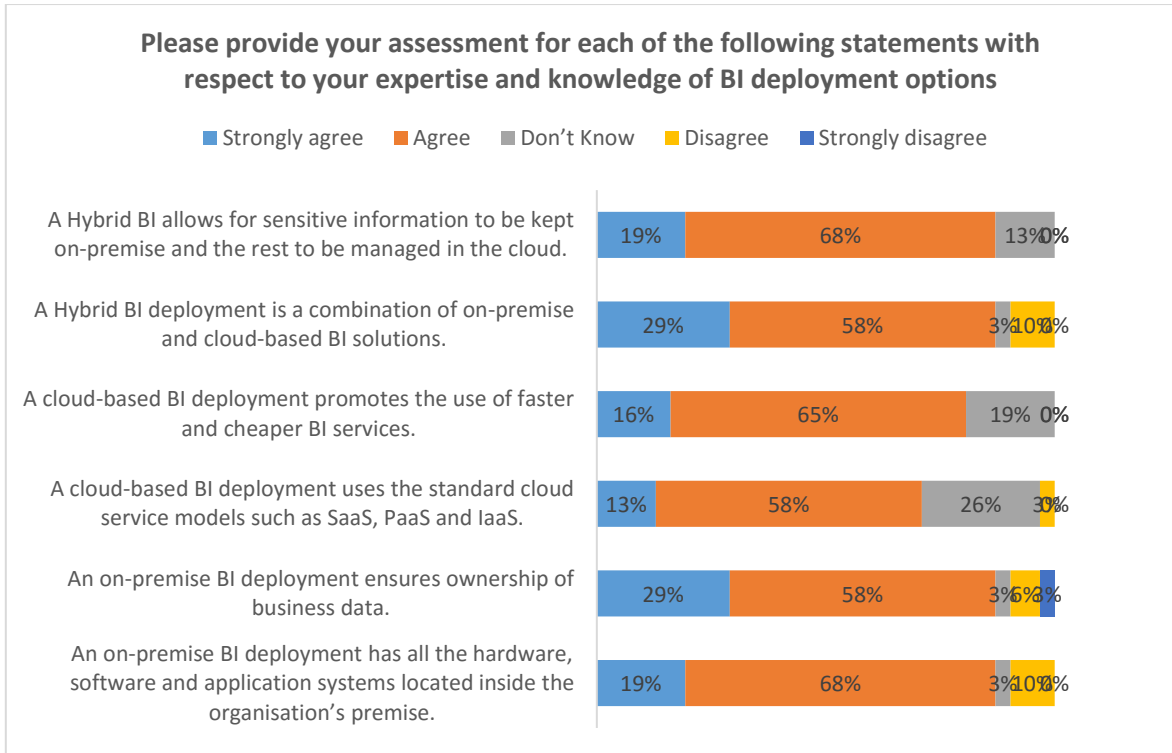


Figure 8-5: Expertise and Knowledge of BI Deployment Options

8.3.3.2 Assessment of pricing models BI solutions

Respondents were asked to give their assessment of pricing models for BI solutions. Figure 8-6 shows that 74% of the respondents indicated that subscription-based pricing is the ideal option to take when acquiring a BI solution. More than one-half (58%) indicated server-based, 64% indicated pay-per-use, 71% indicated pay-per-user, and 68% indicated pay-as-you-go as the ideal pricing model. However, as far as the once-off purchase option is concerned, 64% were not in favour of purchasing the BI solution once-off. These results suggest that various organisations use different types of pricing models irrespective of which type of BI has been deployed.

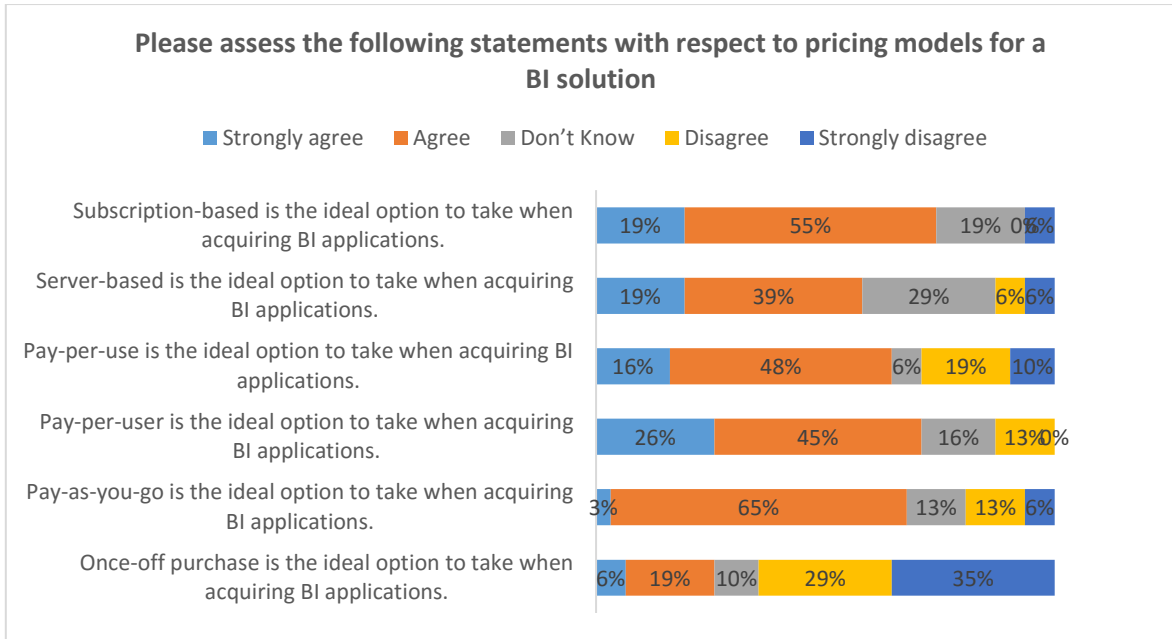


Figure 8-6: Ideal Pricing Models for BI

8.3.4 Section 3: Front-end layer

A front-end layer of a BI architecture acts as the entry point or interface from where business users can access data for enterprise reporting through the use of devices, BI tools and, or applications that reside either on-premise or in the cloud. In order to get an overview of respondents' understanding of devices, BI tools and, or applications that are good candidates for a BI architecture they were asked to assess certain statements in respect of a front-end layer of a BI architecture. The interpretation of the responses is presented below.

8.3.4.1 Assessment of BI architectural front-end layer

As shown in figure 8-7, 90% of the respondents suggest that the interface layer should form part of the front-end layer while 87% selected the presentation layer. The majority (87%) are in favour of having BI tools in the front-end layer. Most (94%) agreed with having applications, such as, OLAP, Data Mining, etc. in the front-end layer. Most (97%) indicated that the devices to be used must be considered when designing the front-end layer. This analysis suggests that a good front-end layer in a BI architecture must consider an interface layer, presentation layer, BI applications and tools, and devices when designing the architecture.

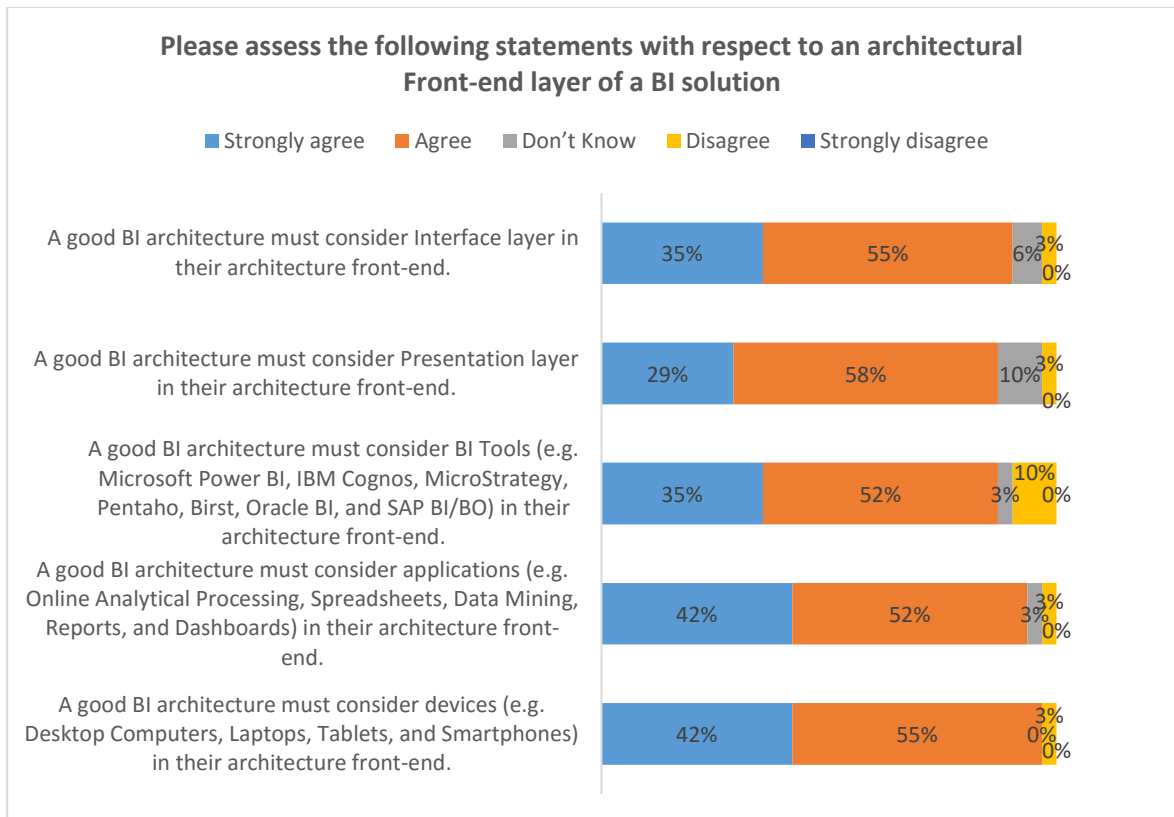


Figure 8-7: Front-end Layer – Proposed Components

8.3.4.2 Additional components for a front-end layer

Figure 8-8 shows that when respondents were asked which additional components should form part of the front-end layer, 39% indicated Business Process Execution Language should be included, 35% selected business process tools, 25% selected Workflow tools, and 4% other components (“Databases, network, location of databases, amount of data in the databases, frequency of collection of data”, “BAPI”). These responses suggest that organisations could be flexible with the types of components which they could include in their front-end layer in order to meet their particular requirements.

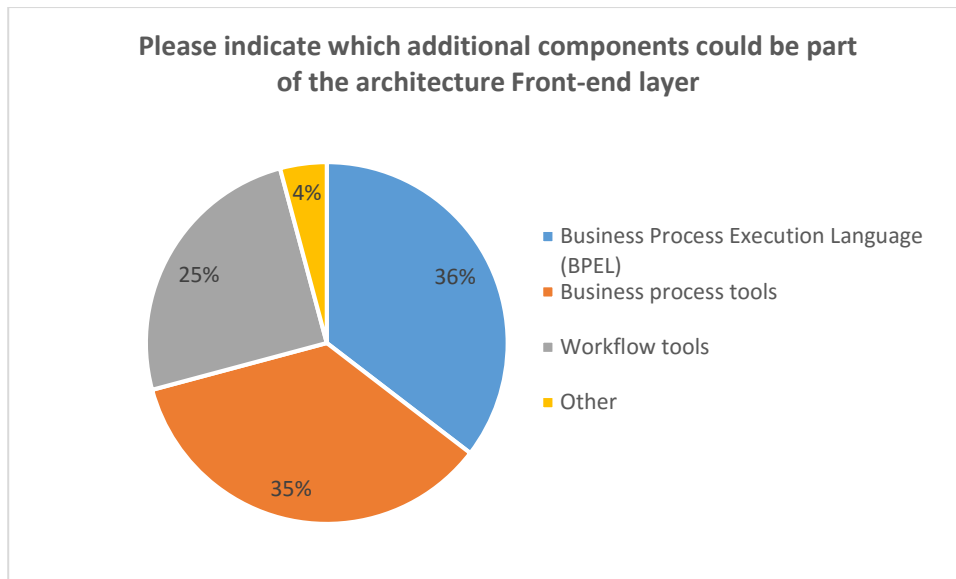


Figure 8-8: Front-end Layer: Additional Components

8.3.5 Section 4: Service layer

The objective of the questions in this section was to evaluate the BI services for validation or improvement of the proposed services layer (Layer 3) in the hybrid BI architecture. A Service layer in a BI architecture comprises the analytic engines that are responsible for executing analytical functions requested by business users in order to achieve and support better decision-making processes.

Respondents were asked to assess certain statements in regard to the architectural service layer of a BI solution. Their responses to the questions are presented and interpreted below.

8.3.5.1 Assessment of BI architectural service layer

As shown in figure 8-9, most (90%) of the respondents agreed that a BI architecture must have a service layer and 87% said it must be a dedicated layer. The majority (84%) of the respondents suggested that the service layer should have an Enterprise Service Bus (ESB) or Service-oriented Architecture (SOA). On average 76% indicated that the service layer should consist of the ESB, BI services, PaaS, DBMS and application server. These results suggest that a number of BI-related systems, servers and services can form part of the service layer in a BI architecture.

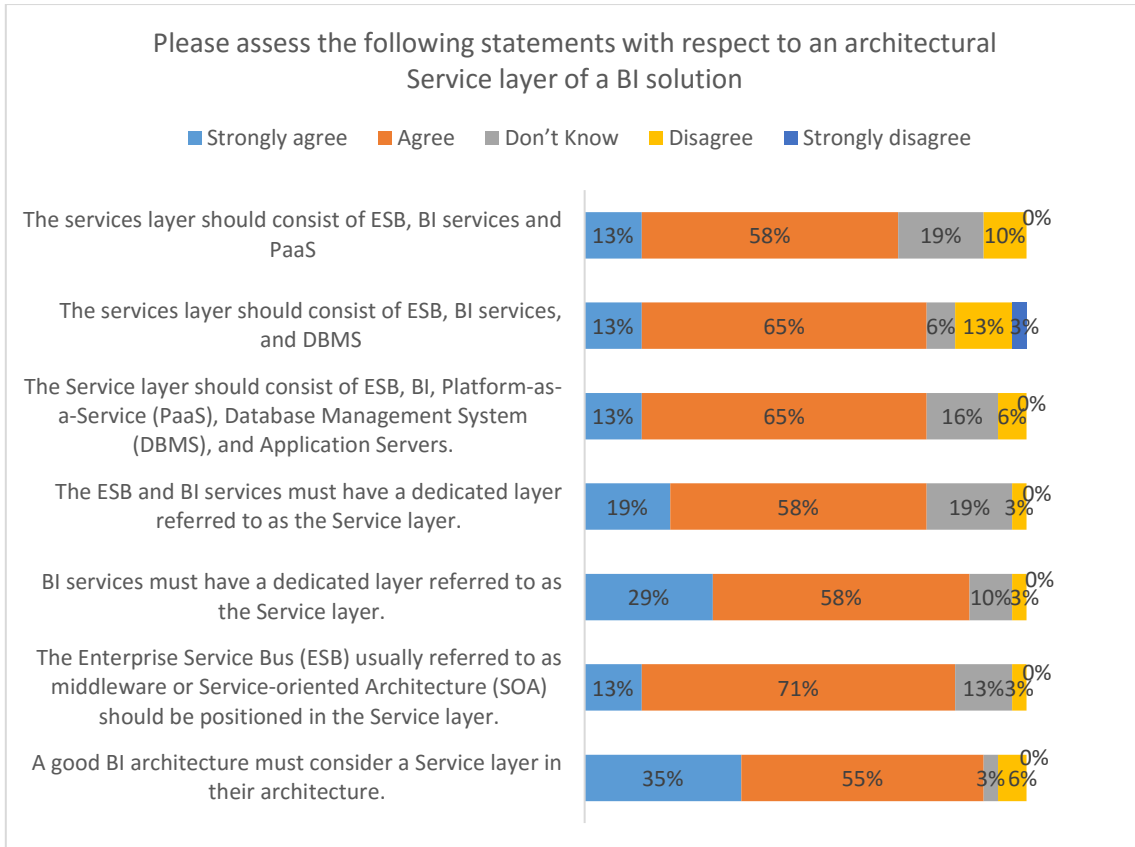


Figure 8-9: Service Layer – Proposed Components

8.3.5.2 Additional components for a service layer

As shown in figure 8-10, 19% of the respondents indicated that visualisation MapReduce services can form part of the Service layer. More than one-third (37%) indicated user applications (including in-memory applications), 10% indicated User and, or Machine interface, 13% indicated Model-View-Controller, 19% indicated data collectors, and 2% indicated Other (“workflow tools”) as ideal components for the service layer. These results also suggest that the service layer can be structured according to the requirements of the organisation.

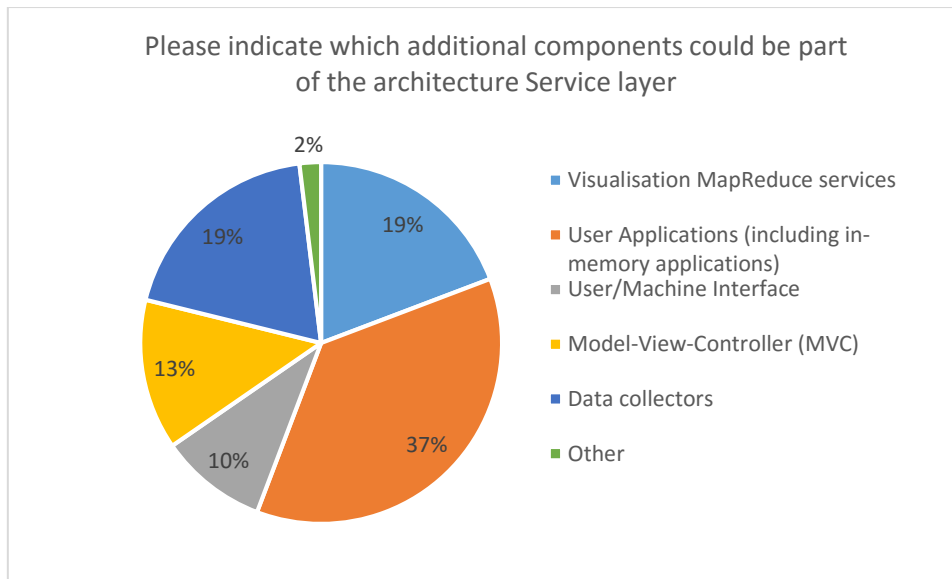


Figure 8-10: Service Layer – Additional Components

8.3.6 Section 5: Back-end layer

The objective of the questions in this section was to evaluate the BI tools for validation or improvement of the proposed presentation layer (layer 2) in the hybrid BI architecture. A back-end layer in a BI architecture is a server side where all the applications are executed in the background without the client's or user's knowledge. Respondents were asked to assess certain statements in respect of the architectural back-end layer of a BI solution. The following are presentations and interpretations of how the respondents responded to the statements.

8.3.6.1 Assessment of BI architectural back-end layer

As shown in figure 8-11, 87% of respondents agreed that a back-end layer should form part of the BI architecture. The majority (81%) indicated that PaaS should be part of the Back-end layer, and less than one-half (42%) of the respondents did not agree that it should have it in its own layer. The majority (81%) indicated that a combination of PaaS, DBMS and application servers should form part of the back-end layer. On the other hand, 45% indicated that DBMS should be a dedicated layer whereas 45% did not agree, leaving 10% who indicated that they 'don't know'. Over one-third (68%) indicated that the application layer should be a dedicated layer. Last but not least, 84% indicated that data source systems (e.g. legacy data systems, OLTP, etc.) should be part of the back-end layer. These results suggest that the recommended back-end

layer applications and systems can be adopted in various ways depending on the expectations from the BI systems.

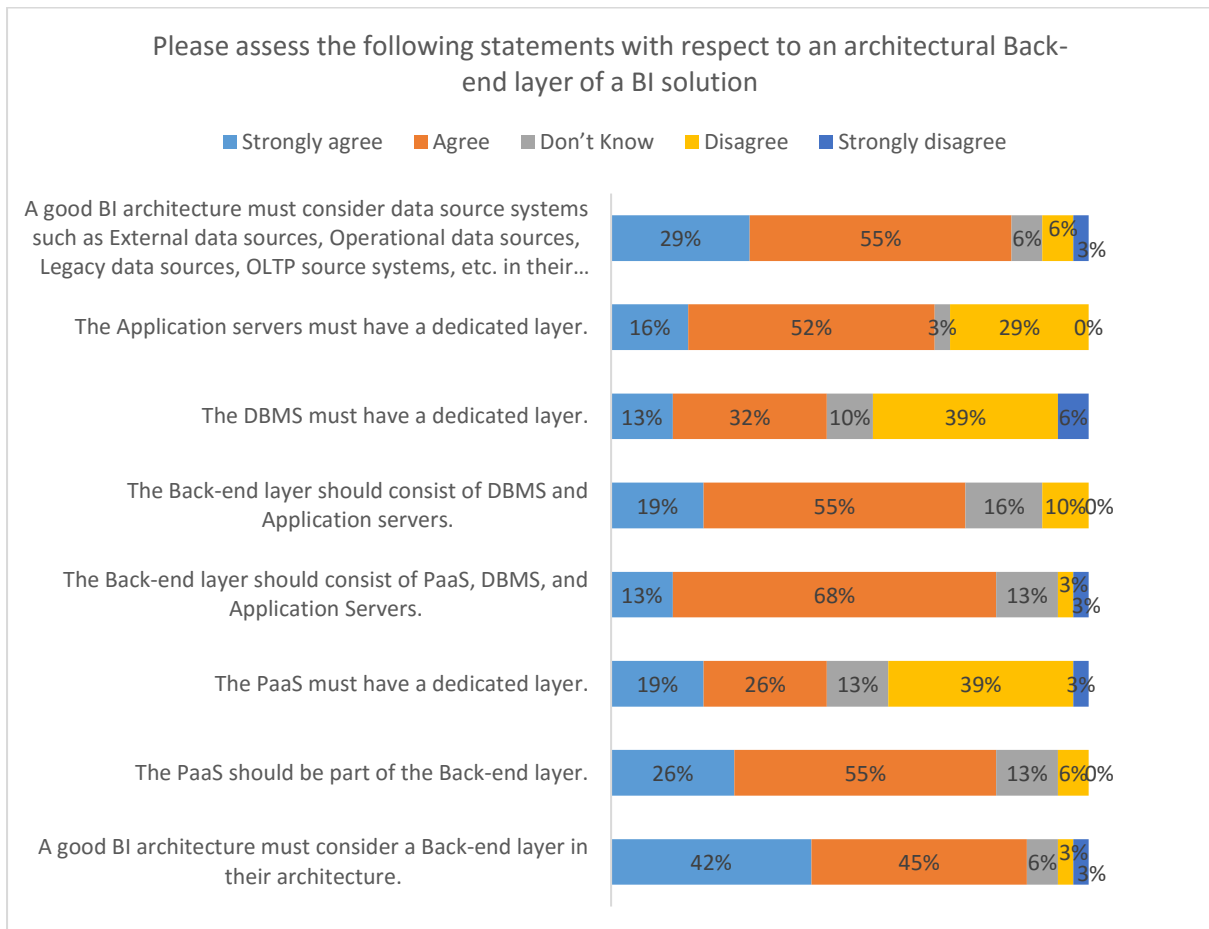


Figure 8-11: Back-end Layer – Proposed Components

8.3.6.2 Additional components of a back-end layer

Respondents were asked to indicate which additional components could be part of the back-end layer of the BI architecture. As shown in figure 8-12, 22% indicated development tools, 16% indicated middleware and In-house API respectively, 18% indicated execution runtime environment, 9% indicated third party API, 11% indicated application server, and 8% web server. These results suggest that most applications or servers can be incorporated into the back-end layer in order to meet certain business requirements.

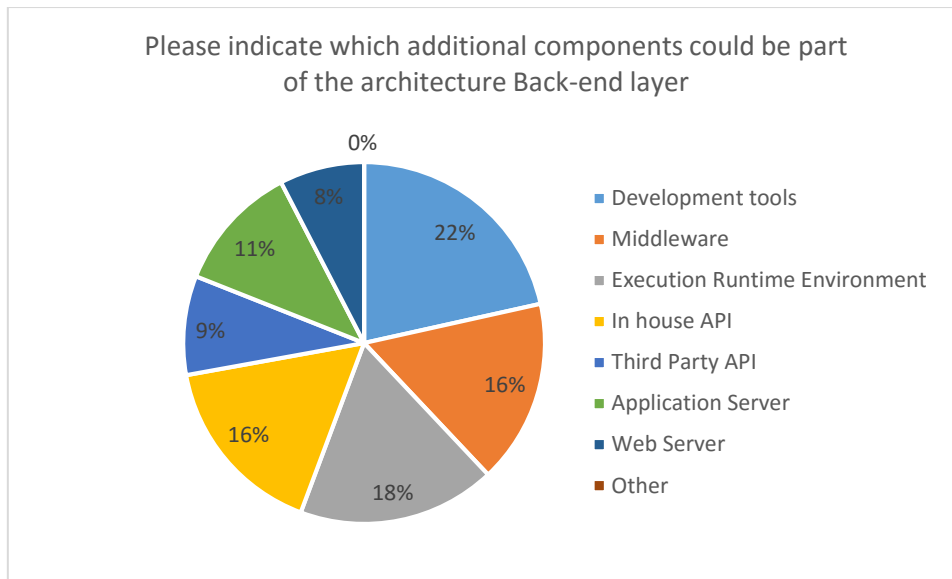


Figure 8-12: Back-end Layer – Additional Components

8.3.7 Section 6: Infrastructure layer

The objective of the questions in this section was to evaluate the infrastructure resources for validation or improvement of the proposed infrastructure layer (layer 5) in the hybrid BI architecture. An Infrastructure layer in a BI architecture comprises BI resources related to hardware and data management systems. The following are presentations and interpretations of how the respondents responded to the questions.

8.3.7.1 Assessment of BI architectural infrastructure layer

Respondents were asked to assess statements in respect of the proposed components for an architectural infrastructure layer of a BI solution. As shown in figure 8-13, most (97%) of the respondents indicated that an infrastructure layer should be part of the BI architecture. The majority of the respondents indicated that hardware resources (87%) and database technologies (90%) should be part of the infrastructure layer. Eighty-one percent (81%) of the respondents indicated that the data management layer should have a dedicated layer. Furthermore, 48% indicated that the network layer should not be on its own, but rather be part of the infrastructure layer as indicated by 78% of the respondents. These results suggest that the infrastructure layer in a BI architecture should consist of at least hardware resources, database technology resources, and network resources.

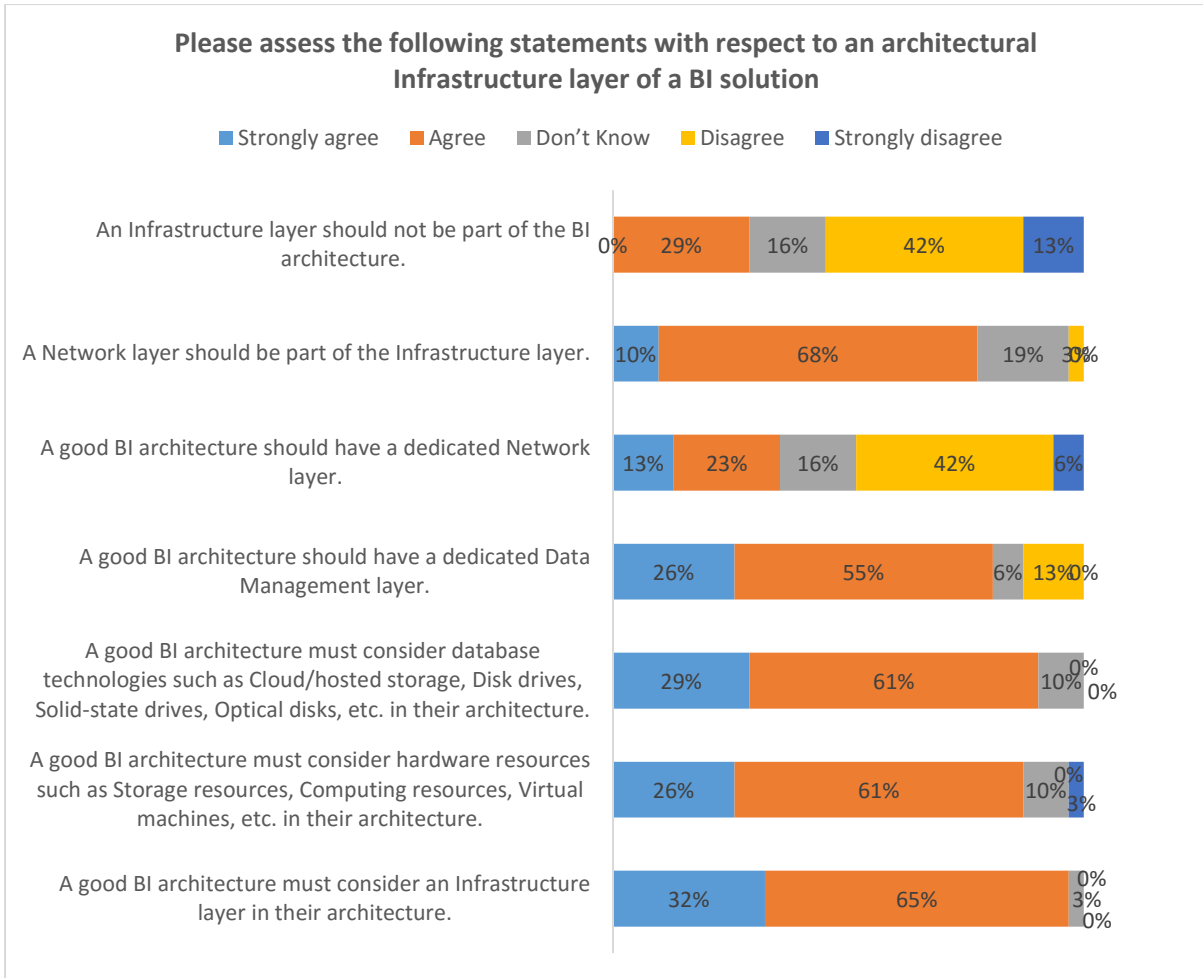


Figure 8-13: Infrastructure Layer – Proposed Components

8.3.7.2 Additional components for an infrastructure layer

Respondents were asked to indicate which additional components could be part of the architecture infrastructure layer. Figure 4-1 shows that 41% of the respondents indicated that IaaS should be part of the infrastructure layer, followed by 28% who indicated Virtual Private Network, 14% who indicated Access Point Name, and 17% who indicated a combination of routers and switches. The results suggest that the proposed additional infrastructure resources could be considered when certain organisational objectives have to be met.

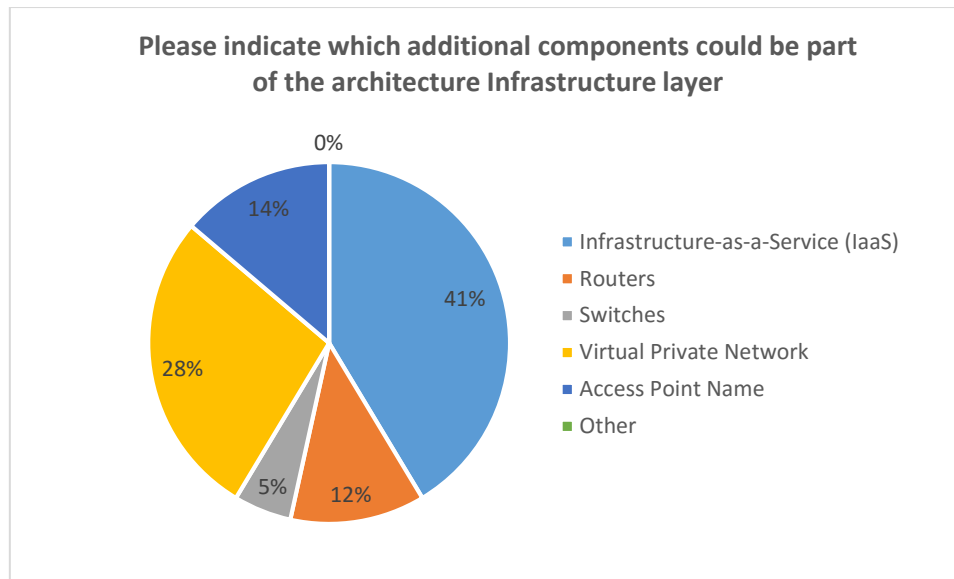


Figure 8-14: Infrastructure Resources

8.4 Research findings

The following is a summary of the key findings related to each sub-objective of the study survey which attempted to validate the proposed hybrid BI architecture by questioning experts in order to suggest improvements in organisations in the South African context.

8.4.1 Degree of BI knowledge and ideal procurement options

In order to determine the degree of BI knowledge and ideal procurement options in organisations, the respondents were asked to assess their expertise and knowledge of their BI solution, and the key findings are listed below.

- It was discovered that on-premise and, or cloud-based BI solutions are widely used in the organisations surveyed. On-premise deployment of BI solutions was found to have BI resources, such as, hardware and software applications and systems housed inside the organisations as well as ensured ownership of business data. It was also found that organisations with cloud-based deployments used standard cloud service models, such as, SaaS, PaaS and IaaS that have the capability of providing quicker and affordable BI services.
- A hybrid BI solution was also found to be popular as it promotes the attainment of the benefits, such as, effective management of sensitive information from

both on-premise and cloud-based deployments. It was learnt that these BI solutions are implemented in such a way that they ensure critical information resources are secure, manageable, and easily accessible.

- The ideal way of procuring BI solutions was found to be through a suitable pricing model that met the organisation's needs and not through a once-off purchase. Of the various models presented, subscription-based pricing was the most favourable followed by pay-per-user. Pay-as-you-go and pay-per-use options are also likely to be considered first before looking at a server-based option.

In summary, BI solution deployment was found to be preferred through a hybrid deployment as it can accommodate both the benefits of an on-premise and cloud-based BI solution. It is believed that the benefits of a hybrid BI deployment outweighs the individual deployment (Srinivasan 2011:344). Furthermore, it was found that an organisation will realise value from reduced IT cost through the use of BI services that are supported by the standard cloud service models. This was found to be possible if an ideal pricing model was selected that meets the organisation's requirements.

8.4.2 Validation of the front-end layer devices and applications

In order to evaluate the front-end layer's devices and applications of the BI architecture, the respondents were asked to assess components ideal for the layer, and the following were the key findings:

- It found that a good design for a BI architecture must consider having a front-end layer. An ideal front-end layer must consider having components, such as, an interface layer, presentation layer, BI applications and tools, devices, Business Process Execution Language, business process tools, and workflow tools.
- Although it was found that devices, applications and interface layers were the most favourable, a combination of the aforementioned components appeared to be critical in facilitating the ideal functionality of the front-end BI solution.

In summary, the findings suggest that the proposed presentation layer and application layer of the architecture need to be improved by possibly merging the two layers in conjunction with other components. A robust front-end layer consisting of the right BI applications and tools and utilising the appropriate technological devices will ensure that organisations are able to analyse business data and generate informative reports or interactive dashboards (Russom 2013:34).

8.4.3 Validation of the BI service layer

In order to evaluate the BI service layer of the BI architecture, the respondents were asked to assess services ideal for the layer, and the key findings are listed below.

- The service layer was found to be an ideal layer for a BI architecture and needed to be a dedicated layer. It was learnt that the ideal components of the Service layer comprised the ESB/SOA, BI services, PaaS, DBMS, application server, Visualisation MapReduce services, user applications (including in-memory applications), user and, or machine interface, Model-View-Controller and data collectors.
- It was learnt that the inclusion of the ESB/SOA was vital for the facilitation of the integration process of a variety of the BI services, applications and systems in an organisation.

In summary, these findings are consistent with the proposed components for the service layer, but the findings suggest that an enhanced layer needs to be proposed with additional components. A service layer that has the capability of hosting on-premise and cloud-based BI solutions enables organisations to use state of the art technology that is accessible anywhere in the world (Rahman et al 2013:542).

8.4.4 Validation of the back-end layer

In order to evaluate the BI back-end layer of the BI architecture, the respondents were asked to assess applications or systems ideal for the layer, and the key findings are listed below.

- The findings show that the back-end layer should be part of the BI architecture as a dedicated layer. It was found that components, such as, PaaS, DBMS and application servers, data source systems, development tools, Middleware, In-house API, Execution Runtime Environment, third party API, application and web servers are ideal for this layer.
- It was also found that the proposal of having application servers, DBMS and PaaS as dedicated layers was favourable, but not as favourable as having them reside within the back-end layer.
- It was learnt that a variety of data source systems, such as, external data sources, operational data sources, Legacy data sources, OLTP systems, etc. are ideal sources for data needed for BI solutions.

In summary, the findings suggest that the current proposed data layer and service layer's components need to be revised to accommodate the suggestions for the back-end layer. A back-end layer is responsible for managing multiple data sources that can be hosted on-premise or in the cloud depending on the level of confidentiality of the information to be processed and stored (Ong et al 2011:4; Zhang & Tu 2009:182).

8.4.5 Validation of the infrastructure layer

In order to evaluate the infrastructure layer of the BI architecture, the respondents were asked to assess hardware resources and data management systems ideal for the layer and the key findings are listed below.

- Both the infrastructure layer and data management layer were found to be ideal dedicated components of a BI architecture. However, the infrastructure layer was the more favoured upper hierarchical title for the layer than the data management layer.

- It was found that the ideal components for the infrastructure layer are hardware resources, database technologies, Network layer, IaaS, Virtual Private Network, Access Point Name, Routers and Switches.

In summary, the findings were similar to those in respect of the initially proposed layer, but additional components were identified as ideal which prompted making adjustments to the current layer. The infrastructure layer has the capability to host BI resources on-premise or in the cloud as IaaS, giving the organisations the benefit of using unlimited storage services for managing data storage challenges (Mvelase et al 2012:215).

Finally, having met the objective of evaluating and validating the layers and components of the initially proposed BI architecture for improvement purposes, it was concluded that the proposed architecture had to be fine-tuned to accommodate the input and opinions received from the BI experts' responses. Therefore, the next section presents an improved hybrid BI architecture based on the literature review and the findings from the study.

8.5 Improved BI architecture

This section presents an improved version of the BI architecture. It was very important that the contributions from the survey respondents should be taken into account as the responses served the purpose of evaluating and validating the proposed architecture. Figure 8-15 demonstrates the enhanced BI architecture followed by a discussion about the changes and vital additional components emanating from the research findings.

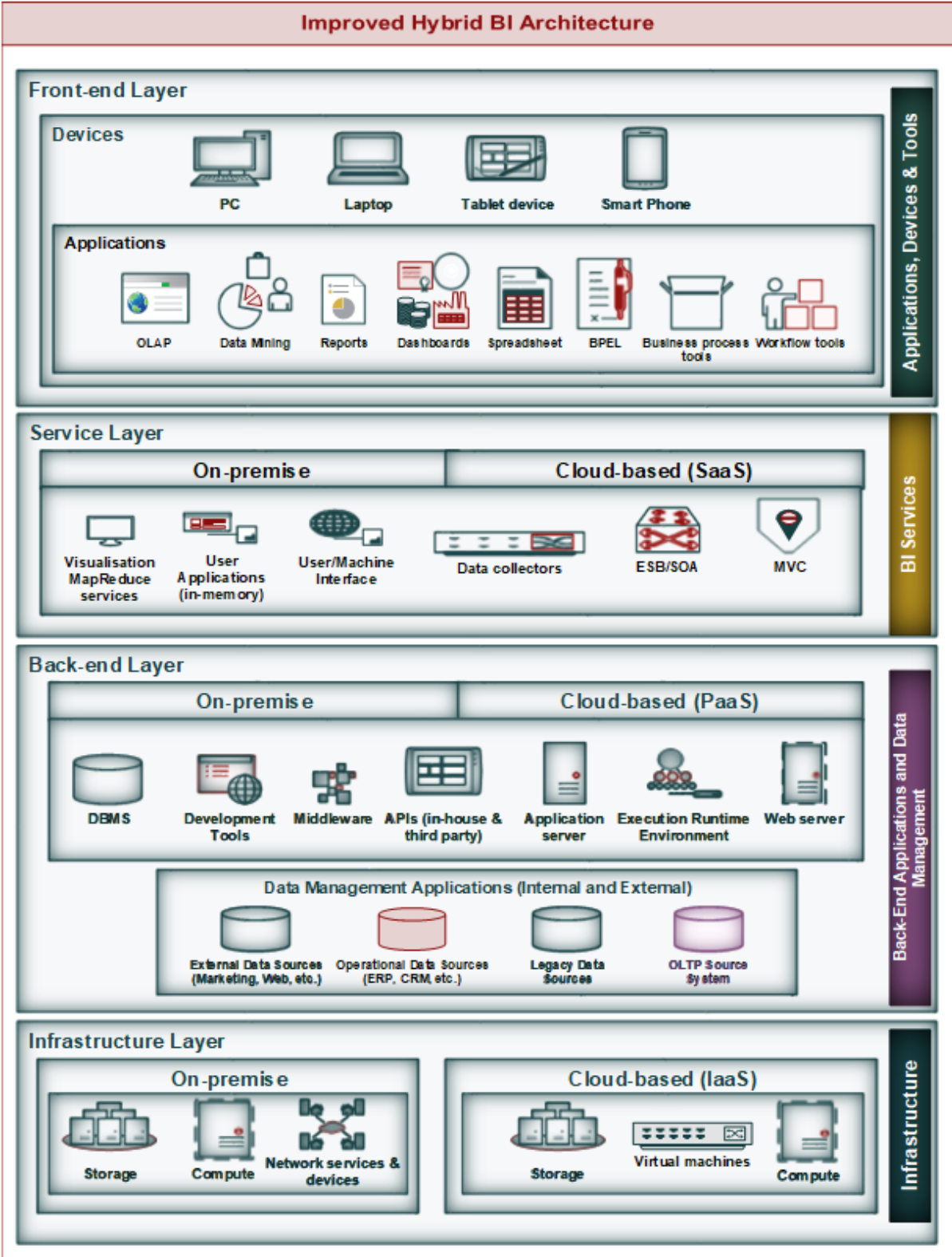


Figure 8-15: Improved Hybrid BI Architecture

8.5.1 Front-end layer

In the initial architecture, the presentation and application layers were independent layers. Based on the recommendations emanating from the results, these layers have been merged together to form what is now known as the front-end layer. The front-end layer will manage access to an on-premise or cloud-based portal for BI solutions through the thin clients that are accessible on any available technological devices, such as, laptops, tablets, smartphones, computers, etc. (Ali et al 2015:360; Garrison et al 2012:63). Upon gaining access, business users can use additional applications, such as, Business Process Executing Language (BPEL), business process and workflow tools to manage business processes in the organisation. BPEL is the de-facto standard workflow language for executable business processes and is suitable for graph representation of activities (Hertis & Juric 2014:740; Distefano, Ghezzi, Guinea, & Mirandola 2014:690). Furthermore, business users can use specific BI functionalities in their applications to perform any required BI-related tasks in order to meet specific business requirements.

8.5.2 Service layer

The service layer maintains the same name, but the components have slightly changed. There is the introduction of a newly recommended component called the Enterprise Service Bus (ESB) that supports Service-oriented Architecture (SOA). ESB is a service-oriented platform or infrastructure for inter-connecting services that can be used to connect several BI applications that have been developed based on various technologies, incompatible formats, data resources, etc. (Orłowski et al 2014:28; Schmidt, Hutchison, Lambros & Phippen 2005:783). ESB and SOA ensure that the technical part of the BI services that are provided remain invisible to the business users requesting the services, and the BI applications providing the service. The ESB and SOA can be utilised to mediate the BI services that are required between on-premise and cloud-based environments.

8.5.3 Back-end layer

The back-end layer is a new layer that has emanated from the combination of data layer and some components from the service layer. BI experts recommended that on-

premise components, that is, development tools, Middleware, DBMS, APIs, etc. and similar PaaS components must be part of the back-end layer. Data management applications (in the previous data layer) are now components of the back-end layer. This ensures that everything happening behind the scene in the on-premise environment remains invisible, as is always the case with cloud-based BI solutions (Ong et al 2011:4; Armbrust et al 2010:55). For example, the data source systems that manage highly sensitive information can be stored on-premise or in the cloud depending on the level of sensitivity.

8.5.4 Infrastructure layer

The infrastructure layer also maintains the same name, but with a few additional components related to network services and devices, that is, Virtual Private Network (VPN), Access Point Name, Routers and Switches. The Network services and devices ensure that data and information are transferable from the source systems into the analytic engine for reporting purposes (SAP 2012:5). For example, VPN can be used to integrate on-premise BI applications with cloud-based applications (Liu et al 2010:403), while the business user maintains full control of the Access Point Name, Routers and Switches residing on-premise. It is critical that the IaaS and Network services and devices ensure data is always safe, secure and readily available to authorised business users in any BI solution implementation (MicroStrategy 2012:141).

In summary, the BI architecture has been improved by introducing a few critical components and moving some of them around as suggested by the findings from the BI experts' responses. This improved BI architecture has been recommended to be used as a guide when implementing a hybrid BI solution based on the business objectives which the organisation needs to meet. In the next section, an example of a practical implementation of the hybrid BI architecture is demonstrated.

8.6 Practical demonstration of the proposed hybrid BI architecture

This section provides an example of how an organisation can use the improved hybrid BI architecture as a guide when implementing a BI solution to resolve real-life

challenges. For the purpose of demonstrating the proposed hybrid BI architecture, an organisation that provides various consulting and recruitment services was chosen and will be referred to as 'Nolex' in order to maintain anonymity.

Nolex needs to implement a BI solution for managing its business that generates multiple data from various divisions. Besides data emanating from cloud-based CRM systems (SalesForce.com, Placement Partner, etc.) and other sales and transaction activities, the organisation has a website and several social media web pages, and, or sites. Nolex also uses a local server with high maintenance costs and a home-grown system to store copies of various documents, such as' consent and application forms, completed and submitted tender documents, SARS documents, and copies of invoices and timesheets, which are kept for compliance purposes or have been submitted to clients and suppliers. A record of such documentation is critical for reporting purposes and future reference should a query arise in regard to a particular submission. Therefore, Nolex faces the challenge of managing big data and successfully reporting on it. To a greater extent than is usually the case a portion of this data needed for further analysis and reporting is often extracted from the CRM solutions and must be stored and managed manually on the premises. To implement a suitable BI solution Nolex can use the improved BI architecture as a guide. An overview of how this could be done is provided below.

It would be ideal for Nolex to consider a hybrid BI solution to address the above-mentioned issues. A hybrid BI solution can be procured at a cheaper price when the various pricing models, such as, subscription-based, pay-as-you-go, pay-per-user, etc. are compared. Selecting any one of these pricing models will mean immediate access to the BI solution at a low, or no investment costs upfront, and with maintenance costs eliminated. In order to ensure success in this hybrid BI deployment, the following scenario is plausible or should be considered.

First, Nolex needs to consider the front-end applications and devices. The organisation can choose between proprietary and, or open source solutions necessary to meet their requirements. The organisation can continue using SalesForce.com – a CRM application to track sales accounts and pipelines and obtain information from consulting services, to benefit from business data stored and managed in the cloud. As an alternative to SalesForce.com, Nolex can pilot test EspoCRM or Zoho which are

open source web-based applications, also available at subscription pricing with similar features, but at a lower cost. Nolex can again continue using Placement Partner for storing and managing confidential information coming from the recruitment business, such as, candidate CV details, placements, gross profit information, placement fees, client and, or candidate interactions, finger print scans, copies of qualifications and IDs, client details, etc. However, as far as storing certain sensitive or non-sensitive but very important information is concerned, various proposed data management options can be considered which could include keeping information on-premise in a secured local server for easier access. This information can include SARS documents, consent forms for criminal and financial checks, application forms, tender documents, invoices and timesheets.

Nolex can install and use a business process tool on their existing laptops or computers to document all functions, processes and activities required to use a BI solution. The relevant business users can then be informed about their responsibilities and this can be linked to specific BI applications for data mining and reporting, such as, Pentaho, an open source application. This can alleviate the challenges associated with generating different types of reports on a monthly basis and the manual intervention to consolidate the reports. By simply selecting the above-mentioned applications and devices, Nolex will be able to use data integration services (to be discussed later) that will consolidate data coming from the CRM applications and store it in a central data repository.

Second, the service layer has to be considered, since the databases and reporting functions that come with Salesforce.com or Placement Partner systems alongside the home-grown system's information are not integrated. A cheap and easy to install ESB/SOA or Middleware, such as, Unified Platform of Talend, an open source-based and freely available ESB that can be used independently, or in combination with other components, or Mule ESB which has similar qualities, such as, one click installation can be used to integrate the various systems.

A hybrid BI solution is an ideal solution to manage this particular organisation's business operations as it offers upgrades to new technology at no additional cost, unlimited storage, and highly scalable applications, etc. Since the organisation cannot do away with storing certain documentation on-premise, a less costly open source

data warehousing solution, such as, Hadoop running Teradata can be used to store all information emanating from the activities done on premise, website and social media sites. Some of the front-end systems can still remain with proprietary software, such as, Windows and Microsoft Office applications which makes it easier to integrate with the data warehouse.

Third, Nolex should consider back-end layer components. Data integration and administration software are not problematic to Teradata as it comes with industry-standard interfaces. The interface can integrate the Teradata system with the rest of the systems in an organisation in order to streamline all ETL processes. Since data originates from multiple sources, utilisation of ETL tools, such as, DataFlux or DataCleaner will ensure that data is transformed into the format that is fit for its purpose. An open source business analytics solution, such as, Pentaho can be used to meet all big data integration, business analytics and BI needs for the organisation. Pentaho, a cloud-based solution, comes with a local client application for use on the premises. It is an ideal BI solution for reporting, visualising and having predictive analytics on Nolex's big data. Pentaho can seamlessly integrate with Teradata systems in order to produce real-time interactive, self-service and visualisation reports and dashboards at reduced operational cost. All these options can be explored when the hybrid BI architecture is carefully used as a guide when selecting the appropriate applications.

Finally, the infrastructure layer has to be considered, since the organisation is required to keep certain information on-premise and the other in the cloud, a robust network infrastructure is crucial in this setup. The Infrastructure layer components can assist Nolex evaluate other infrastructure resources, such as, considering IaaS which offers cheaper unlimited storage, speedy computing resources, and disaster recovery at low cost, etc. An on-premise local network can be setup in such a way that it has access to IaaS services through a VPN so that the organisation's data which comes from various sources can be easily integrated with the various systems and applications mentioned earlier.

The consideration of using the proposed improved hybrid BI architecture will enable Nolex to acquire a clear understanding on some key aspects. These aspects include, guaranteed ownership of business data through clear SLAs; balance of workload

ensuring appropriate applications run either on-premise or in the cloud; use of a parallel processing mechanism to manage big data and to make real-time analytical processing of data easier; access to specialised and adaptive cloud data management and analytic applications which are available immediately with no burden for maintenance cost; and so forth.

Besides the BI architectural considerations, it would be wise to consider TCO associated with such an important and long term BI initiative. TCO was introduced to demonstrate the types of costs that can be associated with each deployment. Considering the various discussions on TCO related to naïve, on-premise, cloud-based and hybrid models, the following formulas have been suggested for cross-comparison so that the organisation can keep the TCO low.

Table 8-1: TCO Calculation for BI Deployment

Direct Costs (DC) – Initial deployment costs	Abbr.
• Cost of hardware (e.g. new servers)	DC_{coh}
• Cost of middleware (operating system, database servers, Web servers)	DC_{com}
• Cost of software licenses	DC_{cosl}
• Cost of consulting services for implementation	DC_{cocs}
• Cost of staff hours to implement the project	DC_{coshr}
• Cost of software quality	DC_{cosq}
• Costs of changes needed in the network	DC_{cocn}
• Cost of additional storage	DC_{soas}
Where: N is the number of user/licenses, HR is the total number of hours spent by staff. $DC = DC_{coh} + DC_{com} + (N)DC_{cosl} + (HR)DC_{cocs} + DC_{coshr} + DC_{cosq} + DC_{cocn} + DC_{soas}$	
Indirect Cost (IC) – Ongoing management costs	
• Cost of upgrading related software	IC_{cous}
• Cost of upgrading related hardware/middleware	IC_{couh_um}
• Cost of customisation/continued application evolution (e.g. new end-user reports/dashboard)	IC_{c_cust}

• Impact on use of existing servers	IC_{i_serv}
• Impact on existing network traffic	IC_{i_net}
• Impact on existing storage	IC_{i_stor}
• Increase in support calls to IT	IC_{i_sup}
• Increase in maintenance of hardware/middleware	IC_{i_maint}
• Downtime cost (loss on productivity costs)	IC_{dt_cost}
<p>Where: N is the number of user licenses/users</p> $RC = IC_{cous} + IC_{cous_um} + (N)IC_{c_cust} + IC_{i_serv} + IC_{i_net} + IC_{i_stor} + IC_{i_sup} + IC_{i_maint} + IC_{dt_cost}$	
Related Costs (RC)	
• Staffing costs	RC_{st_cost}
• Training end-users	RC_{train_EU}
• Training BI professionals, e.g. business analysts, data scientists	RC_{train_BI}
• Training IT staff	RC_{train_IT}
• Creating governance rules	RC_{gov_rules}
• Implementing governance rules	RC_{impl_rules}
• Maintaining governance rules	RC_{maint_rules}
<p>Where: N is the number of user licenses/users</p> $RC = RC_{st_cost} + (N)RC_{train_EU} + (N)RC_{train_BI} + (N)RC_{train_IT} + RC_{gov_rules} + RC_{impl_rules} + RC_{maint_rules}$	
<p>Therefore, the estimated TCO can be achieved by combining the above costs as follows: $TCO = DC + RC + RC$</p>	

8.7 Conclusion

The purpose of this chapter was to present and discuss the findings of the survey conducted on validating the researcher's proposed BI architecture by BI experts within

specific industries in South Africa. The results of the validation of the hybrid BI architecture were quantitatively and qualitatively presented and discussed in detail. It was concluded that the initial proposed layers, namely, the presentation, application, service, data and infrastructure layers of the architecture needed to be enhanced based on the contributions from the BI experts. This resulted in the merging and renaming of some of the current layers into the following: front-end layer, service layer, back-end layer, and infrastructure layer. The improved architecture was discussed in detail mainly highlighting the new components. This was followed by a practical demonstration of how a hybrid BI solution could be deployed together with the associated benefits and TCO calculations based on a real-life case study.

In this chapter the theoretical initial hybrid BI architecture was validated with real-life contributions, and thereafter, improvements based on the empirical findings were suggested. A hybrid BI solution is the recommended solution for the future to address various BI challenges faced by organisations. Therefore, the next chapter seeks to conclude the research and provide some recommendations and closing remarks.

CHAPTER 9 SUMMARY AND CONCLUSION

9.1 Introduction

The previous chapter discussed in detail the data analysis of the last part of the survey, and validated the theoretical initial hybrid BI architecture with real-life contributions. It also presented an improved architecture based on the findings. The purpose of this chapter is to summarise and conclude this research study conducted on a data management and analytic model for BI applications in the South African context.

This chapter summarises the research findings, including findings from the case studies in section 9.2, explains how the objectives of the study were achieved in section 9.3 and discusses possible future research in section 9.4. Finally, conclusions based on the research findings are drawn and recommendations made in section 9.5.

9.2 Research findings

After laying the foundation of the research in the introductory chapter, the key findings, including the views obtained from the survey conducted on BI users and BI experts are summarised and discussed below.

- BI was found to be the process of analysing business data stored by an organisation to determine patterns and trends that support better decision-making for the organisation. In order to achieve these results, an organisation needs an architecture that it could use for the implementation of a BI solution. It was found that a sound technical implementation of a lean BI solution should be based on a BI architecture which is informed by a BI value chain and a naïve model.
- The ideal BI architecture was found to consist of interrelated layers that would enable effective sourcing of data, storing of data and reporting. It was also found that a basic BI solution implementation, OLTP and OLAP systems are critical in managing and analysing data. In order to understand cost implications of implementing a BI solution, TCO was found to be useful by considering direct costs, indirect costs and related costs. In the initial investigation of a naïve model, TCO was found to be near zero.

- It was discovered that as far as deploying a BI solution is concerned, organisations are faced with the challenge of selecting the appropriate deployment model. An on-premise BI deployment model emerged as one of the methods that an organisation could use which entailed that all related software, hardware and application systems are deployed inside the organisation's building. It was found that with this approach organisations were constantly looking for other ways of improving the provision of BI solutions as it was deployed mainly in three modes, namely, the isolated, overlapping and/or centralised or shared modes.
- A BI architecture for an on-premise deployment was found to be more advanced than that of a naïve setup as it had to consider the organisation's structure, departmental needs, and size. Although on-premise BI deployment solutions were able to support the decision-making process, they still had a few challenges relating to managing data and analysing it for reporting purposes. Most TCO elements were applicable which made the implementation a costly exercise.
- Another BI deployment model found to be commonly used was a cloud-based model that mainly utilised cloud computing services, namely, SaaS, PaaS and IaaS. Cloud-based BI services were found to be mainly offered through private, public, hybrid or community cloud deployment modes. These offerings ensure that minimum or no upfront cost is incurred and provides the required services immediately through a subscription pricing method. However, data management issues persist as data resides on-premise while BI services are in the cloud.
- For cloud offerings to be successful in an organisation, a cloud-based BI architecture was found to be a bit more advanced than the on-premise one as it has to address security concerns related to Internet and its cloud computing services. It was found to be also offering other independent architectures, such as, SaaS architecture, SaaS and PaaS architecture, and SaaS and IaaS architecture that can facilitate easy access to cloud BI services. As far as TCO is concerned, it was discovered that cost elements are dependent on the size of the organisation, and hence most cost elements are not applicable.

- In order to address the persistent challenges organisations face in on-premise and cloud-based BI solutions, a hybrid BI solution was proposed. A hybrid BI solution was found to involve a number of fundamental elements that have an influence on the performance of BI solutions. These elements include in-memory data management, in-memory analytical and transactional processing, big data management, and hybrid data management.
- By considering all the on-premise and cloud-based BI solutions along with hybrid fundamental elements, a hybrid BI architecture was found to be more complex than any other architecture that emerged in this study, and was able to manage the persistent challenges. This was achieved through handling emerging issues in their respective domain, that is, on-premise and in the cloud, respectively. This means that organisations that have issues with managing sensitive business data in the cloud are able to do so on-premise and still optimally use cloud-based resources, that is, SaaS, PaaS and IaaS BI computing resources for other business requirements.
- It was discovered that the cornerstone of using a hybrid BI solution was based on the ability of using converged on-premise BI resources and a combination of private cloud and public cloud BI resources that were able to scale up or down in order to meet organisational needs. TCO was found to be dependent on both factors associated with on-premise and cloud-based BI deployments, and the nature and geographical spread of the organisation determined the cost implications by either depending on more cloud adoption and less on-premise, or the other way around. Although the hybrid BI architecture managed to address most of the challenges, it still needed to be validated by BI experts.
- It was found that the key to assessing the state of BI utilisation in organisations and the success to validating the hybrid BI architecture is through a sound data collection methodology. A case study approach with appropriate research questionnaires was conducted to collect the required research data. It was found that the pilot study and pre-testing of the questionnaires played a crucial role in eliminating the weakness of the questionnaires and ensuring that the objectives of the questionnaires were met.
- After conducting the research survey, in terms of assessing the utilisation of BI in organisations by BI users, it was found that a number of industries used well-

structured or advanced BI solutions as compared to a naïve BI solution. It was discovered that a fully-fledged BI solution is the key to successful delivery of insightful information that enables a better decision-making process as most organisations use on-premise and, or cloud-based BI solutions. A range of transactional systems, DBMS systems and ETL tools are utilised to effectively manage and ensure the quality of data for the organisation. Furthermore, the success of utilising the opted BI solutions was based on selecting the appropriate procurement option as various purchasing and licensing options are available that would meet specific business requirements.

- On the other hand, in terms of validating the components of the proposed hybrid BI architecture by BI experts, it was found that a hybrid BI deployment was preferred as the benefits outweigh the individual deployment and is highly dependent on an effective and all-inclusive BI architecture as proposed. All-inclusive hybrid BI architecture means that the front-end layer must consist of technological devices and BI applications or tools that meet the organisation's BI needs; the service layer is important in providing comprehensive on-premise and, or cloud-based BI related services; the back-end layer was found to be appropriate for managing internal and external DBMS systems including other on-premise and cloud-based BI resources; and the infrastructure layer was found to be critical in hosting related BI resources on-premise, and unlimited storage and computing BI resources in the cloud.

9.3 Research summary

This section describes the way in which the main objective and sub-objectives of this research study were achieved, as well as how the sub-research questions were answered. -The main objective was:

- ***To develop a data management and analytic model that can serve as a guide for BI applications implementation.***

The main research question derived from the main objective was:

- ***What is an appropriate data management and analytic model that can serve as a blueprint for BI applications deployment for organisations?***

This study's main objective was achieved through four sub-objectives. The subsections below summarise the way in which the listed sub-objectives were addressed in various chapters in order to achieve the main objective and answer the main research question, respectively.

9.3.1 Sub-objective 1: To investigate and understand BI in General

The objective of understanding BI in general was achieved in Chapter 2. This chapter's purpose was to provide a general understanding of BI, its use and the associated business value. It covered in detail the key concepts and terminologies that are commonly used in order to understand and implement BI successfully. BI was defined as an optimal and effective collection of business data from various sources, which is then transformed through intelligent applications in order to understand information, aimed at supporting decision-making that helps organisations achieve their objectives. This definition contributed to an understanding of the main features and functions that enable utilisation of BI as a business management tool that consists of the applications and advanced technologies required for gathering and analysing information in order to support decision-making and realise business value.

A BI value chain was developed to demonstrate the value generated through the use of BI. This value chain was supported with a naïve BI model. A naïve model formed the basis for the technical implementation of a lean BI solution with the intention of demonstrating how to implement a BI solution based on the value chain. The model covered key aspects of BI, such as, data sources systems, data storages, analysis and reporting tools and decision-making. The naïve model was later used to inform the development of an ideal BI architecture consisting of a presentation layer, service layer, data layer, and source systems layer. The understanding of BI in general was also supplemented with detailed discussions on OLTP, OLAP and TCO.

Considering the detailed discussion on the above-mentioned aspects in chapter 2, and the findings from the survey conducted on the state of BI utilisation in chapter 7, the outcomes have been recognised to have ultimately answered the research question: *What is the current state-of-the-art of BI usage in organisations?*

9.3.2 Sub-objective 2: To investigate and understand the various BI models and their associated advantages and limitations.

This sub-objective was achieved in two parts. First, it was addressed in Chapter 3. This chapter provided a more specific understanding of BI by investigating the on-premise deployment of BI and the business value realised from such implementation. It addressed three on-premise BI deployment modes namely: isolated, overlapping and centralised and, or shared. These deployment modes contributed to a comprehensive understanding of how BI works in an organisation.

The on-premise BI deployment model was demonstrated by discussing four main components, namely, multiple source systems from the independent sources, different data stores, isolated or centralised and, or shared BI tools, and various types of reports. The model was later used to inform the development of an ideal BI architecture consisting of five layers namely: the presentation layer, network layer, service layer, data layer, and source systems layer. The Network layer was the component that made this architecture different as it facilitated the communication of data and access to various BI reports generated. A further understanding of implementing on-premise BI deployment was attained through detailed discussions with regards to the advantages, challenges and related TCO.

The second part was addressed in Chapter 4. In order to establish a broader understanding of the fundamental concepts, this chapter discussed the definitions of cloud computing. It was learnt from the literature review that cloud computing has been dealt with in detail in various information resources in terms of benefits, implications, potential uses, and costs savings through the utilisation of cloud service models, such as, SaaS, PaaS and IaaS. Chapter 4 also discussed various cloud deployment models, that is, private, public, hybrid, and community models in which BI applications can be implemented based on business requirements. It was discovered that various pricing models exist, and that the procurement of cloud-based BI solutions are achieved through commonly used pricing models, such as, subscription-based, pay-as-you-go, pay-per-user, pay-per-use, server-based, etc.

The cloud-based BI deployment model was discussed and demonstrated through the utilisation of the key cloud services model, namely, PaaS, IaaS, and SaaS. An ideal cloud-based BI architecture was developed based on the model and consists of four layers namely: the presentation layer, network layer, security layer, and service layer. The security layer is a new component compared to on-premise BI architecture and addresses important elements related to securing the transmission of data through the Internet, Intranet, Extranet, LAN, or WLAN. The service layer demonstrates various ways of implementing cloud-based BI solutions through SaaS, PaaS and IaaS services.

A further understanding of implementing cloud-based BI deployment was accomplished through detailed discussions on the advantages and challenges related to TCO.

Having discussed the above-mentioned aspects in detail in chapter 3 and chapter 4, and the findings from the survey conducted on the state of BI utilisation in chapter 7, the outcomes of the research have been acknowledged to have answered the research question, that is: *What are the most currently used models at organisation level and the limitations thereof?*

9.3.3 Sub-objective 3: To develop a BI model to overcome the existing BI deployment limitations.

This sub-objective was addressed in chapter 5. Firstly, chapter 5 provided a comprehensive coverage of the hybrid BI deployment solution and the various emerging technologies that can support the hybrid BI deployment solutions were discussed as well as the associated business value propositions.

Hybrid BI deployment was defined in detail in order to clarify the basic concepts. A hybrid BI deployment was defined as a combination of an on-premise BI deployment and a cloud-based BI deployment. This was followed by a discussion on the various emerging technologies, such as, in-memory data management, in-memory analytical and transactional processing, big data management, and hybrid data warehouse that

were found to be useful in augmenting an ideal hybrid BI implementation that meets the business requirements of an organisation.

Chapter 5 discussed the main contribution of this research, that is, the theoretical hybrid BI deployment model and hybrid BI architecture. The model demonstrates the utilisation of the key components, such as, the thin client interface, multiple business intelligence solutions, and multiple data sources and, or storage systems. Based on this model, an ideal theoretical hybrid BI architecture was developed. The architecture consists of five layers, namely, the presentation layer, application layer, network layer, service layer, and data layer. A further understanding of implementing hybrid BI deployment was established through detailed discussions on the advantages, challenges and related TCO. However, since the architecture that was developed was theoretical, it was necessary to evaluate and validate it in real-life by industry BI experts for plausible improvements.

Having discussed the above-mentioned aspects in detail in chapter 5, the outcomes of the research have been acknowledged to have answered the research question, that is: *How can an improved model be developed taking into consideration existing limitations?*

9.3.4 Sub-objective 4: To assess the state of BI utilisations in organisations by business users and validate the proposed BI model by experts.

This sub-objective was addressed in chapter 6, chapter 7 and chapter 8, respectively. Chapter 6 established the foundation in terms of how data will be collected for the evaluation and validation of the proposed model. The first survey covered the validation aspect, but with emphasis on BI in general in order to assess its relevance at organisational level. The second survey dealt with the evaluation of the proposed hybrid model for further improvement. Chapter 7 addressed the analysis and results of the assessment on the utilisation of BI in organisations by BI users in the South African context. It discussed in detail the degree of BI practice in organisations, the data management model used in organisations and its prevalence, the level of understanding of the BI concept, and the current pricing models for procuring BI solutions.

In chapter 8, the analysis and results that formed the basis for improving the proposed hybrid BI architecture was presented. It discussed in detail an improved architecture based on plausible inputs from the findings of the survey conducted. The improved architecture consists of four layers, namely, the front-end layer, service layer, back-end layer, and infrastructure layer. The architecture was further used to demonstrate how it can be used in real-life as a guide to address BI challenges faced by organisations as well as to realise business value from the utilisation of BI solutions.

Therefore, having discussed the above-mentioned aspects in detail in chapter 5, and the findings from the survey conducted in chapter 7 and chapter 8, the outcomes of the research have been declared to have answered the research question, that is: *How can the proposed model be evaluated and validated in a real-life context?*

9.4 Future research

Although the objectives of this research set out in chapter 1 have been achieved, and considering this study's literature review, limitations, and findings, the discussion below highlights some of the research areas that may need further investigation.

- The proposed architecture was developed based on theoretical reviews and was later validated in real life by BI experts. It is recommended that further research could be done by conducting a BI implementation project that follows the Systems Development Life Cycle approach along with the proposed architecture as guide. This could be an ideal postgraduate project at Honours level.
- Big data continues to be a challenge as new forms of data keep emerging. Since the proposed hybrid BI architecture is for BI in general, conducting a further specific investigation on the proposed architecture by focusing only on how to manage big data is another area worth exploring. It would be interesting to uncover what other constraints or adjustments may emerge in order to handle big data. This type of an investigation could be ideal for a research project at Honours or Masters' level.
- The proposed architecture is non-technology dependent meaning any combination of technologies and applications can be tested for implementation.

It would be useful if further investigations could shed some light on whether the architecture was adopted for a specific BI software vendor and whether the BI solution offerings could be implemented. Again, this type of an investigation could be ideal for an honours or Masters' research project.

- Based on the above vendor-specific implementation, another angle that can be considered is doing an investigation on technology acceptance on the implemented vendor-specific technology. This could be done using methodologies, such as, the Technology Acceptance Model (TAM or TAM2), Motivational Model (MM), Unified Theory of Acceptance and Use of Technology (UTAUT), etc. This type of an investigation could be ideal for a Masters' research project.
- The actual cost of implementing a BI solution is vital in determining the return on investment. If further research could be done to determine the TCO of the investment in the BI solution for either a vendor-specific or generic adoption. This type of an investigation can be ideal for a research project at Honours or Masters' level.

9.5 Conclusion

The purpose of this chapter was to provide a summary and conclusion for this research study. The findings on the state of BI utilisation in organisations were reported and it was discovered that most organisations utilise a number of BI solutions that are either deployed on-premise, in the cloud, or as hybrid solutions. A hybrid BI solution was found to be appropriate for managing large volumes of data and critical information residing in various locations. The benefits of using a hybrid BI solution were found to outweigh existing solutions.

The improved and proposed hybrid BI architecture has been recommended to serve as a guide when implementing a hybrid BI solution in an organisation. The design for the hybrid BI architecture includes the most significant technological aspects required for implementing a comprehensive and robust BI solution.

The summary discussion on how the research objectives were achieved covered the four main areas, namely, BI in general, on-premise BI deployment, cloud-based BI deployment, and the development of the hybrid BI solution as well as the resulting

business value respectively. The limitations of this study were highlighted by recommending potential future research that could supplement the current findings.

In conclusion, a hybrid BI solution is highly recommended for addressing a number of BI challenges faced by organisations with assorted solutions. The implementation of such a solution can be successfully achieved by considering the hybrid BI architecture recommended in this study.

REFERENCES

References were managed electronically with the Mendeley citation manager. The Mendeley citation manager offers approximately 2600 citation styles. There are many variations on referencing methods from which to choose. For consistency, the Harvard method of referencing (Harvard reference format 1 author-date) was used throughout this dissertation.

- Abadi, DJ. 2009. Data management in the cloud: limitations and opportunities. *IEEE Data Eng. Bull*, 32(1):3–12.
- Acebo, A. 2013. Experts' perspective. *Business Intelligence Journal*, 18(4) :39–46.
- Acker, O, Grone, F, Blockus, A & Bang C. 2011. In-memory analytics – strategies for real-time CRM. *Journal of Database Marketing & Customer Strategy Management*, 18(2):129–136.
- Alasuutari, P. 2010. The rise and relevance of qualitative research. *International Journal of Social Research Methodology*, 13(2):139–155.
- Ali, M, Khan, SU & Vasilakos, AV. 2015. Security in cloud computing: opportunities and challenges. *Information Sciences*, 305:357–383.
- Al-Roomi, M, Al-Ebrhim, S, Sabika, B & Ahmad. 2013. Cloud computing pricing models: a survey. *International Journal of Grid and Distributed Computing*, 6(5):93–106.
- Andreadis, G, Fourtounis, G & Bouzakis, KD. 2015. Collaborative design in the era of cloud computing. *Advances in Engineering Software*, 81:66–72.
- Applebaum, W. 2010. Managing the total cost of ownership of business intelligence [White Paper]. *SAP*:1–15.
- Armbrust, B, Fox, A, Griffith, R, Joseph, AD, Katz, R, Konwinski, A, Lee, G, Patterson, D, Rabkin, A, Stoica, I & Zaharia, M. 2010. A view of cloud computing. *Communication of the ACM*, 53(4):50–58.
- Asghar, S, Fong, S & Hussain, T. 2009. Business intelligence modeling: a case study of Disaster Management Organization in Pakistan. In *Fourth International Conference on Computer Sciences and Convergence Information Technology*. Ieee:673–678.
- Barnes, S, Ring, T, Yamunan, V, Franks, B, Hiltbrand, T, Taylor, J & Williams, S. 2012. Experts' perspective: requirements for predictive analytics. *Business Intelligence Journal*, 17(1):20–28.
- Barone, D, Topaloglou, T & Mylopoulos, J. 2012. Business intelligence modeling in action : a hospital case study. *CAiSE 2012, LNCS 7328*:502–517.
- Baxter, P & Jack, S. 2008. Qualitative case study methodology: study design and implementation for novice researchers. *The Qualitative Report*, 13(4):544–559.
- Bhatti, N. 2013. Overcoming data challenges with virtualization. *Business Intelligence Journal*, 18(4):25–32.
- Bibi, S, Katsaros, D & Bozanis, P. 2012. Business application acquisition: on-premise

- or SaaS-based solutions? *IEEE Computer Society*:86–93.
- Birst, 2013a. BI Buyer 's Kit - Understanding BI : a quick reference guide: 1–3.
- Birst, 2013b. Comparing the total cost of ownership of business intelligence solutions:1–27.
- Birst, 2013c. New survey results show cloud-based BI satisfaction outpaces on-premise BI according to business intelligence pros. *Birst: News and Events*. Available at: <https://www.birst.com/company/press/new-survey-results-show-cloud-based-bi-satisfaction-outpaces-on-premise-bi-according-to-business-intelligence-pros/> [Accessed June 25, 2015].
- Birst, 2015. Six symptoms of Life Science companies suffering from severe analytics deficiency our future needs:1–3.
- Birst, 2014. Why cloud BI? The 10 Ssubstantial Bbenefits of Ssoftware-as-a-Sservice Bbusiness lintelligence [White Paper]:1–6.
- Boateng, O, Singh, J & Singh, GP. 2012. Data warehousing. *Business Intelligence Journal*, 5(2):224–234.
- Boniface, M, Nasser, B, Papay, J, Phillips, SC, Servin, A, Yang, X, Zlatev, Z, Gogouvitis, SV, Katsaros, G, Konstanteli, K, Kousiouris, G, Menychats, A & Kyriazis D. 2010. Platform-as-a-service architecture for real-time quality of service management in clouds. In *IEEE - 2010 Fifth International Conference on Internet and Web Applications and Services*: 155–160.
- Brohman, MK, Parent, M, Pearce, MR & Wade, M. 2000. The business intelligence value chain: data-driven decision support in a data warehouse environment: an exploratory study 1. In *Proceedings of the 33rd Hawaii International Conference on System Sciences - 2000*:1–10.
- Bucur, C. 2012. Implications and directions of development of web business intelligence systems for Bbusiness community. *Economic Insights - Trends and Challenges*, 64(2):96–109.
- BusinessDictionary.com, 2015. Decision making. *BusinessDictionary.com*. Available at: <http://www.businessdictionary.com/definition/decision-making.html> [Accessed February 24, 2015].
- Chaudhuri, S, Dayal, U & Narasayya, V. 2011. An overview of business intelligence technology. *Communications of the ACM*, 54(8):88–98.
- Chavan, CR. 2012. Insight of business intelligence for strategic and corporate management. *International Journal of Research in Management, Economics and Commerce*, 2(11):335–350.
- Chen, H, Chiang, RHL & Storey, VC. 2012. Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, 36(4):1165–1188.
- Chen, H, Chiang, RHL & Storey, VC. 2012. Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, 36(4):1165–1188.
- Chuah, M. 2010. An enterprise business intelligence maturity model (EBIMM): conceptual framework. *IEEE*:303–308.
- Chuah, M & Wong, KL. 2013. An enterprise business intelligence maturity model: case study approach. *IEEE*: 1–4.
- Cooper, DR. & Schindler, PS. 2006. Business research methods. 9th Edition. McGraw

– Hill International Edition.

- Cortellessa, V, Marinelli, F & Potena, P. 2008. An optimization framework for “build-or-buy” decisions in software architecture. *Computers and Operations Research*, 35:3090–3106.
- Cruz, Z, Fernández-Alemán, J & Toval, A. 2015. Security in cloud computing: a mapping study. *Computer Science and Information Systems*, 12(1):161–184.
- Daneshgar, F, Low, GC & Worasinchai, L. 2013. An investigation of “build vs. buy” decision for software acquisition by small to medium enterprises. *Information and Software Technology*, 55(10):1741–1750.
- Davenport, TH & Harris, JG. 2007. The architecture of business intelligence. *Harvard Business School Press*:1–12.
- Davis, JR, Imhoff, C & White, C. 2009. Operational business intelligence: the state of the art:1–75.
- Dawson, CD. 2002. *Practical research methods: a user-friendly guide to mastering research techniques and projects*. Oxford:Newtec Place: HowToBooks Ltd.
- Dictionary.com, 2014. Business Intelligence. *Dictionary.com*. Available at: <http://www.dictionary.com/browse/business-intelligence?s=t> [Accessed June 12, 2014].
- Dimensional Research, 2013. Comparing cloud and on-premise BI: a survey of business intelligence users. *Dimensional Research*, (September):1-11.
- Distefano, S, Ghezzi, G, Guinea, S & Mirandola R. 2014. Dependability assessment of web service orchestrations. *IEEE Transactions on Reliability*, 63(3):689–705.
- Dupada, S, Gedela, RK, Aryasri, RC, Acharya, R. 2013. Building value chain through actionbale benchmarking for sustainability and excellence. In *2nd International Conference on Information Management in the Knowledge Economy*: 24–30.
- Dyché, J, Geiger, JG, McGivern, M, McKnight, W & Zeid, A. 2011a. BI experts’ perspective - BI Center of Excellence. *Business Intelligence Journal*, 16(1):42–50.
- Elbashir, MZ, Collier, PA, Sutton, SG, Davern, MJ & Leech, SA. 2013. Enhancing the business value of business intelligence: the role of shared knowledge and assimilation. *Journal of Information Systems*, 27(2):87–105.
- Elbashir, MZ, Collier, P & Davern, MJ. 2008. Measuring the effects of business intelligence systems: the relationship between business process and organizational performance. *International Journal of Accounting Information Systems*, 9(3):135–153.
- Evelson, B., 2013. The Forrester Wave: enterprise business intelligence platforms , Q4. *The Forrester*, (4):1–21.
- Evelson, B. 2014. The Forrester Wave™: Agile business intelligence platforms , Q3 2014:1–22.
- Forrester, 2015. *In-memory computing platforms simplify and accelerate application development*:1–2.
- Funahashi, M & Yoshikawa, S. 2011. Fujitsu’s approach to hybrid cloud systems. *Fujitsu Science Technology Journal*, 47(3):285–292.
- Garrison, G, Kim, S & Wakefield, RL. 2012. Success factors for deploying cloud

- computing. *Communication of the ACM*, 55(9):62–68.
- Gartner, 2015a. Architecture. *IT Glossary*. Available at: <http://www.gartner.com/it-glossary/?s=architecture>. [Accessed June 20, 2015].
- Gartner, 2017. Big data. *Gartner - IT Glossary*. Available at: <http://www.gartner.com/it-glossary/big-data> [Accessed May 18, 2017].
- Gartner, 2013. Business intelligence. *Gartner - IT Glossary*. Available at: <http://www.gartner.com/it-glossary/business-intelligence-bi/> [Accessed May 20, 2013].
- Gartner, 2014. Software as a service. *Gartner - IT Glossary*. Available at: <http://www.gartner.com/it-glossary/software-as-a-service-saas/> [Accessed June 22, 2014].
- Gartner, 2015b. Total cost of ownership. *Gartner - IT Glossary*. Available at: <http://www.gartner.com/it-glossary/total-cost-of-ownership-tco/> [Accessed June 20, 2015].
- Gartner, 2015c. Wireless Local-Area Network (WLAN). *Gartner - IT Glossary*. Available at: <http://www.gartner.com/it-glossary/wlan-wireless-local-area-network/> [Accessed June 16, 2015].
- Gash, D, Ariyachandra, T & Frolick, M. 2011. Looking to the clouds for business intelligence. *Journal of Internet Commerce*, 10(4):261–269.
- Gastermann, B, Stopper, M, Kossik, A & Katalinic, K. 2015. Secure implementation of an on-premises cloud storage service for small and medium-sized enterprises. *Procedia Engineering*, 100:574–583.
- Ghazanfari, M, Jafari, M & Rouhani, S. 2011. A tool to evaluate the business intelligence of enterprise systems. *Scientia Iranica*, 18(6):1579–1590.
- Gollapudi, K, Jangeti, SK & Kotapati, AR. 2012. Analysis of using a business intelligence tool (Cognos) in a company to result in more efficient and intuitive company in the current era. *Business Intelligence Journal*, 5(2):406–409.
- Gonzales, M. 2011. Success factors for business intelligence and data warehousing maturity and competitive advantage. *Business Intelligence Journal*, 16(1):22–30.
- Gray, J. 1996. Evolution of data management. *IEEE*, 29(10):38–46.
- Hagerty, J, Sallam, RL & Richardson, J. 2012. Magic quadrant for business intelligence platforms. *Gartner*, (February):1–54.
- Halper, F., 2014. BI StatShots. *Business Intelligence Journal*, (19)1:56-57.
- Han, Y. 2013. IaaS cloud computing services for libraries: cloud storage and virtual machines. *OCLC Systems & Services: International Digital Library Perspectives*, 29(2):87–100.
- Harriott, J. 2013. 7 pillars for successful analytics implementation. *Marketing Research*, 25(1):34–41.
- Hazen, BT, Boone, CA, Ezell, JD & Jones-Farmer LA. 2014. Data quality for data science, predictive analytics, and big data in supply chain management: an introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154:72–80.
- Henry, D & Nys, D. 2013. *Enterprise data warehouse optimization with Hadoop Big Data*: 1–24.

- Hertis, M & Juric, M. 2014. An empirical analysis of business process execution language usage. *IEEE Transactions on Software Engineering*, 40(8):738–757.
- Hewlett-Packard, 2012. *Shift from data to insight: A new approach to business intelligence and analytics*, rest of entry?
- Hobek, R, Ariyachandra, TR & Frolick, MN. 2006. The importance of soft skills in business intelligence implementations. *Business Intelligence Journal*, 14(1):28–37.
- Hortonworks & Denodo, 2015. A modern data architecture with Apache Hadoop - The Hybrid Data Warehouse:1–11.
- Hostmann, B. 2007. BI competency centers: bringing intelligence to the business. *Business Performance Management*, 5(4):4–10.
- Huang, M, Serres, O & Narayana, VK 2010. Efficient cache design for solid-state drives. In, *Proceedings of the 7th ACM international conference on Computing frontiers - CF '10*:41. Available at: <http://portal.acm.org/citation.cfm?doid=1787275.1787282>.
- IBM. 2012. In-memory processing: enhancing performance and analytics capabilities:1–8.
- IBM. 2015. What is cloud computing? *IBM Cloud*:1–7. Available at: <http://www.ibm.com/cloud-computing/us/en/what-is-cloud-computing.html> [Accessed May 26, 2015].
- Iyer, B & Henderson, JC. 2012. Business value from clouds: learning from users. *MIS Quarterly Executive*, 11(1):51–60.
- Jansen, W & Grance, T. 2011. Guidelines on security and privacy in public cloud computing. *National Institute of Standards and Technology Special Publication 800-144*, 800–144:1–80.
- Jiang, L, Barone, D, Amyot, D & Mylopoulos, J. 2011. Strategic models for business intelligence. *LNCS 6998*:, 429–439.
- Kalyvas, BJR, Overly, MR & Karlyn, MA. 2013a. Cloud computing: a practical framework for managing cloud computing risk, part I. *Intellectual Property & Technology Law Journal*, 25(3):7–18.
- Kalyvas, BJR, Overly, MR & Karlyn, MA, 2013b. Cloud computing: a practical framework for managing cloud computing risk, part II. *Intellectual Property & Technology Law Journal*, 25(4):19–28.
- Karim, AJ. 2011. The value of competitive business intelligence system (CBIS) to stimulate competitiveness in global market. *International Journal of Business and Social Science*, 2(19):196–204.
- Kaur, K, Azad, N & Sing, P. 2013. Cost effective cloud based business intelligence model for small- scale organizations. *International Journal of Computer Applications (0975-8887)*, 63(8):24–31.
- Kempler, S, Lynnes, C, Vollmer, B, Alcott, G & Berrick, S. 2009. Evolution of information management at the GSFC Earth Sciences (GES) Data and Information Services Center (DISC): 2006-2007. In, *IEEE Transactions on Geoscience and Remote Sensing*, 47(1):21–28.
- Khalid, K, Hilman, H & Kumar, DM. 2012. Get along with quantitative research

- process. *International Journal of Research in Management*, 2(2):15–29.
- Khan, AMA, Amin, N & Lambrou, N. 2010. Drivers and barriers to business intelligence adoption: a case of Pakistan. In *Proceedings of the European and Mediterranean Conference on Information Systems (EMCIS2010), Abu Dhabi, UAE*: 1–23.
- Khan, RA & Quadri, SMK. 2014. Business intelligence: an integrated approach. *International Journal of Management and Innovation*, 6(2):21–32.
- King, E. 2007. Key metrics for determining ROI for business intelligence implementations. *Unisphere Research*, (August):1–16.
- Kocakoç, ID. & Erdem, S. 2010. Business intelligence applications in retail business: OLAP, Data Mining & Reporting Services. *Journal of Information & Knowledge Management*, 9(2):171–181.
- Kolb, J. 2012. What is analytics? *Applied Data Labs*:1–13.
- Kothari, CR. 2004. *Research methodology: methods and techniques*. 2nd rev ed. New Delhi: New Age International (P) Limited, Publishers.
- Kruger, J. 2010. A study of strategic intelligence as a strategic management tool in the long-term insurance industry in South Africa. *University of South Africa*:17–376.
- Kwon, O, Lee, N & Shin, B. 2014. Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 34(3):387–394.
- Lind, DA, Marchal, WC & Wathen, SA. 2008. *Statistical techniques in business and economics with global data sets*. 13th Ed. McGraw: Hill Education.
- Van der Linde, J. 2013. *An investigation into the impact of Enterprise Architecture decisions on the responsibilities of software developers on the companies that develop software*. University of South Africa:1-174.
- Liu, F, Guo, W, Zhao, ZQ & Chou, W. 2010. SaaS integration for software cloud. In, *IEEE 3rd International Conference on Cloud Computing*:402–409.
- Liyang, T, Zhiwei, N, Zhzngjun, W & Li, W. 2011. A conceptual framework for business intelligence as a service (SaaS BI). In, *Fourth International Conference on Intelligent Computation Technology and Automation*:1025–1028.
- Lonnqvist, A & Pirttimaki, V. 2006. The measurement of business intelligence. *Information Systems Management*, 23(1):32–40.
- Lopez, K & Antoni, JD. 2014. The modern data warehouse: how big data impacts analytics architecture. *Business Intelligence Journal*, 19(3):8–16.
- Madaleno, R & Marinheiro, RN. 2013. Identification and ranking of relevant criteria for the selection of Software as a Service. In, *Information Systems and Technologies (CISTI), 2013 8th Iberian Conference*:.Publisher and place?: 1–6.
- Mantha, B. 2014. Five guiding principles for realizing the promise of big data. *Business Intelligence Journal*, 19(1):8–12.
- Manyika, J, Chui, M, Brown, B, Bughin, J, Dobbs, R, Roxburgh, C & Byers, AH. 2011. Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute*, (June):1–156.
- Marinela, M, Bogdan, G & Marian, S. 2011. Combining business intelligence with cloud computing to delivery agility in actual economy. *Economic Computation & Economic Cybernetics Studies & Research*, 45(1):1–16.

- Mate, A, Llorens, H & De Gregorio, E. 2012. An integrated multidimensional modeling approach to access big data in business intelligence platforms. *ER Workshops 2012, LNCS 7518*:1–120.
- Mateescu, G. & Vladescu, M., 2013. Secure on premise -on demand services communication. In, *System Theory, Control and Computing (ICSTCC), 2013 17th International Conference*. IEEE: 613–618.
- Mavodza, J. 2013. The impact of cloud computing on the future of academic library practices and services. *New Library World*, 114(3/4):132–141.
- Mell, P & Grance, T. 2011. The NIST definition of cloud computing. *National Institute of Standards and Technology*:1–7.
- Melville, N, Kraemer, K & Gurbaxani, V. 2004. Review: information technology and organizational performance: an integrative model of IT business value. *MIS Quarterly*, 28(2):283–322.
- Menon, L & Rehani, B. 2012. Business intelligence on the cloud: overview and use cases:1–15.
- MicroStrategy, 2012. Architecture for enterprise business intelligence: an overview of the MicroStrategy platform architecture for big data, cloud BI and mobile applications:1–443.
- MicroStrategy, 2008. *Reducing total cost of ownership: delivering cost effective enterprise business intelligence*:1–28.
- Minoli, D. 2008. *Enterprise architecture A to Z: frameworks, business process modeling, SOA, and infrastructure technology*. Boca Raton: Auerbach Publications.
- Moghimi, F & Zheng, C. 2009. A decision-making model to choose business intelligence platforms for organizations. In, *Third International Symposium on Intelligent Information Technology Application*:73–77.
- Mvelase, P, Dlodlo, N, Makitla, I, Sibiyi, G & Adigun, M. 2012. An architecture based on SOA and virtual enterprise principles: opennebula for cloud deployment. In, *Proceedings of the International Conference on Information Management & Evaluation*.: 214–223.
- Najjar, MS & Kettinger, WJ. 2013. Data monetization: lessons from a retailer's journey. *MIS Quarterly Executive*, 2013(December):21–32.
- Nunan, D. & Domenico, M. Di, 2013. Market research and the ethics of big data. *International Journal of Market Research*, 55(4), pp.2–14.
- O'Donnell, P, Sipsma, S & Watt, C. 2012. The critical issues facing business intelligence practitioners. *Journal of Decision Systems*, 21(3):203–216.
- Olsen, W., 2004. Triangulation in social research: qualitative and quantitative methods can really be mixed. *Developments in Sociology*:1–30.
- Olszak, C.M. & Ziemba, E., 2007. Approach to building and implementing business intelligence systems. *Interdisciplinary Journal of Information, Knowledge and Management*, 2:135–148.
- Ong, IL, Siew, PH & Wong, SF. 2011. A five-layered business intelligence architecture. *Communications of the IBIMA*:1–11.
- Oracle, 2014a. *Analysis of SAP HANA High Availability Capabilities*,

- Oracle, 2014b. Oracle Cloud Solution Overview:1–21.
- Oracle, 2012. Oracle OLAP:1–5.
- Orłowski, C, Szczerbicki, E. & Grabowski, J. 2014. Enterprise service bus architecture for the big data systems. *Management and Production Engineering Review*, 5(1):28–31.
- Ouf, S & Nasr, M. 2012. Business Intelligence in the Cloud. *IEEE*:650–655.
- Pant, P. 2009. Business intelligence (BI): how to build successful BI strategy:1–19.
- Parziale, L, Avramenko, A, Chan, S, De Valence, F, Dziekan, C, Dziekan, M, Greggo, A, Hagen, C, Lin, D, Machung, J & Roik, N. 2010. IBM Smart Analytics Cloud. *IBM:RedBooks*. Available at: <http://www.redbooks.ibm.com/redbooks/pdfs/sg247873.pdf>.
- Paul, K, Govindarajan, S & Chatterjee, D. 2013. Cloud computing : emphasizing hybrid cloud computing on Android computing platform – an overview. *International Journal of Applied Sciences & Engineering (IJASE)*, 1(1):21–28.
- Pentaho, 2013. The ultimate guide to buying business analytics:1–10.
- Phatak, M & Kamalesh, V. 2010. On cloud computing deployment architecture. In, *International Conference on Advances in ICT for Emerging Regions (ICTer)*, :11–14.
- Plattner, H & Zeier, A. 2012. *In-memory data management: technology and applications*. 2nd ed. Berlin,Heidelberg: Springer Verlag.
- Plattner, H. 2009. A common database approach for OLTP and OLAP using an in-memory column database. *ACM*:1–7.
- Pombriant, D. 2012. *Developing for the cloud: challenges and best practices*.TechTarget:1–12.
- Poonen, JS. 2011. The business information revolution. *SAPinsider Special Report*:1–20.
- Popovi, A, Turk, T & Jaklic, J. 2010. Conceptual model of business value of business intelligence systems. *Management*, 15(1):5–30.
- Porter, ME & Millar, VE. 1985. How information gives you competitive advantage. *Harvard Business Review*, (July-August):149–174.
- Power, DJ. 2008. Understanding data-driven decision support systems. *Information Systems Management*, 25(2):149–154.
- Qionqwei, Y, Yaotang, L & Qiuyun, N. 2012. Research on the application model of business intelligence (BI) in e-business from the perspective of chinese culture. In, *2nd International Conference on Computer Science and Network Technology*:907–911.
- Rahman, N, Aldhaban, F & Akhter, S. 2013. Emerging Ttechnologies in Bbusiness lintelligence. In, *Proceedings of PICMET '13: Technology Management for Emerging Technologies*: 542–547.
- Rahman, NHA & Choo, KKR. 2014. A survey of information security incident handling in the cloud. *Computers & Security*, 49:45–69.
- Ramamurthy, K, Sen, A & Sinha, AP. 2008. An empirical investigation of the key determinants of data warehouse adoption. *Decision Support Systems*, 44(4):817–

841.

- Read, K. 2013. Is “in-memory” always the right choice? *Business Intelligence Journal*, 18(1):46–51.
- Reddy, PA & Rao, KRS. 2012. Metrics and quality economics for defect and detection techniques for measuring the software cost. *International Journal of Computers & Communications*, 1(1):40–46.
- Rosslyn Analytics. 2014. Why on-premise business intelligence software and traditional return-on- investment models are costing your company millions (preface from Jeffrey Kaplan of THINKstrategies):1–8.
- Rugg, G & Petre, M. 2007. *A gentle guide to research methods*, Maidenhead, Berkshire: Open University Press.
- Russom, P. 2013. Managing big data. *TDWI Research*:1–39.
- Safeer, M & Zafar, S. 2011. Impact of business intelligence competency center in success/ failure of B.I. applications. In, *IEEE 14th International Multitopic Conference*. IEEE:267–272.
- Sallam, RL, Hostmann, B, Schlegel, K, Tapadinhas, J & Oestrelch, TW. 2015. *Magic quadrant for business intelligence and analytics platforms*. Gartner:1-85.
- SAP. 2013. Delivering information faster: in-memory technology reboots the big data analytics world , (February):1–6.
- SAP. 2012. SAP HANA ® Database for next-Generation Business Applications and Real-Time Analytics:1–18.
- SAP. 2014. SAP OLAP. *SAP Help Portal*:1–136. Available at: http://help.sap.com/saphelp_erp60_sp/helpdata/en/7c/c3e60666cd9147bb6242dc6500cd77/content.htm [Accessed June 22, 2014].
- SAP, 2015. SAP S/4HANA Frequently asked questions , (May):1–18.
- Schauer, J. 2004. *The new era of BI and business analytics*, DM Review: 28-30.
- Schlegel, K, Milbury, O, Buytendijk, F & Sommer, D. 2014. Create a centralized and decentralized organizational model for business intelligence. Gartner:1–7.
- Schmidt, MT, Hutchison, B, Lambros, P & Phippen, R. 2005. The Enterprise Service Bus: making service-oriented architecture real. *IBM Systems Journal*, 44(4):781–797.
- Sen, A, Ramamurthy, K & Sinha, AP. 2012. A model of data warehousing process maturity. *IEEE Transactions on Software Engineering*, 38(2):336–353.
- Sibanda, ND. 2009. Quantitative research. In, *Quantitative Research*:1–33.
- Sikka, V. 2013. Re-thinking the performance of information processing systems. *IEEE 29th International Conference on Data Engineering (ICDE)*. IEEE:9–13.
- Siller, GG. 2013. *FeatureIT: a platform for collaborative software development*. University of South Africa:1-147.
- Silva, E, Vollino, B, Becker, K & Galante, R. 2012. A business intelligence approach to support decision making in service evolution management. In, *IEEE Ninth International Conference on Services Computing*. IEEE:41–48.
- Silva, E, Becker, K & Galante, R. 2013. Supporting strategic decision making on service evolution context using business intelligence. In, *Services Computing*

- (SCC), *2013 IEEE International Conference*:240–247.
- Srinivasan, M. 2011. Cloud-based email architecture for higher education institutions. *Issues in Information Systems*, 12(1):339–345.
- Suciu, G, Halunga, S, Apostu, A, Vulpe, A & Todoran, G. 2013. Cloud computing as evolution of distributed computing: a case study for SlapOS Distributed Cloud Computing Platform. *Informatica Economica*, 17(4):109–122.
- Sukamolson, S. 2007. Fundamentals of quantitative research. *Language Institute Chulalongkorn University*:1–20.
- Summar, V. 2012. Magic quadrant for business intelligence platforms. Gartner:1–27.
- Sun, X, Gao, B, Fan, L & An W. 2012. A cost-effective approach to delivering analytics as a service. In, *IEEE 19th International Conference on Web Services*, (1):512–519.
- TEC. 2011. *Business intelligence buyer's guide: BI for everyone*. Technology Evaluation Centers:1–101.
- TEC, 2012. *Business intelligence and data management: buyer's guide*. Technology Evaluation Centers:1–126.
- TEC, 2014. *Mobile BI buyer's guide*. Technology Evaluation Centers:1–80.
- TEC, 2015. *Cloud BI and analytics buyer's guide*. Technology Evaluation Centers:1–95.
- Techopedia, 2015. Network layer. *Techopedia*:1. Available at: <http://www.techopedia.com/definition/24204/network-layer> [Accessed June 16, 2015].
- Thomsen, C & Pedersen, TB. 2008. A survey of open source tools for business intelligence. *International Journal of Data Warehousing and Mining*, 5(3):56–75.
- Trikha, B. 2010. A journey from floppy disk to cloud storage. *International Journal on Computer Science & Engineering*, 2(4):1449–1452.
- Wagner, HT & Weitzel, T. 2012. How to achieve operational business-IT alignment: insights from a global aerospace firm. *MIS Quarterly Executive*, 11(1):25–36.
- Watson, HJ. 2011. Business analytics insight: hype or here to stay? *Business Intelligence Journal*, 16(1):4–9.
- Watson, HJ. 2012a. The necessary skills for advanced analytics. *Business Intelligence Journal*, 17(4):4–8.
- Watson, HJ. 2012b. This isn't your mother's BI architecture. *Business Intelligence Journal*, 17(1):4–7.
- Watson, HJ & Wixom, BH. 2007. The current state of business intelligence. *IEEE Computer Society*,:96–99.
- Webb, B. 2008. *SAS in-database processing with Teradata: an overview of foundation technology*. SAS Institute Inc:1-22.
- Weber, M. 2013. Keys to sustainable self-service business intelligence. *Business Intelligence Journal*, 18(1):18–24.
- Wegner, T. 2007. *Applied business statistics: methods and excel based applications*. 2nd Ed.:Juta & Co. Ltd.

- Wikipedia, 2014. On-premises software. *Wikipedia, the free encyclopedia*. Available at: http://en.wikipedia.org/wiki/On-premises_software [Accessed December 2, 2014].
- Williams, S & Williams, N. 2003. The business value of business intelligence. *Business Intelligence Journal*, (301):1–11.
- Williams, S & Williams, N. 2007. *The profit impact of business intelligence*. San Francisco:Morgan Kaufmann.
- Winkler, TJ & Brown, CV. 2013. Horizontal allocation of decision rights for on-premise applications and Software-as-a-Service. *Journal of Management Information Systems*, 30(3):13–48.
- Wixom, BH, Watson, HJ & Werner, T. 2011. Developing an enterprise business intelligence capability: the Norfolk southern journey. *MIS Quarterly Executive*, 10(2):61–72.
- Yang, SQ. 2012. Move into the Cloud, shall we? *Library HI-TECH NEWS*, (1):4–7.
- Zainal, Z. 2007. Case study as a research method. *Academy of Management Review*:1–6.
- Zhang, L & Tu, X. 2009. A feasible enterprise business intelligence design model. In, *International Conference on Management of e-Commerce and e-Government*. IEEE:182–187.
- Zhang, X. 2009. Evolution of business information systems on the integration and intelligence dimensions. In, *Proceedings: International Conference on Management and Service Science, MASS*. IEEE:1–5.
- ZiffDavis. 2012. Maximizing performance and minimizing cost with interactive business intelligence. Ziff Davis Inc.:1–3.
- Zimmerman, DK. 2014. Five cloud essentials for the boardroom : what banking and financial markets executives need to know about cloud computing. *Journal of Payment Strategy & Systems*, 8(1):84–93.

APPENDIX 1: ETHICAL CLEARANCE DOCUMENTATION

UNISA

college of
science, engineering
and technology

Dear Mr. Misheck Banda (39151123)

Date: 2015-11-24

Application number:
124/MB/2015/CSET_SOC

REQUEST FOR ETHICAL CLEARANCE: (A Big Data Management and Analytic Model for Cloud-based Business Intelligence Applications)

The College of Science, Engineering and Technology's (CSET) Research and Ethics Committee has considered the relevant parts of the studies relating to the abovementioned research project and research methodology and is pleased to inform you that ethical clearance is granted for your research study as set out in your proposal and application for ethical clearance.

Therefore, involved parties may also consider ethics approval as granted. However, the permission granted must not be misconstrued as constituting an instruction from the CSET Executive or the CSET CRIC that sampled interviewees (if applicable) are compelled to take part in the research project. All interviewees retain their individual right to decide whether to participate or not.

We trust that the research will be undertaken in a manner that is respectful of the rights and integrity of those who volunteer to participate, as stipulated in the UNISA Research Ethics policy. The policy can be found at the following URL:

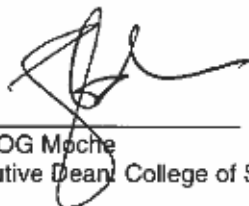
http://cm.unisa.ac.za/contents/departments/res_policies/docs/ResearchEthicsPolicy_apprvCounc_21Sept07.pdf

Please note that the ethical clearance is granted for the duration of this project and if you subsequently do a follow-up study that requires the use of a different research instrument, you will have to submit an addendum to this application, explaining the purpose of the follow-up study and attach the new instrument along with a comprehensive information document and consent form.

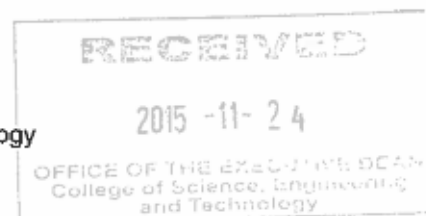
Yours sincerely



Prof Ernest Mnkandla
Chair: College of Science, Engineering and Technology Ethics Sub-Committee



Prof IOG Mophe
Executive Dean, College of Science, Engineering and Technology



APPENDIX 2: RESEARCH QUESTIONNAIRE – BI EXPERTS

[Soft and hard copy survey administration]



M Banda – Std. No. 39151123

A LETTER REQUESTING AN ADULT TO PARTICIPATE IN A QUESTIONNAIRE

Dear participant,

This is an invitation to consider participating in a study I, Misheck Banda, am conducting as part of my research as a Master's student entitled A Big Data Management and Analytic Model for Cloud-based Business Intelligence Applications at the University of South Africa. Permission for the study has been given by School of Computing, College of Science, Engineering and Technology and the Ethics Committee of the College of Education, UNISA. I have purposefully identified you as a possible participant because of your valuable experience and expertise related to my research topic.

I would like to provide you with more information about this project and what your involvement would entail if you should agree to take part. The importance of having a big data management and analytic model is that it can support organisations with making informed and evidence based decisions. In this interview/questionnaire I would like to have your views and opinions on this topic. This information will be used to improve my theoretical hybrid BI model and architecture in my research.

Your participation in this study is voluntary. It will involve answering a questionnaire of approximately 10-15 minutes in length at a time convenient to you. Furthermore, you may decide to withdraw from this study at any time without any negative consequences.

All information you provide is considered completely confidential. Your name will not appear in any publication resulting from this study and any identifying information will not be collected. However, with your permission, anonymous quotations may be used. Data collected during this study will be retained on a password protected computer for 12 months. There are no known or anticipated risks to you as a participant in this study.

If you have any questions regarding this study, or would like additional information to assist you in reaching a decision about participation, please contact me at 082 798 0264 or by e-mail at 39151123@mylife.unisa.ac.za.

I look forward to receiving your response and thank you in advance for your assistance in this project.

Yours sincerely
Misheck

Please complete the questionnaire by **marking (X)** on your selection – double click on your answer and insert **X**.

Section 1: Demographic Information

1.1 Years of experience in your field:

- 2 to 4 years
- Between 5 and 7 years
- Between 8 and 10 years
- More than 10 years

1.2 Industry of your organisation:

- | | |
|--|--|
| <input type="radio"/> Automotive | <input type="radio"/> Marketing & Media |
| <input type="radio"/> Finance (Banking, Insurance) | <input type="radio"/> Medical |
| <input type="radio"/> Government & Parastatals | <input type="radio"/> Property |
| <input type="radio"/> Information Communication & Technology | <input type="radio"/> Retail |
| <input type="radio"/> Legal | <input type="radio"/> Travel & Tourism |
| <input type="radio"/> Logistics & Transport | <input type="radio"/> Other (please specify) |
| <input type="radio"/> Manufacturing | <input type="text"/> |

1.3 Please indicate the level of deployment of BI in your organisation (you can select more than one):

- Centralised
- Distributed
- Isolated
- Other

If other, please specify:

1.4 Please indicate the geographical spread of your organisation:

- | | | | |
|-------------------------------------|---------------------------------------|--|--|
| <input type="checkbox"/> Gauteng | <input type="checkbox"/> Eastern Cape | <input type="checkbox"/> Western Cape | <input type="checkbox"/> International |
| <input type="checkbox"/> Limpopo | <input type="checkbox"/> North West | <input type="checkbox"/> KwaZulu Natal | |
| <input type="checkbox"/> Mpumalanga | <input type="checkbox"/> Free State | <input type="checkbox"/> Northern Cape | |

Section 2: Familiarity with BI Deployment and Pricing Options

Most of the organisations these days are deploying a combination of BI solutions such as On-premise, Cloud-based or Hybrid in order to manage their large volumes of data, critical information, and make informed decisions among other benefits.

2.1. Please provide your assessment for each of the following statements with respect to your expertise and knowledge of BI deployment options:

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
An on-premise BI deployment has all the hardware, software and application systems located inside the organisation's premise.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
An on-premise BI deployment ensures ownership of business data.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A cloud-based BI deployment uses the standard cloud service models such as SaaS, PaaS and IaaS.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A cloud-based BI deployment promotes the use of faster and cheaper BI services.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A Hybrid BI deployment is a combination of on-premise and cloud-based BI solutions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A Hybrid BI allows for sensitive information to be kept on-premise and the rest to be managed in the cloud.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2.2. Please assess the following statements with respect to pricing models for a BI solution.

Question	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
Once-off purchase is the ideal option to take when acquiring BI applications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pay-as-you-go is the ideal option to take when acquiring BI applications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pay-per-user is the ideal option to take when acquiring BI applications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pay-per-use is the ideal option to take when acquiring BI applications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Server-based is the ideal option to take when acquiring BI applications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Subscription-based is the ideal option to take when acquiring BI applications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Section 3: Front-end Layer

A Front-end layer of a BI architecture acts as the entry point or interface from where business users can access data for enterprise reporting through the use of BI tools that resides either on-premise or in the cloud.

3.1. Please assess the following statements with respect to an architectural Front-end layer of a BI solution.

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
A good BI architecture must consider devices (e.g. Desktop Computers, Laptops, Tablets, and Smartphones) in their architecture front-end.	0	0	0	0	0
A good BI architecture must consider applications (e.g. Online Analytical Processing, Spreadsheets, Data Mining, Reports, and Dashboards) in their architecture front-end.	0	0	0	0	0
A good BI architecture must consider BI Tools (e.g. Microsoft Power BI, IBM Cognos, MicroStrategy, Pentaho, Birst, Oracle BI, and SAP BI/BO) in their architecture front-end.	0	0	0	0	0
A good BI architecture must consider Presentation layer in their architecture front-end.	0	0	0	0	0
A good BI architecture must consider Interface layer in their architecture front-end.	0	0	0	0	0

3.2. Please indicate which additional components could be part of the architecture Front-end layer:

- Business Process Execution Language (BPEL)
- Business process tools
- Workflow tools
- Other

If other, please specify:

Section 4: Service Layer

A Service layer in a BI architecture comprise of the analytic engines that are responsible for executing analytical functions requested by business users in order to achieve and support better decision making processes.

4.1. Please assess the following statements with respect to an architectural Service layer of a BI solution.

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
A good BI architecture must consider a Service layer in their architecture.	0	0	0	0	0
The Enterprise Service Bus (ESB) usually referred to as middleware or Service-oriented Architecture (SOA) should be positioned in the Service layer.	0	0	0	0	0
BI services must have a dedicated layer referred to as the Service layer.	0	0	0	0	0
The ESB and BI services must have a dedicated layer referred to as the Service layer.	0	0	0	0	0
The Service layer should consist of ESB, BI, Platform-as-a-Service (PaaS), Database Management System (DBMS), and Application Servers.	0	0	0	0	0
The services layer should consist of ESB, BI services, and DBMS	0	0	0	0	0
The services layer should consist of ESB, BI services and PaaS	0	0	0	0	0

4.2. Please indicate which additional components could be part of the architecture Service layer:

- Visualisation MapReduce services
- User Applications (including in-memory applications)
- User/Machine Interface
- Model-View-Controller (MVC)
- Data collectors
- Other

If other, please give examples:

Section 5: Back-end Layer

A Back-end layer in a BI architecture is a server side where all the applications are executed in the background without the client user's knowledge.

5.1. Please assess the following statements with respect to an architectural Back-end layer of a BI solution.

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
A good BI architecture must consider a Back-end layer in their architecture.	0	0	0	0	0
The PaaS should be part of the Back-end layer.	0	0	0	0	0
The PaaS must have a dedicated layer.	0	0	0	0	0
The Back-end layer should consist of PaaS, DBMS, and Application Servers.	0	0	0	0	0
The Back-end layer should consist of DBMS and Application servers.	0	0	0	0	0
The DBMS must have a dedicated layer.	0	0	0	0	0
The Application servers must have a dedicated layer.	0	0	0	0	0
A good BI architecture must consider data source systems such as External data sources, Operational data sources, Legacy data sources, OLTP source systems, etc. in their Back-end layer.	0	0	0	0	0

5.2. Please indicate which additional components could be part of the architecture Back-end layer:

- Development tools
- Middleware
- Execution Runtime Environment
- In house Application Programming Interface (API)
- Third Party API
- Application Server
- Web Server
- Other

If other, please give examples:

Section 6: Infrastructure layer

An Infrastructure layer in a BI architecture comprise of BI resources related to hardware and data management systems.

6.1. Please assess the following statements with respect to an architectural Infrastructure layer of a BI solution.

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
A good BI architecture must consider an Infrastructure layer in their architecture.	0	0	0	0	0
A good BI architecture must consider hardware resources such as Storage resources, Computing resources, Virtual machines, etc. in their architecture.	0	0	0	0	0
A good BI architecture must consider database technologies such as Cloud/hosted storage, Disk drives, Solid-state drives, Optical disks, etc. in their architecture.	0	0	0	0	0
A good BI architecture should have a dedicated Data Management layer.	0	0	0	0	0
A good BI architecture should have a dedicated Network layer.	0	0	0	0	0
A Network layer should be part of the Infrastructure layer.	0	0	0	0	0
An Infrastructure layer should not be part of the BI architecture.	0	0	0	0	0

6.2. Please indicate which additional components could be part of the architecture Infrastructure layer:

- Infrastructure-as-a-Service (IaaS)
- Routers
- Switches
- Virtual Private Network
- Access Point Name
- Other

If other, please give examples:

*****END*****

Responses cannot be edited

Business Intelligence Questionnaire (Experts)

This questionnaire is aimed at obtaining your opinion on what would be the best way of implementing a Business Intelligence (BI) solution. An example of a BI solution is any type of software application that has the capability to report, analyse, and present information that can assist an organisation make better decisions. This questionnaire can be answered by anyone performing the roles such as Business Intelligence, Business Architect, Project Manager, Product Manager, Data Analytics, Business Analyst, System Analyst, etc.

Section 1: Demographic Information

Please complete the questionnaire by selecting the answer from the available options.

1.1. Years of experience in your field:

- 2 to 4 years
- Between 5 and 7 years
- Between 8 and 10 years
- More than 10 years

1.2. Industry of your organisation: *

- Automotive
- Finance (Banking, Insurance)
- Government & Parastatals
- Information Communication & Technology
- Legal
- Logistics & Transport
- Manufacturing
- Marketing & Media
- Medical
- Property
- Retail
- Travel & Tourism
- Other :

1.3. Please indicate the level of deployment of BI in your organisation (you can select more than one): *

- Centralised
- Distributed
- Isolated
- Other: _____

1.4. Please indicate the geographical spread of your organisation: *

- Gauteng
- Limpopo
- Mpumalanga
- Eastern Cape
- North West
- Free State
- Western Cape
- KwaZulu Natal
- Northern Cape
- International
- Other: _____



Section 2: Familiarity with BI Deployment and Pricing Options

Most of the organisations these days are deploying a combination of BI solutions such as On-premise, Cloud-based or Hybrid in order to manage their large volumes of data, critical information, and make informed decisions among other benefits.

2.1. Please provide your assessment for each of the following statements with respect to your expertise and knowledge of BI deployment options:

An on-premise BI deployment has all the hardware, software and application systems located inside the organisation's premise. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

An on-premise BI deployment ensures ownership of business data. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A cloud-based BI deployment uses the standard cloud service models such as SaaS, PaaS and IaaS. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A cloud-based BI deployment promotes the use of faster and cheaper BI services. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A Hybrid BI deployment is a combination of on-premise and cloud-based BI solutions. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A Hybrid BI allows for sensitive information to be kept on-premise and the rest to be managed in the cloud. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

2.2. Please assess the following statements with respect to pricing models for a BI solution.

Once-off purchase is the ideal option to take when acquiring BI applications. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

Pay-as-you-go is the ideal option to take when acquiring BI applications. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

Pay-per-user is the ideal option to take when acquiring BI applications. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

Pay-per-use is the ideal option to take when acquiring BI applications. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

Server-based is the ideal option to take when acquiring BI applications. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

Subscription-based is the ideal option to take when acquiring BI applications. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

Section 3: Front-end Layer

A Front-end layer of a BI architecture acts as the entry point or interface from where business users can access data for enterprise reporting through the use of BI tools that resides either on-premise or in the cloud.

3.1. Please assess the following statements with respect to an architectural Front-end layer of a BI solution.

A good BI architecture must consider devices (e.g. Desktop Computers, Laptops, Tablets, and Smartphones) in their architecture front-end. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture must consider applications (e.g. Online Analytical Processing, Spreadsheets, Data Mining, Reports, and Dashboards) in their architecture front-end. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture must consider BI Tools (e.g. Microsoft Power BI, IBM Cognos, MicroStrategy, Pentaho, Birst, Oracle BI, and SAP BI/BO) in their architecture front-end. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture must consider Presentation layer in their architecture front-end. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture must consider Interface layer in their architecture front-end. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

3.2. Additional components - Front-end layer

Please indicate which additional components could be part of the architecture Front-end layer: *

- Business Process Execution Language (BPEL)
- Business process tools
- Workflow tools
- Other: _____

Section 4: Service Layer

A Service layer in a BI architecture comprise of the analytic engines that are responsible for executing analytical functions requested by business users in order to achieve and support better decision making processes.

4.1. Please assess the following statements with respect to an architectural Service layer of a BI solution.

A good BI architecture must consider a Service layer in their architecture. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The Enterprise Service Bus (ESB) usually referred to as middleware or Service-oriented Architecture (SOA) should be positioned in the Service layer. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

BI services must have a dedicated layer referred to as the Service layer. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The ESB and BI services must have a dedicated layer referred to as the Service layer. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The Service layer should consist of ESB, BI, Platform-as-a-Service (PaaS), Database Management System (DBMS), and Application Servers. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The services layer should consist of ESB, BI services, and DBMS. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The services layer should consist of ESB, BI services and PaaS. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

4.2. Additional components - Service layer

Please indicate which additional components could be part of the architecture Service layer: *

- Visualisation MapReduce services
- User Applications (including in-memory applications)
- User/Machine Interface
- Model-View-Controller (MVC)
- Data collectors
- Other: _____



Section 5: Back-end Layer

A Back-end layer in a BI architecture is a server side where all the applications are executed in the background without the client user's knowledge.

5.1. Please assess the following statements with respect to an architectural Back-end layer of a BI solution.

A good BI architecture must consider a Back-end layer in their architecture. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The PaaS should be part of the Back-end layer. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The PaaS must have a dedicated layer. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

The Back-end layer should consist of PaaS, DBMS, and Application Servers. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The Back-end layer should consist of DBMS and Application servers. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

The DBMS must have a dedicated layer. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

The Application servers must have a dedicated layer. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture must consider data source systems such as External data sources, Operational data sources, Legacy data sources, OLTP source systems, etc. in their Back-end layer. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

5.2 Additional components - Back-end layer:

Please indicate which additional components could be part of the architecture Back-end layer: *

- Development tools
- Middleware
- Execution Runtime Environment
- In house Application Programming Interface (API)
- Third Party API
- Application Server
- Web Server
- Other: _____

Section 6: Infrastructure layer

An Infrastructure layer in a BI architecture comprise of BI resources related to hardware and data management systems.

6.1. Please assess the following statements with respect to an architectural Infrastructure layer of a BI solution.

A good BI architecture must consider an Infrastructure layer in their architecture. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture must consider hardware resources such as Storage resources, Computing resources, Virtual machines, etc. in their architecture. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture must consider database technologies such as Cloud/hosted storage, Disk drives, Solid-state drives, Optical disks, etc. in their architecture. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture should have a dedicated Data Management layer. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A good BI architecture should have a dedicated Network layer. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

A Network layer should be part of the Infrastructure layer. *

	1	2	3	4	5	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

An **Infrastructure layer** should not be part of the BI architecture. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

6.2. Additional components - Infrastructure layer

Please indicate which additional components could be part of the architecture **Infrastructure layer**: *

- Infrastructure-as-a-Service (IaaS)
- Routers
- Switches
- Virtual Private Network
- Access Point Name
- Other:

This content is neither created nor endorsed by Google.

Google Forms

APPENDIX 3: RESEARCH QUESTIONNAIRE – BI USERS



M Banda – Std. No. 39151123

A LETTER REQUESTING AN ADULT TO PARTICIPATE IN A QUESTIONNAIRE

Dear participant,

This is an invitation to consider participating in a study I, Misheck Banda, am conducting as part of my research as a Master's student entitled A Big Data Management and Analytic Model for Cloud-based Business Intelligence Applications at the University of South Africa. Permission for the study has been given by School of Computing, College of Science, Engineering and Technology and the Ethics Committee of the College of Education, UNISA. I have purposefully identified you as a possible participant because of your valuable experience and expertise related to my research topic.

I would like to provide you with more information about this project and what your involvement would entail if you should agree to take part. The importance of having a big data management and analytic model is that it can support organisations with making informed and evidence based decisions. In this interview/questionnaire I would like to have your views and opinions on this topic. This information will be used to improve my theoretical hybrid BI model and architecture in my research.

Your participation in this study is voluntary. It will involve answering a questionnaire of approximately 10-15 minutes in length at a time convenient to you. Furthermore, you may decide to withdraw from this study at any time without any negative consequences.

All information you provide is considered completely confidential. Your name will not appear in any publication resulting from this study and any identifying information will not be collected. However, with your permission, anonymous quotations may be used. Data collected during this study will be retained on a password protected computer for 12 months. There are no known or anticipated risks to you as a participant in this study.

If you have any questions regarding this study, or would like additional information to assist you in reaching a decision about participation, please contact me at 082 798 0264 or by e-mail at 39151123@mylife.unisa.ac.za.

I look forward to receiving your response and thank you in advance for your assistance in this project.

Yours sincerely
Misheck

Please complete the questionnaire by **marking (X)** on your selection – double click on your answer and insert **X**.

Section 1: Demographic information

1.1. Years of experience in your field:

- 2 to 4 years
- Between 5 and 7 years
- Between 8 and 10 years
- More than 10 years

1.2. Size of your organisation:

- Micro
- Small
- Medium
- Large

1.3. Industry of your organisation:

- | | |
|--|--|
| <input type="radio"/> Automotive | <input type="radio"/> Marketing & Media |
| <input type="radio"/> Finance (Banking, Insurance) | <input type="radio"/> Medical |
| <input type="radio"/> Government & Parastatals | <input type="radio"/> Property |
| <input type="radio"/> Information Communication & Technology | <input type="radio"/> Retail |
| <input type="radio"/> Legal | <input type="radio"/> Travel & Tourism |
| <input type="radio"/> Logistics & Transport | <input type="radio"/> Other (please specify) |
| <input type="radio"/> Manufacturing | <input type="text"/> |

1.4. Please indicate the number of business users that have access to BI services:

- Less than 10 Users
- Between 10 and 50 Users
- Between 51 and 100 Users
- More than 100 Users

Section 2: Business Intelligence Deployment

Q.1. Please provide your assessment for each of the following statements with respect to software capabilities that support your BI solution.

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
We use a single point of entry to access our BI solution.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We only have one data generating system and one user performing all our BI related functions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We have multiple data generating systems and multiple users performing complex BI related functions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We use simple tools such as Microsoft Excel, Google Sheets, etc. to store data, retrieve, analyse and produce reports.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We use complex BI tools such as OLAP server, SAS BI, SAP BI, Oracle BI, etc. to store data, retrieve, analyse and produce reports.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our reporting and analysis tool allows us to perform analytical functions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our reporting and analysis tool produce information that assist us to make informed decisions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We use self-service BI solution with timely insight and decision making.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We have experienced BI professionals that use our BI solution.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
You do not need to be a BI professional or expert in order to use our BI solution.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q.2. Please provide your assessment for each of the following statements with respect to management of data source/generating systems for your BI solution.

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
Our data is managed through the use of On-line Transaction Processing systems such as ERP, CRM, Sales, etc.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data source systems are managed internally.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data source systems are managed externally by our service provider.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data from our source systems is stored in a simple database.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data from our source systems is stored in a complex data warehouse.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
Data is accessed from a central repository at our office.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data is easily accessible from the data repository residing at our service provider.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our data goes through a data cleansing process (Extract, Transform, Load) before we gain access to it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q.3. Please provide your assessment for each of the following statements with respect to location and management of your BI solution.

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
Our BI solution is installed and run on computers located in our building.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our BI solution is installed and run on servers located at our service provider.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We own all the hardware and software used to support our BI solution.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We have full control of our in-house data processing applications and development environment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We have full control of our remote data processing applications and development environment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We have full control of our storage and computing resources that reside at a remote data centre.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We have full control of our storage and computing resources that reside at our office.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q.4. Please provide your assessment for each of the following statements with respect to buying or licensing options for your BI solution.

Statement	Strongly agree	Agree	Don't Know	Disagree	Strongly disagree
We use a free version of any available BI applications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our BI solution was an off-the-shelf purchased.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our BI user account is a Pay-as-you-go.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our BI user account is a Pay-per-user.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Our BI user account is a Pay-per-use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our BI user account is a Subscription-based.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*****END*****

[Online survey administrations – printable version]

Responses cannot be edited

Business Intelligence Questionnaire (BI User)

This questionnaire is aimed at understanding the types of Business Intelligence (BI) solutions that are currently been used by business users. An example of a BI solution is any type of software application that has the capability to report, analyse, and present information that can assist an organisation make better decisions.

Section 1: Demographic Information

Please complete the questionnaire by selecting the answer from the available options.

1.1. Years of experience in your field:

- 2 to 4 years
- Between 5 and 7 years
- Between 8 and 10 years
- More than 10 years

1.2. Size of your organisation: *

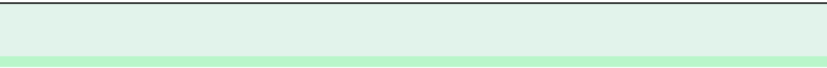
- Micro
- Small
- Medium
- Large

1.3. Industry of your organisation: *

- Automotive
- Finance (Banking, Insurance)
- Government & Parastatals
- Information Communication & Technology
- Legal
- Logistics & Transport
- Manufacturing
- Marketing & Media
- Medical
- Property
- Retail
- Travel & Tourism
- Other :

1.4. Please indicate the number of business users that have access to BI services: *

- Less than 10 Users
- Between 10 and 50 Users
- Between 51 and 100 Users
- More than 100 Users



2.1. Please provide your assessment for each of the following statements with respect to software capabilities that support your BI solution.

We use a single point of entry to access our BI solution. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

We only have one data generating system and one user performing all our BI related functions. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

We have multiple data generating systems and multiple users performing complex BI related functions. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

We use simple tools such as Microsoft Excel, Google Sheets, etc. to store data, retrieve, analyse and produce reports. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

We use complex BI tools such as OLAP server, SAS BI, SAP BI, Oracle BI, etc. to store data, retrieve, analyse and produce reports. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

Our reporting and analysis tool allows us to perform analytical functions.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

Our reporting and analysis tool produce information that assist us to make informed decisions.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

We use self-service BI solution with timely insight and decision making.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

We have experienced BI professionals that use our BI solution.

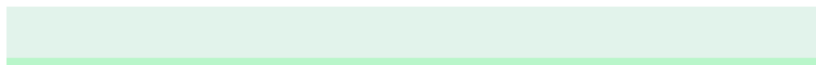
*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

You do not need to be a BI professional or expert in order to use our BI solution.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree



Section 3: Management of data source or data generating systems

Please provide your assessment for each of the following statements with respect to management of data source/generating systems for your BI solution.

Our data is managed through the use of On-line Transaction Processing systems such as ERP, CRM, Sales, etc.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

Data source systems are managed internally.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

Data source systems are managed externally by our service provider.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

Data from our source systems is stored in a simple database.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

Data from our source systems is stored in a complex data warehouse.

*

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

Data is accessed from a central repository at our office.

*

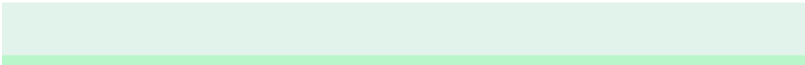
	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

Data is easily accessible from the data repository residing at our service provider. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

Our data goes through a data cleansing process (Extract, Transform, Load) before we gain access to it. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree



Section 4: Location and management of BI solution

Please provide your assessment for each of the following statements with respect to location and management of your BI solution.

Our BI solution is installed and run on computers located in our building. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

Our BI solution is installed and run on servers located at our service provider. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

We own all the hardware and software used to support our BI solution. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

We have full control of our in-house data processing applications and development environment. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

We have full control of our remote data processing applications and development environment. *

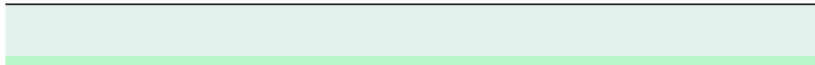
	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

We have full control of our storage and computing resources that reside at a remote data centre. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree

We have full control of our storage and computing resources that reside at our office. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree



Section 5: Buying or licensing options

Please provide your assessment for each of the following statements with respect to buying or licensing options for your BI solution.

We use a free version of any available BI applications. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

Our BI solution was an off-the-shelf purchased. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Strongly Disagree

Our BI user account is a Pay-as-you-go. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

Our BI user account is a Pay-per-user. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

Our BI user account is a Pay-per-use. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

Our BI user account is a Subscription-based. *

	1	2	3	4	5	
Strongly Agree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Strongly Disagree

This content is neither created nor endorsed by Google.

Google Forms

APPENDIX 4: LETTER FROM EDITOR

Prof EK Ngassam
Department of Computer Science
University of South Africa

14 January 2017

Re: EDITING OF DISSERTATION BY MISHECK BANDA (98961)

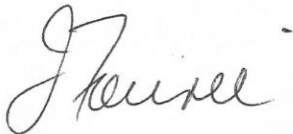
Dear Sir

This is to confirm that I edited and proofread the dissertation entitled:

'A data management and analytic model for business intelligence applications'

by the abovementioned student of your department.

Yours faithfully



Dr JA Fourie

Member of the Professional Editors' Guild

APPENDIX 5: TURNITIN SUMMARY REPORT

Secure | https://turnitin.com/s_class_portfolio.asp?r=41.4675386510105&svr=301&lang=e

M BANDA User Info Messages Student English Help Logout

turnitin

Class Portfolio Peer Review My Grades Discussion Calendar

NOW VIEWING: HOME > M&D STUDENT SUBMISSIONS 2017: KETCHA NGASSAM, E

Welcome to your new class homepage! From the class homepage you can see all your assignments for your class, view additional assignment information, submit your work, and access feedback for your papers. ✕

Hover on any item in the class homepage for more information.

Class Homepage

This is your class homepage. To submit to an assignment click on the "Submit" button to the right of the assignment name. If the Submit button is grayed out, no submissions can be made to the assignment. If resubmissions are allowed the submit button will read "Resubmit" after you make your first submission to the assignment. To view the paper you have submitted, click the "View" button. Once the assignment's post date has passed, you will also be able to view the feedback left on your paper by clicking the "View" button.

Assignment Inbox: M&D Student submissions 2017						
	Info	Dates			Similarity	
Research proposal	ⓘ	Start	10-Feb-2016	12:00AM		Submit View ↓
		Due	30-Jan-2018	11:59PM		
		Post	11-Feb-2016	12:00AM		
Chapters	ⓘ	Start	10-Feb-2016	12:00AM	18% ■	Resubmit View ↓
		Due	30-Jan-2018	11:59PM		
		Post	11-Feb-2016	12:00AM		
Final submission for examination	ⓘ	Start	10-Feb-2016	12:00AM		Submit View ↓
		Due	30-Jan-2018	11:59PM		
		Post	11-Feb-2016	12:00AM		