Guidelines for the migration of business intelligence in a developing organisation

A van der Walt

orcid.org 0000-0002-6406-2409

Dissertation accepted in fulfilment of the requirements for the degree *Master of Science in Computer Science* at the North-West University

Supervisor: Dr S Gilliland

Co-supervisor: Ms I Smit

Graduation: October 2019

Student number: 23972092

PREFACE

When I started working in the business intelligence (BI) field, within an organisational environment, one thing became very clear: due to the fast development of BI in all sectors of industry, most organisations do not have the knowledge, technical abilities or personnel capacity to implement and maintain a thorough BI solution. Furthermore, with production data growth, organisations sometimes neglect to also maintain their implemented BI solution. This study aims to provide organisations experiencing large amounts of data with guidelines as to how a thorough BI solution can be integrated into a growing organisation. It aims to provide information to organisations that are planning to migrate large amounts of data into their BI solution with minimal impact to organisational operations.

I would like to first acknowledge the North-West University for granting me the permission to conduct the study. Secondly, I would like to acknowledge the organisation used in this study for providing me with full access and authority to its premises, staff, databases and BI solution to conduct the study. Next, I would like to acknowledge Dr Sonja Gilliland and Mrs Imelda Smit from the North-West University for agreeing to be supervisor and co-supervisor to the study conducted. A researcher could not ask for more friendly, cooperative and insightful supervisors! I would also like to acknowledge my parents for always supporting me in my studies and providing me the aid to overcome all educational obstacles that I faced through the years. Last but not least, I would like to acknowledge God for providing me with such a great talent to be able to conduct this study. May this research prove that, with God in your corner, nothing is impossible.

ABSTRACT

Business intelligence (BI) is an analytical method used by organisations to extract production data, analyse it and present business users with meaningful insight into operations to support decision-making. The study aimed to provide guidelines for migrating a BI architecture into a growing organisation. Furthermore, the study aimed to provide means of integrating large amounts of data into the organisation's BI architecture with minimal impact on business operations.

Literature was reviewed on concepts such as BI, data warehouses and data migration in accordance with the study. An interpretivist research paradigm was used to conduct a descripto-explanatory case study of a single phenomenon. Qualitative data were collected from participants in the case study environment by means of semi-structured interviews. The data were analysed using ATLAS.ti and the results presented.

The results from data collected and analysed from participants concluded that the old BI solution presented several complications such as a poor quality of data, duplicated data and insufficient information presented to the organisation for decision-making purposes. Results further indicated that the data migration from the acquired organisation's database management system to the organisation used for the research's database management system, showed complications such as timeline scheduling and resource limitations. Results indicated that the new BI solution provided more accurate data, was cleaned from anomalies, and was easily accessible from one central location. In accordance with the literature reviewed, the case study method utilised, the researcher's interpretation of the organisation and data gathered, guidelines for the integration of BI architecture in a developing organisation were established. With the successful introduction of the BI solution, management was enabled to predict business decisions more accurately through the utilisation of the migrated data.

Keywords: Business intelligence, business intelligence architecture, data warehouse, data migration, analysis solution

TABLE OF CONTENTS

PREFACE ..... **I**
ABSTRACT ..... **II**

CHAPTER 1: CHAPTER FLOW **1**
CHAPTER 1 INTRODUCTION **2**
1.1 CONCEPTS KEY TO THE STUDY **2**
1.1.1 Business intelligence 3
1.1.2 Data warehouse 4
1.1.3 Data migration 7
1.2 PROBLEM STATEMENT **7**
1.3 OBJECTIVES OF THE STUDY **8**
1.3.1 Primary objectives 8
1.3.2 Theoretical objectives 8
1.3.3 Empirical objectives 9
1.4 RESEARCH DESIGN **9**
1.4.1 Paradigms 9
1.4.2 Qualitative strategies of inquiry 10
1.4.3 Methodologies 12
1.5 LITERATURE REVIEW **14**
1.6 EMPIRICAL STUDY **15**
1.6.1 Participant selection 15
1.6.2 Data collection methods 16
1.6.3 Data analysis methods 16
1.6.4 Rigour and evaluation of method 17
1.6.5 Limitations of the study 18
1.7 ETHICAL CONSIDERATIONS **18**
1.8 CHAPTER CLASSIFICATION **18**

CHAPTER 2: CHAPTER FLOW **20**
CHAPTER 2 RESEARCH DESIGN **21**
2.1 INTRODUCTION **21**
2.2 RESEARCH PARADIGMS **21**
2.2.1 Positivism 22
2.2.2 Critical social theory 22

2.2.3 Design science research ... 23
2.2.4 Interpretivism ... 24
2.3 QUALITATIVE STRATEGIES OF INQUIRY ... 26
2.3.1 Ethnography ... 26
2.3.2 Action research ... 27
2.3.3 Grounded theory ... 28
2.3.4 Case study research ... 30
2.4 DATA COLLECTION ... 33
2.4.1 Quantitative research ... 33
2.4.2 Qualitative research ... 34
2.5 METHODS ... 35
2.5.1 Document analysis ... 35
2.5.2 Surveys ... 36
2.5.3 Questionnaires ... 36
2.5.4 Focus groups ... 37
2.5.5 Observations ... 38
2.5.6 Interviews ... 39
2.6 DATA ANALYSIS ... 40
2.7 STUDY PLAN ... 42
2.8 CONCLUSION ... 46

CHAPTER 3: CHAPTER FLOW ... 48
CHAPTER 3 LITERATURE REVIEW ... 49
3.1 INTRODUCTION ... 49
3.2 DATA WAREHOUSE ... 50
3.2.1 Data warehouse data flow ... 50
3.2.2 Data warehouse maturity model ... 51
3.2.3 Extract, transform and load flow ... 53
3.2.4 Dimensional modelling ... 55
3.2.5 The Kimball methodology ... 57
3.2.6 Data warehouse automation ... 59
3.3 BUSINESS INTELLIGENCE ... 60
3.3.1 Business intelligence evolution ... 60
3.3.2 Business intelligence maturity model ... 62
3.3.3 Business intelligence lifecycle ... 63
3.3.4 Analysis services ... 66
3.4 DATA MINING ... 68

3.4.1 Data scientist 68
3.4.2 Data mining lifecycle 70
3.4.3 Data mining phases 71
3.5 CONCLUSION 74

CHAPTER 4: CHAPTER FLOW 75
CHAPTER 4 CASE STUDY 76
4.1 INTRODUCTION 76
4.2 CASE STUDY ENVIRONMENT 76
4.3 BUSINESS INTELLIGENCE ARCHITECTURE 77
4.3.1 Initial business intelligence architecture 79
4.3.1.1 Initial server structure 79
4.3.1.2 Initial reporting solution 80
4.3.2 Business intelligence development project plan 81
4.3.3 Current business intelligence architecture 82
4.3.3.1 Business intelligence server structure implemented 82
4.3.3.2 Data warehouse implemented 83
4.3.3.3 Analysis solution implemented 84
4.3.3.4 Reporting structure implemented 85
4.3.3.5 Automation process implemented 87
4.4 DATA MIGRATION 89
4.5 CONCLUSION 93

CHAPTER 5: CHAPTER FLOW 94
CHAPTER 5 RESEARCH IMPLEMENTATION, RESULTS AND FINDINGS 95
5.1 INTRODUCTION 95
5.2 PARTICIPANT SELECTION 95
5.2.1 Senior management 95
5.2.2 Technical staff 96
5.2.3 Data capturers 96
5.3 DATA COLLECTION 96
5.3.1 Semi-structured interviews 97
5.4 DATA ANALYSIS 97
5.5 RESULTS OF THE RESEARCH 103
5.5.1 Senior management participant results 103
5.5.2 Technical staff participant results 106
5.5.3 Data capture participant outcomes 110

5.5.4 Conclusion of the results ... 112
5.5.4.1 Senior management participants ... 112
5.5.4.2 Technical staff participants ... 113
5.5.4.3 Data capture participants ... 114
5.6 GUIDELINES FOR BUSINESS INTELLIGENCE MIGRATION IN A DEVELOPING ORGANISATION ... 114
5.6.1 Business intelligence architecture design ... 115
5.6.2 Data warehouse design ... 116
5.6.3 Analysis design ... 118
5.6.4 Reporting design ... 120
5.6.5 Business intelligence automation design ... 121
5.7 CONCLUSION ... 122

CHAPTER 6: CHAPTER FLOW ... 124
CHAPTER 6 CONCLUSION AND RECOMMENDATIONS ... 125
6.1 INTRODUCTION ... 125
6.2 STUDY AND DISSERTATION SUMMARY ... 125
6.3 OBJECTIVES ACHIEVED ... 127
6.3.1 Primary objectives ... 128
6.3.2 Theoretical objectives ... 129
6.3.3 Empirical objectives ... 130
6.4 THEORETICAL AND PRACTICAL CONTRIBUTIONS ... 131
6.5 RIGOUR OF QUALITATIVE RESEARCH ... 132
6.6 RECOMMENDATIONS FOR FURTHER RESEARCH ... 134
6.7 LIMITATIONS OF THE STUDY ... 135
6.8 CONCLUSION ... 135
BIBLIOGRAPHY ... 136
APPENDIX A: ANALYSIS DIMENSION STRUCTURE ... 148
APPENDIX B: ANALYSIS TABLE STRUCTURE ... 149
APPENDIX C: SENIOR MANAGEMENT BUSINESS INTELLIGENCE ARCHITECHTURE INTERVIEW QUESTION LIST ... 150
APPENDIX D: SENIOR MANAGEMENT DATA MIGRATION INTERVIEW QUESTION LIST ... 152
APPENDIX E: TECHNICAL STAFF BUSINESS INTELLIGENCE ARCHITECHTURE INTERVIEW QUESTION LIST ... 154
APPENDIX F: TECHNICAL STAFF DATA MIGRATION INTERVIEW QUESTION LIST ... 156
APPENDIX G: DATA CAPTURER INTERVIEW QUESTION LIST ... 158

APPENDIX H: SENIOR MANAGEMENT PARTICIPANT FULL CODE NETWORK 159
APPENDIX I: TECHNICAL STAFF PARTICIPANT FULL CODE NETWORK 160
APPENDIX J: DATA CAPTURER PARTICIPANT FULL CODE NETWORK 161
APPENDIX K: RIGOUR 162

LIST OF TABLES

Table 2-1: Guidelines and principles directing the rigour of this study 46

Table 4-1: Management – Old BI solution 80

Table 4-2: Management – New BI solution..... 84

Table 4-3: Developer – Data Migration Preparation..... 92

Table 4-4: Developer – Data Migration..... 92

Table 5-1: Theme identified: Old BI..... 99

Table 5-2: Theme identified: Old reporting 99

Table 5-3: Theme identified: New BI 99

Table 5-4: Theme identified: New reporting..... 101

Table 5-5: Theme identified: Data migration..... 101

Table 5-6: Management – Old business intelligence solution that was implemented... 103

Table 5-7: Management – Reason for new business intelligence solution implementation..... 104

Table 5-8: Management – Contributions of the new business intelligence solution 104

Table 5-9: Management – Complications with the data migration project..... 105

Table 5-10: Management – Stakeholder impression regarding the data migration into the organisation's business intelligence solution 106

Table 5-11: Management – Stakeholder impression surrounding the business intelligence solution)..... 106

Table 6-1: Guidelines and principles of rigour adhered to by this study..... 132

LIST OF FIGURES

Figure 1-1: Implementation phases of a BI architecture as adapted from Sirin and Karacan (2017:210) 4

Figure 1-2: Kimball data warehouse development lifecycle as adapted from Kimball and Ross (2015:33) 6

Figure 2-1: Action research phases (Coghlan & Brannick, 2019:9) 27

Figure 2-2: Grounded theory process (Silverman, 2016:349) 29

Figure 2-3: Grounded theory methods (Birks & Mills, 2015:13) 30

Figure 2-4: Parts needed for a case study research (Thomas, 2015:16) 31

Figure 2-5: Features of designing a case study research (Thomas, 2015:26) 32

Figure 2-6: Features of designing a case study research (Thomas, 2015:26) 42

Figure 2-7: Study plan diagram 45

Figure 3-1: Enterprise DW as adapted from Sherman (2014:69) 51

Figure 3-2: Data warehouse maturity capability model as adapted from Spruit and Sacu (2015:1512) 52

Figure 3-3: Extract, transform and load process flow as adapted from Andersen *et al.* (2018:5) 53

Figure 3-4: Extract, transform and load process flow as adapted from Meehan *et al.* (2017:2) 54

Figure 3-5: Bus matrix as adapted from Kimball *et al.* (2015:16) 56

Figure 3-6: Star schema as adapted from Iyer and Lakhtaria (2017:64) 57

Figure 3-7: Kimball lifecycle as adapted from Kimball and Ross (2015:33) 58

Figure 3-8: Business intelligence conceptual model as adapted from Stockdale *et al.* (2015:263) 62

Figure 3-9: Business intelligence maturity model as adapted from Tavallaei *et al.* (2015:1015) 63

Figure 3-10: Data flow in a simple BI architecture as adapted from Sirin and Karacan (2017:210) 64

Figure 3-11: Business intelligence lifecycle as adapted from Stefanovic (2015:916) 64

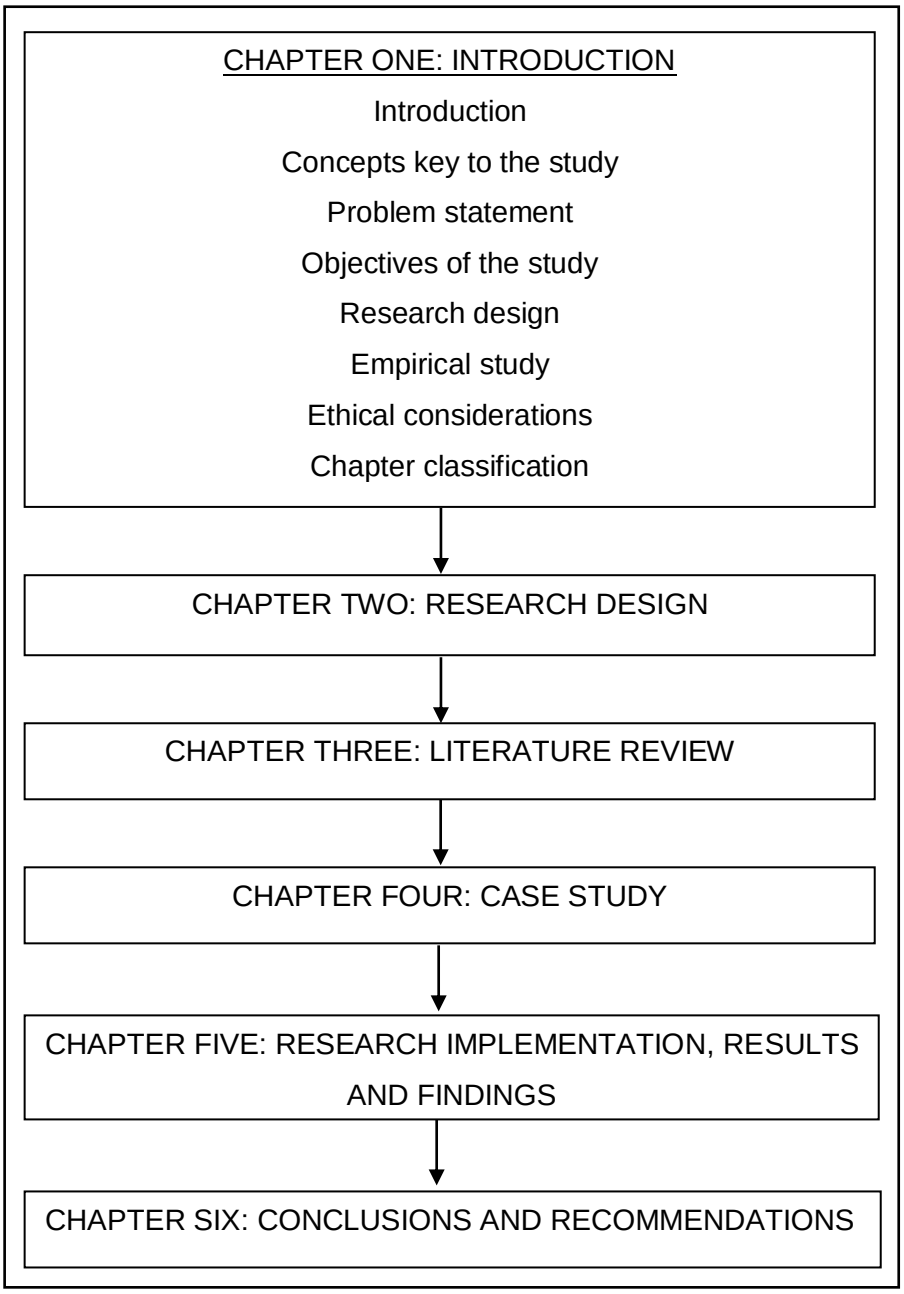
Figure 3-12: Business intelligence lifecycle as adapted from Łęgowik-Świącik *et al.* (2016:5) 65

Figure 3-13: Multidimensional cube (Rad, 2014:1) 67

Figure 3-14: MDX coding example (Rad, 2014:2) 67
Figure 3-15: DAX coding example (Cuevas, 2017:3) 68
Figure 3-16: Data scientist work flow as adapted from Agasisti and Bowers (2017:189) 70
Figure 3-17: Data mining lifecycle as adapted from Witten *et al.* (2016:29) 71
Figure 3-18: Data mining phases as adapted from Stefanovic (2015:918) 72
Figure 3-19: Phases of data mining as adapted from Rokach and Maimon (2008:3) 73
Figure 3-20: Data mining phases as adapted from Rokach and Maimon (2008:5) 73
Figure 4-1: Case study event flow 78
Figure 4-2: Old server structure 79
Figure 4-3: Business intelligence architecture Gantt chart 81
Figure 4-4: New server structure 82
Figure 4-5: New data warehouse data flow structure 83
Figure 4-6: Analysis data flow 85
Figure 4-7: Reporting data flow structure 87
Figure 4-8: Automation data flow 89
Figure 4-9: Data migration plan 91
Figure 4-10: Data migration project Gantt chart 91
Figure 5-1: Technical staff total number of comments regarding the old business intelligence solution 107
Figure 5-2: Technical staff total number of comments regarding the new business intelligence solution 108
Figure 5-3: Technical staff total number of comments regarding the data migration project 109
Figure 5-4: Data capturer total number of comments regarding the old business intelligence solution 111
Figure 5-5: Data capturer total number of comments regarding the new business intelligence solution 112
Figure 5-6: Business intelligence architecture guideline 116
Figure 5-7: Data warehouse structure guideline 117
Figure 5-8: Analysis structure guideline 120
Figure 5-9: Business intelligence automation guideline 122

Figure 6-1: Case study event flow 128
Figure 6-2: Business intelligence architecture design guidelines 129

CHAPTER 1: CHAPTER FLOW



CHAPTER 1 INTRODUCTION

In today's business environment, organisations have become more and more reliant on business intelligence (BI) to accurately measure operational performance. Business intelligence is gathered from a BI solution, which is supported by a developed BI architecture. A BI solution is defined as goal-driven techniques or practices used to analyse production data sets and present an application or method to the organisation for insightful decision-making (Wazurkar *et al.*, 2017:367). A BI architecture is a set of tools and processes needed for a BI solution to function accurately, such as a central data warehouse (DW), analysis solution and reporting solution (Larson & Chang, 2016:700).

Business intelligence enables management to accurately predict business decisions based on reliable operational data. This has resulted in greater emphasis being placed on data integrity and visibility throughout organisations (Huang *et al.*, 2017:1230). Business intelligence solutions allow for the reliable and accurate analysis of day-to-day operations, resulting in profitable and successful business outcomes. Giménez-Figueroa *et al.* (2018:115) describe BI solutions as an increasingly important aspect of decision-making that an organisation can undertake to gain an advantage and surpass competition. The use of BI also supports decision-making and, as such, BI plays a vital role in the growth of an organisation. Olszak (2014:116) states that introducing an integrated BI architecture into an organisation can result in effective decision-making operations for distributed organisations. However, as an organisation matures, the process of maintaining these systems becomes more challenging due to the volume and complexity of the data.

The aim of this study is to investigate, analyse and report on the way a BI architecture may be integrated into a growing organisation. Furthermore, the objectives, timelines, challenges and accomplishments during the transformation of the BI architecture in an expanding organisational environment are discussed. The study will also investigate and discuss the results of data integrated from two organisations' DBMSs into one and the impact of the additional data on the organisation's new BI solution. The outcome of the study may inform entities in need of improving a BI architecture or those that are planning to implement data migration from one database management system (DBMS) to another.

1.1 CONCEPTS KEY TO THE STUDY

Key concepts that are discussed in this study concern the investigation, development and analysis of a BI architecture being integrated into a growing organisation environment. These

concepts include BI, DWs and their development, and the migration of data into the BI architecture.

1.1.1 Business intelligence

Business intelligence forms the data hub of any organisation, and it is because of BI that important operational business decisions can be made with consideration of up-to-date operational data (Wieder & Ossimitz, 2015:1164). To develop a comprehensive BI solution, a BI architecture must be designed and thoroughly constructed to ensure accurate data flow and minimise data anomalies. The BI architecture portrays the environment in which all components associated with the BI solution work together such as the extract, transform and load (ETL) process, DW, analysis solution, reporting solution and automation process.

The decision to implement a BI architecture could be considered either by an organisation seeking to find new decision-making possibilities from analysing operational data, or to overcome a problematic business operational situation. To accomplish this, implementation of a BI architecture comprising three phases should be completed. Firstly, a well-structured data ETL process should integrate data from multiple sources into one central location. Secondly, a comprehensive data analysis is performed on data extracts from the DW to provide a solution aligned with business rules associated with organisational objectives. Lastly, a reporting solution based on problem solving, which supports managerial decision-making, is enriched (Kimball & Ross, 2015:3; Sale *et al.*, 2018:16018; Soon & Fraser, 2011:190; Wang *et al.*, 2018:3). The end result of a well-developed BI architecture may be optimised business performance and profitability by decision-making through trustworthy and comprehensive data analytics done on operational data.

The BI architecture suggested by Sirin and Karacan (2017:210), as seen in Fig. 1-1, is a more detailed variation of the three BI architecture phases portrayed by (Kimball & Ross, 2015:3; Sale *et al.*, 2018:16018; Soon & Fraser, 2011:190; Wang *et al.*, 2018:3). Phases 1 and 2 of this architecture correspond to the ETL process, phases 3 and 4 correspond to the comprehensive data analysis and phase 5 reflects the reporting solution. The phases are explained below:

Phase 1 - Source data are identified from numerous data sources (this may include various DBMSs, Excel files and text files).

Phase 2 - The data identified in the source data stage are placed through an ETL process where the data are cleaned of all anomalies as well as irrelevant and corrupted data. The processed data are then loaded into a singular DBMS instance.

Phase 3 - The DW stage entails the upload of data into a DW, where the data are prepared to be analysed so that business rules may be developed for future application.

Phase 4 - A sample of the data in the DW is then extracted and analysed to identify whether the data are concrete and clean. Through the extracted sample data, meaningful knowledge can be gained from the information in the DW.

Phase 5 - The data from the DW is analysed and presented to the organisation via statistical reports and presentations in order to assist management with meaningful business decision-making.

Figure 1-1: Implementation phases of a BI architecture as adapted from Sirin and Karacan (2017:210)

Distinctively, the explanation of the process to implement a BI architecture as described by Kimball and Ross (2015:3); Sale *et al.* (2018:16018); Soon and Fraser (2011:190); Wang *et al.* (2018:3) differs slightly from the depiction of Sirin and Karacan (2017:210), as seen in Fig. 1-1. The description of the process as portrayed by the latter includes data mining and machine learning techniques.

1.1.2 Data warehouse

A well-developed DW provides a central location of accurate and easy to maintain data, accumulated through the ETL process from various sources. A BI solution can retrieve data from the DW with ease and minimal complexity.

A DW can be described as a repository of data collected from heterogeneous sources. The data are cleaned and developed to find one consolidated result of truth regarding organisational information, and is then stored in a central structured database (Iyer & Lakhtaria, 2017:61). The DW lifecycle is focused around the design, development and implementation of a DW and the data flowing through the DW. In an organisation, the DW objectives and sub-objectives are set to correspond with business objectives and sub-objectives. Inmon and Linstedt (2014:97) propose that a DW should be designed and developed starting with the main objective towards the sub-objectives. Kimball and Ross (2015:32) contradict them by stating that the DW lifecycle should first focus on the most detailed level, beginning by identifying the most important

business process, then identifying the business rules and dimensional data, and lastly identifying the factual data. The organisation investigated in the study made use of the Kimball methodology of designing and developing a DW by addressing the sub-objectives first. For this reason, the literature review of the study focuses on the Kimball methodology of DW development.

The Kimball DW lifecycle of development as described by Kimball and Ross (2015:33) is widely used around the world for its simplicity and detailed phases to design and develop a DW solution. Kimball and Ross (2015:33) visually express that a thoroughly designed DW, as seen in Fig. 1-2, contains the following phases:

Phase 1 - The programme and/or project plan to design a DW is drafted by the project manager and approved by the project owner and stakeholders.

Phase 2 - All business requirements needed from the organisation are gathered by the project manager and integrated into the DW solution.

Phase 3 - The BI architect executes three sections of the BI development lifecycle concurrently, each with its own respective steps as mentioned below:

Section 1: Technical architecture design

Step 1: All technical components of the DW architecture are identified.

Step 2: The needed products associated with the design and development of the DW are identified.

Section 2: Dimensional modelling

Step 1: The DW dimensional model is created. The dimensional model contains the dimensions and facts as prescribed by the business requirement definitions.

Step 2: The physical design of the DW dimensional model is drafted.

Step 3: The ETL process plans how data extracted from the DW will be cleaned and transformed into meaningful data that can be used by the DW.

Section 3: BI application design

Step 1: The BI architect develops an appropriate BI application design by referencing the business requirements definition.

Step 2: The BI application is developed as designed by the BI architect.

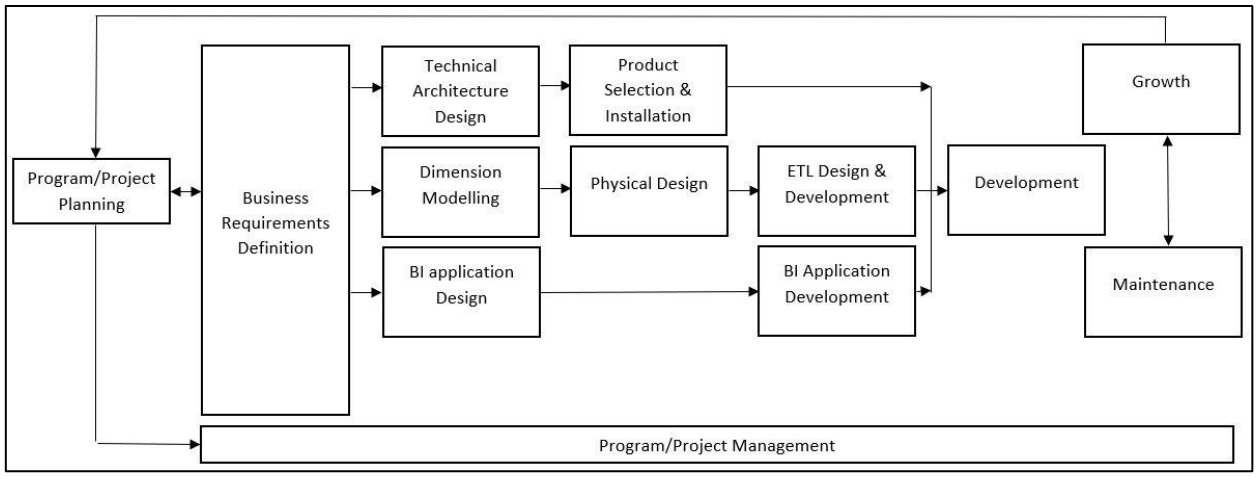


Figure 1-2: Kimball data warehouse development lifecycle as adapted from Kimball and Ross (2015:33)

Phase 4 - The BI architect works with a BI developer to deploy the DW into the required environment.

Phase 5 - When the DW solution is in operation, the BI developer does continuous maintenance and DW maturity inspections to ensure the DW solution performs as intended.

Phase 6 - This phase is optional and only occurs when further development is needed after implementation of the DW solution has been completed. In this phase, the project manager investigates the DW design and drafts a project management plan to implement changes to the currently implemented DW. Thereafter, the DW lifecycle is reiterated.

After a DW is implemented, the next logical step for an organisation to consider is to populate the DW by integrating data from multiple systems. A database management system can be defined as software that allows users to access and manage a system of databases by either a graphical user interface (GUI) or by structured query language (SQL) commands (Fredstam & Johansson, 2019:4). Velimeneti (2016:10) states that, when migrating data from one system to another, it is important to answer questions such as:

- What (information) is needed from the organisation to integrate databases?
- How will the migration benefit the organisation?
- How can the organisation guarantee that only high-quality data will be migrated?
- What is the percentage of downtime the organisation will experience during data migration?

By following a complete data migration plan as suggested in the following section, all of the questions as described will be answered.

1.1.3 Data migration

Moving data from one DBMS platform to another can be a complex task, especially in an organisation that does not use a well-structured DBMS. Data imperfections, security constraints and user knowledge are just a few of the challenges a database administrator (DBA) can expect when migrating data from one DBMS platform to another (Mateen & Ali, 2017:96). The steps to design a data migration plan vary from organisation to organisation, according to the data needs and rules related to the data migration (Kumar *et al.*, 2015:1).

When an organisation is tasked to migrate data from one DBMS platform to another, there are three types of data migration (Andersen *et al.*, 2018:5):

- Migration of updated facts. Data from existing factual sources are migrated to an existing fact table already in place.
- Migration of new facts. New facts are introduced into the fact tables to produce new meaning in the data.
- Migration of deleted facts. Deleted facts are migrated into the historical fact tables for historical archive purposes.

These three types of data migration mentioned by Andersen *et al.* (2018:5) can be implemented according to the requirements of the organisation.

1.2 PROBLEM STATEMENT

The organisation in which the research was conducted was started early in the 21st century, and quickly grew to become one of the largest competitors in the value added tax (VAT) recovery industry. The organisation is now providing VAT recovery services to more than 16 000 clients worldwide. Because of the organisation's growth and success in its ventures, it was possible for the organisation to acquire divisions of its direct competition, resulting in the organisation becoming one of the largest global leaders in the VAT recovery industry. Along with the increasing growth in the organisation's client base, came exponential growth in its data. The organisation soon realised that exponential data consumption would occur, and exclusive BI development was needed to support its growing business endeavour.

Due to business decisions being made through statistical reports, BI forms an important part of any medium to corporate size organisation (Yasser & Zota, 2016:35). There are, however, reasons why organisations fail to fully implement BI solutions or successfully maintain BI

solutions as they mature. A thoroughly designed BI architecture is necessary in any organisation and, as an organisation grows, the implemented BI architecture needs to be maintained and developed, as well as accepted by the business managers and stakeholders. Also, most of the time when an organisation implements a BI architecture with minimal or no BI implementation planning, more resources are used than anticipated, resulting in higher than expected cost implications, as well as business users' demands and expectations for relevant and updated information not being met (Obeidat *et al.*, 2015:47).

This study investigates the process followed to successfully integrate BI into an organisation experiencing exponential growth. The study focuses on challenges such as upgrading the current BI solution, migrating data from one organisation to another with exponential data growth, and the impact of the newly developed BI architecture on users' utilisation of organisational data and reports.

The emphasis of the study is to compile guidelines to support the implementation of BI into growing organisations, or further BI development. Furthermore, the study provides encouragement for organisations planning to perform data migration from one DBMS platform to another.

1.3 OBJECTIVES OF THE STUDY

The study uses the term BI architecture, as it refers to the strategies and technology driven techniques (such as data warehousing, analysis services and reporting solution) to transform production data into meaningful information that the organisation can use to derive business decisions from (Larson & Chang, 2016:705).

The following objectives have been formulated for the study:

1.3.1 Primary objectives

The primary objective of the study is to compile guidelines to assist developing organisations to overcome challenges of data growth while successfully implementing a BI architecture.

1.3.2 Theoretical objectives

To achieve the primary objective, the following theoretical objectives have been formulated for the study:

T01: To perform a literature review in support of all concepts key to the research.

T02: To use the literature to describe research paradigms, designs and methods used for the research.

1.3.3 Empirical objectives

In accordance with the primary objective of the study, the following empirical objectives have been formulated:

E01: To describe the BI architecture of the small organisation in its first year of existence.
E02: To describe the newly implemented BI architecture.
E03: To determine and report how data from two different database platforms (Microsoft and Oracle) could be integrated into one BI architecture.
E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI architecture in the organisation.

1.4 RESEARCH DESIGN

In this section, a number of topics concerning research design are addressed, namely paradigms, qualitative strategies in the information systems (IS) field, methodology and methods.

1.4.1 Paradigms

Interpretivism was identified as the most appropriate research paradigm for the study since the empirical focus of the study was primarily to gather qualitative data by means of interviews to describe how a BI architecture was implemented into a growing organisation. Therefore, although four research paradigms, namely interpretivism, positivism, critical social theory (CST) and design science research (DSR), are relevant to the research field, the latter three paradigms are defined here for the purpose of thorough clarification of the research paradigms available in the IS field, while the former is discussed in more detail.

Positivism is a research paradigm designed with the goal of measuring. The positivist researcher perceives the study environment from an objective perspective, and research results are collected and analysed in a quantitative way (Orlikowski & Baroudi, 1991:5).

Critical social theory has as goal the deconstruction of reality in terms of oppressing structures in order to reconstruct the environment, excluding the oppressing structures, with the purpose of determining whether a feasible solution can be met (Myers, 1997:5).

Design science research is a research design with the goal to learn about a problem environment, to then develop an artefact with the sole purpose to change the affected participants' situation, which is impacted by the study environment. Design science research requires the researcher to consider users of an artefact and involve them in the development of the artefact (Simon, 1996:4).

Interpretivism is a research paradigm where the belief is that reality is socially constructed, and the goal is therefore to understand the study environment. Interpretive research design requires the researcher to be a learner. The researcher should be participative as well as subjective to conduct the needed research. The end result of interpretive research is the development of guidelines, rules or a theory (Myers, 2004:117). When performing interpretive research, the researcher usually gathers qualitative data from participants directly impacted by the study environment in order for the researcher to support the end result (Gibbons, 1987:5). In an interpretive study, the researcher aims to understand the views, experiences and perceptions through the eyes of the participant (Victor *et al.*, 2016:24). Interpretivism was suitable as the research paradigm for this study as its intention is to understand the context of the study environment. Semi-structured interviews were used as a qualitative data collection method to collect data from participants directly impacted by the study environment. The results have been used to compile guidelines for developing organisations to overcome challenges of data growth while successfully implementing a BI architecture.

1.4.2 Qualitative strategies of inquiry

The strategy of any study is simply the map that the research follows to conclude the research done. A well-planned strategy builds a foundation for a good study and notifies the reader of how the study came to its conclusion (Hofstee, 2006:107). Case study research was indicated as the most appropriate strategy for this study as a case study explores a natural single phenomenon that is considered as a 'case' by obtaining comprehensive knowledge through various methods (Collis & Hussey, 2013:340). A descriptive case study is a type of case study where the researcher uses a clear picture of the phenomena to collect data with the aim to accurately describe an event or phenomenon. An explanatory case study, on the other hand can be described as when a researcher studies a situation, aiming to provide a relationship between variables in a given study environment. When a descriptive and explanatory case study is combined, a descripto-explanatory type of case study is conducted (Saunders *et al.*, 2009:140).

According to Myers (1997:6), there are four strategies which are relevant to the IS field, namely case study, ethnography, grounded theory and action research. The latter three strategies,

defined here for the purpose of thorough clarification of all qualitative strategies associated with the IS field, are briefly discussed. The case study research strategy, however, is discussed in more detail.

Ethnography can be defined as a strategy used to study a cultural tradition and its practices by interpreting the social interactions of the study environment (Denscombe, 2014:4). In agreement, Hancock and Algozzine (2016:9) state that, in an ethnographic strategy of inquiry, the researcher aims to describe the beliefs and values on which languages and interactions in social and cultural groups are based. The ethnographic research strategy aims to gather quantitative data from the study environment (Hammersley & Atkinson, 2007:5).

Action research comprises two different focuses, namely "action" in terms of what needs to be done, and "research" by finding out what is necessary to complete the study by conducting research. Researching the event and then taking action to improve the study environment is the fundamental purpose of the action research strategy (McNiff, 2009:7). Action research is designed for and focuses on developing the organisational environment (Baskerville & Wood-Harper, 1996:235).

Grounded theory, as opposed to other research strategies, does not start with using other theories but rather extracts or composes a theory from the results of the chosen area of inquiry (Zikmund *et al.*, 2013:139). Theory evolves during the continuous actions of analysis and collection of data, where the researcher asks questions about the data to achieve deeper understanding and a better explanation (Strauss & Corbin, 1994:273). According to Myers (1997:9), the purpose of grounded theory research in business and management is to develop new concepts and theories of business-related phenomena, where these concepts and theories are firmly grounded in qualitative data.

Case study research can be defined as a strategy with the primary objective of understanding how a social setting with complex relationships operates (Denscombe, 2014:4). In agreement with Denscombe (2014:4), Yin (2011:5) states that an explanatory case study is regarded as an investigation into a specific set of occurring events with the intent to justify the events and provide guidelines as to how those events may be applied to a different situation. Furthermore, Hancock and Algozzine (2016:15) explain the different levels of case study research as follows:

Level 1: Case study research is focused on an individual's (researcher or observer) representation of how a phenomenon impacted a group of participants.

Level 2: A natural context is needed for the research method. The research method is also bound by time and space.

Level 3: Various and descriptive resources of information are needed for a thorough case study to be conducted.

This study was classified as a descripto-explanatory study which means that the study environment was described as a precursor to an explanation of the singular phenomenon that had been investigated with the purpose of building guidelines to be used in other contexts (Saunders *et al.*, 2007:140). This type of case study tends to gather qualitative information regarding the study environment.

1.4.3 Methodologies

Methodology can be defined as how a group of methods, integrated within a singular system, are applied within a desired field of interest (Kothari, 2004:7). Research methodologies provide researchers with means to find the answers to research questions asked (Kumar, 2019:4).

Principles of interpretive research will first be discussed followed by the methods of collecting qualitative data. According to Klein and Myers (1999:72), when conducting interpretive research, there are generally several evaluation principles to keep in mind. The principles and how they have been implemented in this study are listed:

- The hermeneutic circle principle is primarily a representation of the human understanding of independent sections or parts of meaning from participants in the study. Human understanding is fundamental to all the other principles listed below. The researcher needs to consider and portray the participants' emotional state with regard to the study environment.
- The principle of contextualisation requires the researcher to have a critical reflection of the historical and social background of the study environment in order to be able to understand how the current situation emerged. This principle requires that the researcher understands the BI architecture migration in its context, from its origin to its current state.
- The principle of interaction between the researcher and the participants is reflected in the structuring of the data received from the participants. The researcher needs to be able to relate to the participant while conducting the research to truly understand what is being described by the participant. In this study, the researcher was exposed to the study environment during the BI architecture evolution and data migration from one DBMS platform to another.

- The principle of abstraction and generalisation entails understanding the idiographic ideals relating to the social and cultural actions of the data collected from the participants. This principle needs to be correlated with the first two principles. The researcher needs to analyse data collected from participants by abstracting meaning in the context of the study, leading to the extraction of guidelines applicable in other contexts.
- The principle of dialogical reasoning revises the actual findings and research design, guided by sensitively identifying differences between preconceptions ("stories told through the data"). The researcher should be aware that the conclusion of the research design and research data, analysed from participants, might be different. In the case of this study, implementations regarding BI architecture and data migration might differ from the opinions of the participants directly impacted.
- The principle of multiple interpretations involves different perspectives received from different participants in the same study environment or sequences of events expressed in multiple stories or narratives. Participants will describe the study environment as they experience it. The researcher should be aware that the opinions of participants impacted by the same study environment might differ.
- The principle of suspicion requires the researcher to be sensitive to possible distortions or biases in data collected from the participants. The researcher should not only rely on the words of the participants during the interview but should also look at behaviour such as facial expressions and body language when questions are answered.

Qualitative research methods can be defined as techniques used to collect qualitative data from participants (Myers, 1997:2). Taylor *et al.* (2015:7) state that qualitative data can be defined as producing descriptive data through research. Researchers cannot possibly know which questions to ask or what data to retrieve until they have spent some time in the environment.

In this section, different types of qualitative research methods are defined, namely document analysis, questionnaires, focus groups, observations and interviews. Interviews were used as a qualitative data collection method in the study and are therefore discussed in more detail.

Document analysis is a form of data collection by which the researcher systematically evaluates documents, whether electronic or printed. This method, like other qualitative analysis methods, require the researcher to examine and interpret the documents to gather explicit meaning from them (Bowen, 2009:27).

Questionnaires are a data collection method where the researcher creates a form using predefined questions that a respondent needs to complete and return to the researcher after completion (Moser & Kalton, 2017:257).

Focus groups are when an individual researcher communicates with a group of participants at the same time. This data collection technique allows the researcher to have the participants discuss the study among themselves, providing the researcher with relevant aspects and ideas regarding the study and their grouped experiences (Parahoo, 2014:320).

Observations are a data collection technique used where the researcher observes the participants in a study environment while collecting data from the participants' actions, expressions and experiences inside the study environment (Jamshed, 2014:88). Observations usually occur in a naturally selected environment, as opposed to interviews, which are conducted in a designated area selected by the researcher (Merriam & Tisdell, 2015:137).

Interviews are the process during which a researcher, called the interviewer, collects information from a participant, formally known as the interviewee, regarding a specific study environment (Parahoo, 2014:308). According to Spradley (1979:3), to efficiently conduct an interview, the researcher needs to be able to learn from interviewees instead of studying them. Recording devices are used to keep track of all key points discussed during the interview. After the interview, the interviewer goes through the recordings of the interview to make sure that detailed data were collected from the participant (Jamshed, 2014:88). The first thing to consider when conducting an interview is the interview process to follow. The interview process allows the participant and researcher to exchange predefined dialogue. The responses from the participant is captured by the researcher, either on paper or by a recording device, then later reviewed and analysed. There are two types of interviews that can be conducted. The first type involves structured interviews that consist of the researcher constructing predefined questions. The second interview type, semi-structured interviews, allows the participant to answer the predefined questions in an open dialogue (Smith, 2015:30).

1.5 LITERATURE REVIEW

The contribution of a literature review in any study is to define current knowledge regarding a research topic or area. Furthermore, the literature review allows the researcher to formulate an argument towards a convincing thesis dispute by using logical arguments (Machi & McEvoy, 2016:1). Aveyard (2014:2) agrees and states that the literature may provide insight as to why the research in the particular study is needed.

The literature review for this study focuses on data sources such as relevant textbooks, journal articles and the internet. The literature review associated with the study primarily concentrates on literature regarding BI architecture, DW, and data migration. Academic resources such as the North-West University library academic catalogue and databases such as Google Scholar and ScienceDirect were used to gather literature for this study.

1.6 EMPIRICAL STUDY

The empirical study is discussed in more detail in this section. A case study can be classified as an empirical study when data regarding a real-life contemporary phenomenon is collected and examined, and the boundary between the context and the phenomenon is unclear (Myers, 1997:7). The descripto-explanatory case study to be discussed is the phenomenon of a BI architecture migration integrated into a growing organisational environment.

The empirical part of this study comprises the following methodological dimensions: participant selection and participants, data collection methods, data analysis methods, rigour and evaluation of method, and the limitations of the study.

1.6.1 Participant selection

Selecting the appropriate group of participants for this study was essential as the participants needed to have comprehensive knowledge of either the business itself or the technical implementation of the business. The technique of purposive sampling was used to identify participants for this study as the researcher already knew where the participants would be selected from (Sharma, 2017:749). The primary participants of this study were the following groups of people:

- The senior management team of the organisation, who could provide valuable information regarding the business. This group consisted of senior management associated with the study environment since the beginning of the organisation's existence. Furthermore, this group had the most insight into the organisation with reference to data growth and business decisions.
- The technical staff responsible for the BI architecture and the data integration between the two organisations would be able to provide data regarding the growth of data, the entire BI architecture, as well as the data migration project. This group had the most knowledge and understanding of the whole project, its costs and projections. This group of participants was also responsible for the data and database optimisation during the organisational growth

process. They had the most interaction with the database, data capturers and the enterprise resource planning (ERP) system responsible for the day-to-day functions.

- The data capturing staff that engaged with the ERP system on a regular basis. This group of participants had the most knowledge about day-to-day data capturing operations.

1.6.2 Data collection methods

To be able to collect comprehensive data from participants, a decision needs to be made about which qualitative data collection methods to use. According to Smith (2015:2), naturalistic verbal reports such as interviews, case studies and written accounts are generally used to collect qualitative data. In agreement, Silverman (2015:163) states that conducting interviews is the generally preferred way to gather qualitative information.

For this study, semi-structured interview questions based on the literature reviewed were used to collect data for the purpose of addressing the empirical objectives associated with the study. The data collected from all participants in the case study were analysed and used to compile guidelines for the migration of BI architecture in a developing organisation.

1.6.3 Data analysis methods

The data collected from participants needs to be analysed and transformed into meaningful information before a conclusion to the study can be reached. According to Hofstee (2006:117), data by itself holds little meaning until it is analysed and transformed to present meaningful information. To comprehensively analyse the data, the following methods are used:

- Content analysis. This is when behavioural and verbal data are categorised into interpretative and descriptive levels to be summarised or classified.
- Narrative analysis. This involves that, for every interview or all data collected, the researcher identifies what is meant by the participants and how to present the data collected in the best way possible. Different stories from different participants need to be reformulated to present the different experiences of the various participants in the same manner.
- Discourse analysis. This is where the researcher needs to study the way participants express themselves during the interview or case study, and to interpret the resulting data accordingly.

For the purpose of this study, content analysis as mentioned by Hofstee (2006:117) were applicable and were implemented.

According to Elo *et al.* (2014:1), there are four phases to qualitative analysis that need to be adhered to when conducting a qualitative analysis:

- Data organisation: The recorded interview data were transcribed into text.
- Framework identification: The research actions were decided as soon as ethical clearance for the study was received.
- Data sorting: Interview data were transcribed and analysed using ATLAS.ti. Codes were defined according to the empirical objectives of the research.
- Descriptive analysis: Codes were categorised into themes and recurrent themes were identified.

The literature was used to guide the themes identified in the empirical study and the themes were compared with information related to the research objectives.

1.6.4 Rigour and evaluation of method

The last step for the researcher to consider before a thoroughly developed conclusion can be reached is to evaluate the research method used to conduct the study.

Detailed understanding of some phenomena in a social connected environment can be examined by researchers conducting research by means of a qualitative methodology (Myers, 1997:2; Zikmund *et al.*, 2013:132). The interpretive research principles of Klein and Myers (1999:72) guided the research and were adhered to. According to Neergaard *et al.* (2009:3), the following criteria should be met for a thoroughly rigorous study to be conducted:

- Authenticity: The participants must be able to speak freely, their voices must be heard when spoken, and an accurate perception must be presented.
- Credibility: A true researcher's perspective must be captured and portrayed.
- Integrity: Bias must be reflected, informants must be validated, and peer review of data collected and analysed must be conducted.

Brannen (2017:5), in agreement with Myers (1997:2), states that qualitative research gathers data not in number format or as statistical data, but from the viewpoint of a participant involved in the study. In this approach, the emotional state of the participant in the study environment is taken into consideration. It is therefore important for qualitative research to be trustworthy and plausible. There are measures available to ensure this, which include ethical conduct of the research and the use of qualitative data analysis tools.

After the data had been collected from participants following the interpretive research principles as described by Klein and Myers (1999:72), the researcher interpreted the data collected and memos compiled from the interview using ATLAS.ti, a computer application.

1.6.5 Limitations of the study

The first limitation is that the data policies, server structures and documented processes of other organisations might differ from the organisation used in this study. A second limitation of the study is that some of the key personnel, that was involved with the organisation's BI architecture's transformation process have left the organisation since the transformation process took place. Lastly, the case study was performed in one organisation and the guidelines developed are therefore associated with a single organisation's experience of implementing a BI architecture and conducting data migration from one DBMS platform to another.

1.7 ETHICAL CONSIDERATIONS

When conducting a qualitative study, the researcher needs to take into consideration the fact that human participants are being used. Participants need to give permission to the researcher before any data can be collected. The researcher should also make it known to the participants that no information provided by them will be changed (Munhall, 1988:150).

The ethical considerations for this study included first obtaining ethical clearance from the North-West University to perform the research. In addition, permission from the management of the organisation used in the study was necessary. Lastly, the researcher needed to get informed consent from every senior manager, technical staff and data capturers who participated in this study. Names of the participants and the organisation are not disclosed, and data have been treated as confidential.

1.8 CHAPTER CLASSIFICATION

This study comprises the following chapters:

Chapter 1 – Introduction

This chapter introduced the study and provided the reader with the background of the research environment, research methodology to be used and literature review. This chapter included sections on key concepts to the study, problem statement, objectives of the study, research design, literature review, empirical study and ethical considerations.

Chapter 2 – Research design

The different paradigms applicable to IS, as well as the corresponding methodologies and methods, are discussed in this chapter. The research design for this study is described and justified, and the chapter includes sections on research paradigms, qualitative strategies of inquiry, data collection, methods, data analysis and study plan.

Chapter 3 – Literature review

This chapter describes current BI architecture literature, with a focus on research conducted on BI architecture in growing organisations. Literature about DW, BI architecture, and data mining is also reviewed.

Chapter 4 – Case study

This chapter describes the organisation's database platform structure and the BI architecture implementation when the organisation started. Thereafter, the way the current BI solution was implemented is defined. Lastly, the manner that data from two organisations were integrated into the organisation's current BI architecture is described.

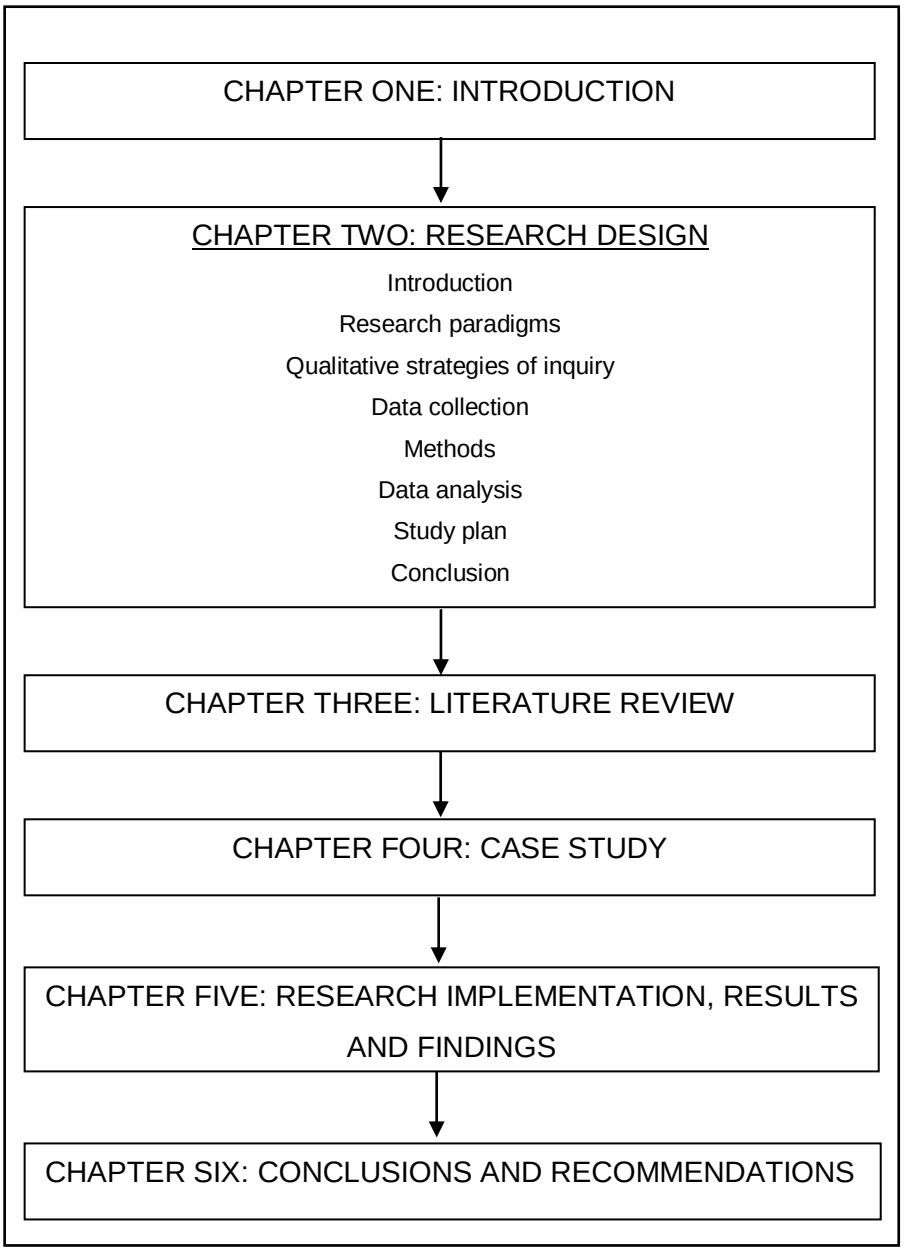
Chapter 5 – Research implementation, results and findings

This chapter describes the participants selected for the study, the manner in which all data were collected from the participants and how the data obtained from participants was analysed. Furthermore, the results of the research are presented and guidelines are formulated.

Chapter 6 – Conclusions and recommendations

A summary of the study and dissertation, how the objectives were achieved, theoretical and practical contributions of the study, recommendations for further research, and lastly limitations of the study are described in this chapter.

CHAPTER 2: CHAPTER FLOW



CHAPTER 2 RESEARCH DESIGN

2.1 INTRODUCTION

The main objective of the study is to compile guidelines to assist developing organisations to overcome challenges of data growth while successfully implementing a BI architecture. The aim of this chapter is to distinguish between different research methodologies available in the information systems (IS) research environment and their respective research frameworks. Furthermore, the chapter aims to determine the most appropriate research methodology and framework for this study (as stated by theoretical objective T01). Lastly, the chapter aims to describe a study plan depicting how the study was designed, providing the flow of chapters and topics discussed in each chapter.

There are different paths to thorough research, and each option is unique. Before conducting research, the researcher needs to select an appropriate research technique suited for the study environment (Denscombe, 2014:3). Research design is the researcher's plan regarding 'how', 'when' and 'where' the researcher will collect data from a study environment, analyse the data and draw a conclusion regarding the study findings. To have a thoroughly conducted research design, the researcher needs to have an appropriate approach to the research, a method of collecting data, relevant study environment, source to the study, time to conduct the study and a data analysis method (Parahoo, 2014:165). In agreement, Rudestam and Newton (2014:1) state that a research methodology can be seen as a researcher's plan regarding the 'why', 'what' and 'how' of conducting the study.

In the subsequent sections, the following topics are addressed: the different research paradigms related to the IS environment (section 2.2); the different qualitative strategies of inquiry regarding the IS environment (section 2.3); the data collection method and the most appropriate data collection technique for the research being conducted (section 2.4); and the different methods mostly used in the IS research environment (section 2.5). The data analysis phase following data collection is also discussed (section 2.6), an appropriate study plan for this study is presented (section 2.7), and lastly, a conclusion to the chapter is given (section 2.8).

2.2 RESEARCH PARADIGMS

This section first discusses the most-used research paradigms relevant to IS, followed by a discussion regarding the selected research paradigm for the study and the reason why this paradigm was suitable for the research.

When a researcher conducts a study, the most relevant research paradigm needs to be selected in order to depict the views and beliefs of the researcher regarding the study in the most appropriate manner possible. It is also important for the researcher to have willing participants for the study (Parahoo, 2014:165). Hancock and Algozzine (2016:4) state that research involves the determination of what one wants to study, the importance of why 'it' needs to be studied, how 'it' is going to be studied and how to acquire information in the best possible way to support the study. Additionally, the type of information and how the information is gathered, analysed and interpreted are important. Lastly, it must be decided who will conduct the research and how the findings will be verified and shared.

2.2.1 Positivism

The first research paradigm commonly used in the IS field is positivism. Positivism was described by Bacon (1561-1626), contributed to by Descartes (1596-1650) and further developed by Comte (1798-1857). Positivist research tends to measure things in the study environment by reducing the problem in the study to a measurable size and formulating hypotheses. Data are collected from respondents that are directly impacted by the study environment, to try and prove the hypotheses (Straub *et al.*, 2004:381). Instead of identifying essences and enhanced vision, positivist research aims to summarise observations (Lunt, 2014:48).

In social sciences, positivism involves a scientific approach to study an independent physical reality using facts and figures, with a focus on the causes and consequences of a specified problem (Denscombe, 2014:2). When conducting positivist research, the researcher should be objective. Experiments and surveys are common methodologies associated with positivism, collecting data mainly from questionnaires (Kitchener, 2004:38; O'Brien, 1998:1). Positivist research is normally conducted when the researcher collects quantitative data from a large population of respondents.

While positivism is an important research paradigm in IS, this was not the ideal research paradigm for the study. However, it would be ideal for a study interested in gathering quantitative research from a large population of respondents.

2.2.2 Critical social theory

Critical social theory (CST) is another commonly used research paradigm in the IS field. CST was developed by Habermas (1929), with the core focus to change the study environment through intervention and emancipation (Held, 1980:3). With CST, the oppressing structures in

the study environment are deconstructed and reconstructed to remove negative impacts caused by the oppressing structures currently in place (Myers & Klein, 2011:20).

The CST research paradigm consists of contradiction and change, imagination, critical reasoning, self-determination, and emancipation (Box, 2015:5). According to Myers and Klein (2011:25), when conducting CST research, there are several principles to consider, namely:

- The principle of core concepts used from CST. One or more core CST research aspects should be used in the researcher's data analysis and collection approach.
- The principle of value position taking. Discursive ethics, equal opportunity and open democracy should be the core values driving a researcher.
- The principle of revealing challenging social practices and beliefs. The researcher should attempt to challenge arguments using evidence identified through the analytical data opposing social practices and beliefs.
- The principle of individual emancipation. Self-transformation, self-reflection and human needs should be facilitated through the CST research provided by the researcher.
- The principle of society improvements. The researcher's focus is to prove that changing the study environment is possible by overcoming the current forms of oppressing power.
- The principle of society theory improvements. By using theories, the researcher needs to identify manners to improve them through analytical findings.

The action research qualitative strategy of inquiry is commonly associated with the CST research paradigm. The focus of CST research is to change the study environment, collecting data mainly from focus groups and observations. While CST is an important research paradigm in IS, it was not the ideal research paradigm for the study. However, it would be beneficial for any study meaning to change the study environment.

2.2.3 Design science research

A new type of research design being used in the IS environment is design science research (DSR). Founded by Simon (1916-2001), DSR can be defined as a paradigm in which the researcher investigates the problem area, develops an artefact and uses the artefact to improve the problem environment (Vaishnavi & Kuechler, 2015:9). By solving subsets of the problem, the researcher uses DSR cycles to gather new knowledge regarding the study environment and improve the artefact.

While DSR is a different approach to research in IS, there is a dispute among researchers that DSR might not be considered as a legitimate IS research method. According to Gregor and Hevner (2013:338), there has been some dispute in the IS field as experts in the field of IS have difficulty in identifying how the DSR research paradigm relates to human knowledge. Many authors have failed to correctly interpret the DSR paradigm and make clear how it contributes to existing knowledge. Hevner and Chatterjee (2010:24), in contradiction to Gregor and Hevner (2013:338), state that DSR is an important research paradigm in IS and is generally accepted in the academic community. The DSR paradigm contributes to human knowledge by creating an artefact to understand and change the study environment. Venable (2006:184), in agreement with Hevner and Chatterjee (2010:24), portrays that research professionals, when conducting research, sometimes overlook the DSR paradigm, which is especially oriented to creative problem solving. In IS, the DSR paradigm is mostly applicable to problem solving in society and industry practice.

The focus of the DSR paradigm is to create an artefact to improve the study environment. While this is an important research paradigm in IS, it was not the ideal research paradigm for this study.

2.2.4 Interpretivism

The fourth commonly used research paradigm in the IS field is interpretivism. The interpretive research paradigm was introduced by Dilthey (1831-1911) with the mindset to understand the study environment (Walsham, 1995:376). In the interpretive paradigm, research is conducted with the goal of creating guidelines or a theory regarding “why” the study environment exists or “how” it functions (Klein & Myers, 1999:69).

The researcher needs to keep in mind the values, thoughts and actions of participants in interpretive paradigm research. Interpretivist theories focus on intersubjective interpretation and understanding of people’s norms, values and symbols to make sense of how they experience their world (Remler & Van Ryzin, 2014:51). The researcher usually reports on the research environment and findings in a subjective manner, which is in contrast to the positivist paradigm where the study environment is explained in an objective manner (Denscombe, 2014:2). The researcher needs to understand the views, experiences and perceptions of the participant (Victor *et al.*, 2016:24). Grounded theory and case study are commonly associated strategies used in the interpretive research paradigm. Oates (2006:292) lists characteristics of interpretivism as:

- Multiple subjective realities. Different people or groups of people will perceive the world differently.
- Dynamic, socially constructed meaning: The reality as perceived by individuals or groups can only be accessed or transferred through social interaction, language, shared meaning and understanding.
- Researcher reflexivity: Researchers acknowledge their role and influence in the research.
- People are studied in their natural social settings: The aim of the research is to understand people and their perceptions in their worlds.
- Qualitative data analysis: There is a strong preference for gathering and analysis of qualitative data.
- Multiple interpretations: Researchers expect that there will often be more than one fixed explanation or outcome of the research and that the evidence can be discussed.

When conducting interpretive research, the researcher needs to consider the core principles associated with the interpretive paradigm. Klein and Myers (1999:72) state that, when conducting interpretive research, there are several principles to keep in mind, namely:

- The hermeneutic circle principle. The researcher's human understanding of independent sections or parts of meaning from participants. (Note: human understanding is fundamental to all the other principles listed below).
- The principle of contextualisation. This principle requires the researcher to have a critical reflection of the historical and social background of the study environment in order to understand how the current study environment situation emerged.
- The principle of interaction between the researcher and the subjects. Interaction between the researcher and participants and how the data are received from the participant must be discussed.
- The principle of abstraction and generalisation. The core aspect of this principle is understanding the idiographic ideals relating to the social and cultural actions of data collected from participants, and correlating these with the first two principles.
- The principle of dialogical reasoning. This involves the revision of actual findings and research design guided by sensitively identifying differences between preconceptions ("stories told through the data").
- The principle of multiple interpretations. Different perspectives can be received from various participants, with the same study environment or sequence of events expressed in multiple stories or narratives. Participants will describe the study environment as they experience it.

- The principle of suspicion. The researcher should be sensitive to possible distortions or biases in data collected from the participants.

As the focus of the interpretive paradigm is to develop guidelines or a theory as to how the study environment came to be, this paradigm correlates with the research objectives for this study as the researcher aimed to learn from the study environment and develop guidelines to assist in the migration of business intelligence (BI) architecture in developing organisations.

2.3 QUALITATIVE STRATEGIES OF INQUIRY

The qualitative strategy of inquiry for any study is simply the map that the research follows to conclude the research done. A well-planned qualitative strategy of inquiry builds a foundation for a good study and notifies the reader of how the study came to its conclusion (Hofstee, 2006:107).

2.3.1 Ethnography

The first qualitative strategy of inquiry to consider in the IS study environment is the ethnography methodology. Ethnography is used to study a culture's traditions and practices by interpreting the environment's social interactions. The researcher conducting a ethnography strategy of inquiry to describe the social and cultural based values of languages and interactions (Hancock & Algozzine, 2016:9). Hammersley (2016:8), states that there are three types of assumptions normally associated with ethnography research, namely:

- Naturalism. The researcher must capture the natural human behaviour of the participants in the study environment and not in an artificial setting. This can be achieved by face-to-face communication between the researcher and participant.
- Understanding. Human actions differ from participant to participant and by their interaction with the study environment. The researcher needs to be able to understand the data collected from the participant in that environment. The researcher also needs to understand the cultural perspectives of the participant.
- Discovery. The researcher needs to be able to discover the true nature of the phenomenon by analysing the study environment and data collected from participants. The researcher should conduct research without making too many assumptions.

After researcher has made these assumptions, thorough ethnographic research can be conducted.

2.3.2 Action research

The next important qualitative research strategy of inquiry to consider when conducting a study in the IS field of study is the action research qualitative strategy of inquiry. Action research, by definition, is used to create guidelines for the best practices to solve practical problems (Denscombe, 2014:4). Action research is commonly associated with the CST paradigm and comprises two different focuses, namely “action”, or what needs to be done, and “research”, or finding out what one needs to do. Researching the problem and then taking action to improve the study environment by implementing “measures to improve a situation” is what action research is all about (McNiff, 2009:7). Action research allows the researcher to view the research process and findings in a narrative style, gaining propositional knowledge of the study (Herr & Anderson, 2014:2). What differentiates action research from other research is the reflection of commitment, generation of knowledge, collaborative participative working and transformational practice (McNiff, 2009:20).

A logical question asked when considering action research is “how is action research conducted?” Coghlan and Brannick (2019:9) explain in the action research phases in Fig. 2-1 as follows:

Phase 1 - Constructing. The researcher constructs the issues that reside in the study environment.

Phase 2 - Planning action. The researcher plans what actions are needed.

Phase 3 - Taking action. The researcher implements the action as planned in phase 2.

Phase 4 - Evaluating action. The researcher evaluates the problem area after phase 3, identifying a probable solution. If a solution is not found, then the cycle begins again from phase 1.

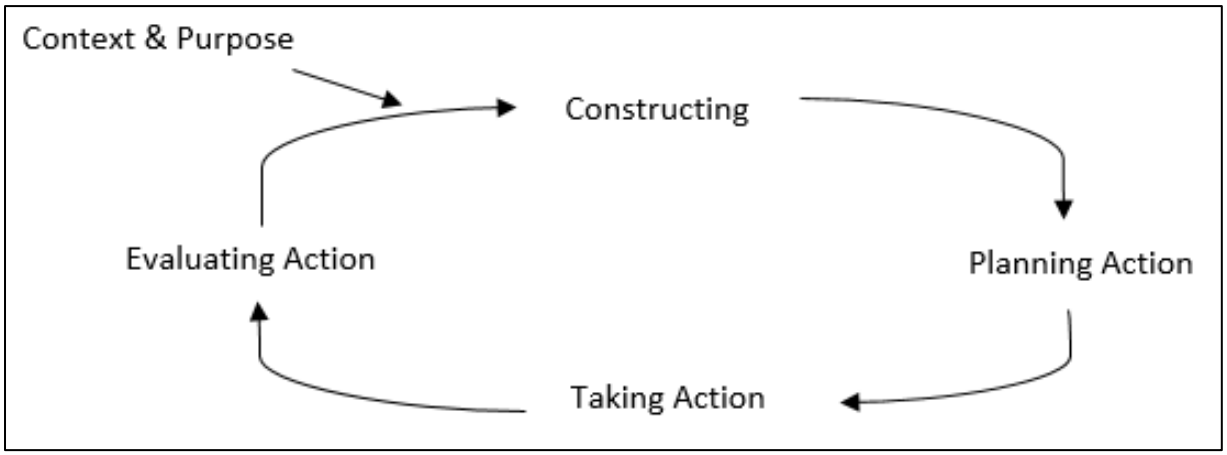


Figure 2-1: Action research phases (Coghlan & Brannick, 2019:9)

Mertler (2019:4) proposes four steps to follow when conducting action research, namely:

1. Identify area of focus. The researcher identifies the possible problem in the study environment.
2. Collect data. The researcher uses data collection methods such as interviews, questionnaires and surveys to collect data form the problem area.
3. Interpret and analyse data. The researcher analyses the data collected in step 2 in the study environment and interprets the data accordingly.
4. Develop an action plan. From the interpretation made in step 3, the researcher can formulate an action plan accordingly.

The steps as described by Mertler (2019:4) are explained in more understandable detail than those of (Coghlan & Brannick, 2019).

2.3.3 Grounded theory

Another research methodology in the IS field that this study considered was the grounded theory strategy of inquiry. Created in the 1960s by Strauss and Glaser, grounded theory is defined as providing new insights and exploring new topics by creating new theories (Denscombe, 2014:4). In agreement Glaser and Strauss (2017:2) states that, in grounded theory, the researcher can develop a new theory based on systematically obtained data from a social environment. In grounded theory, the researcher tends to explain some interaction, action or process by developing a theory, collecting data from the problem environment and deriving meaning from the data collected (Hancock & Algozzine, 2016:9). For this reason, the grounded theory research strategy of inquiry is commonly associated with the positivist and interpretive research paradigms (Smith, 2015:30).

The first step to consider when conducting grounded theory research is to understand how to do the research. Silverman (2016:349) visually presents grounded theory processes as shown in Fig. 2-2:

- Research question. The researcher asks a question relating to a study environment.
- Requirements and sampling of participants. The researcher identifies the requirements for the research and the participants from whom it would be feasible to gather data.
- Data collection. The researcher collects meaningful data from the participants identified and selected for the study.

- Initial coding. The researcher captures and analyses the data collected. Using coding techniques, the researcher then finds meaningful answers from the data.
- Focus coding and categorising. By following a focused approach after coding, the researcher categorises each answer from participants into relevant categories.
- Theory building. The researcher identifies and develops a theory associated with the data analysed.
- Writing up dissertation. The researcher constructs a dissertation from the theory identified, with support from the data analysed.

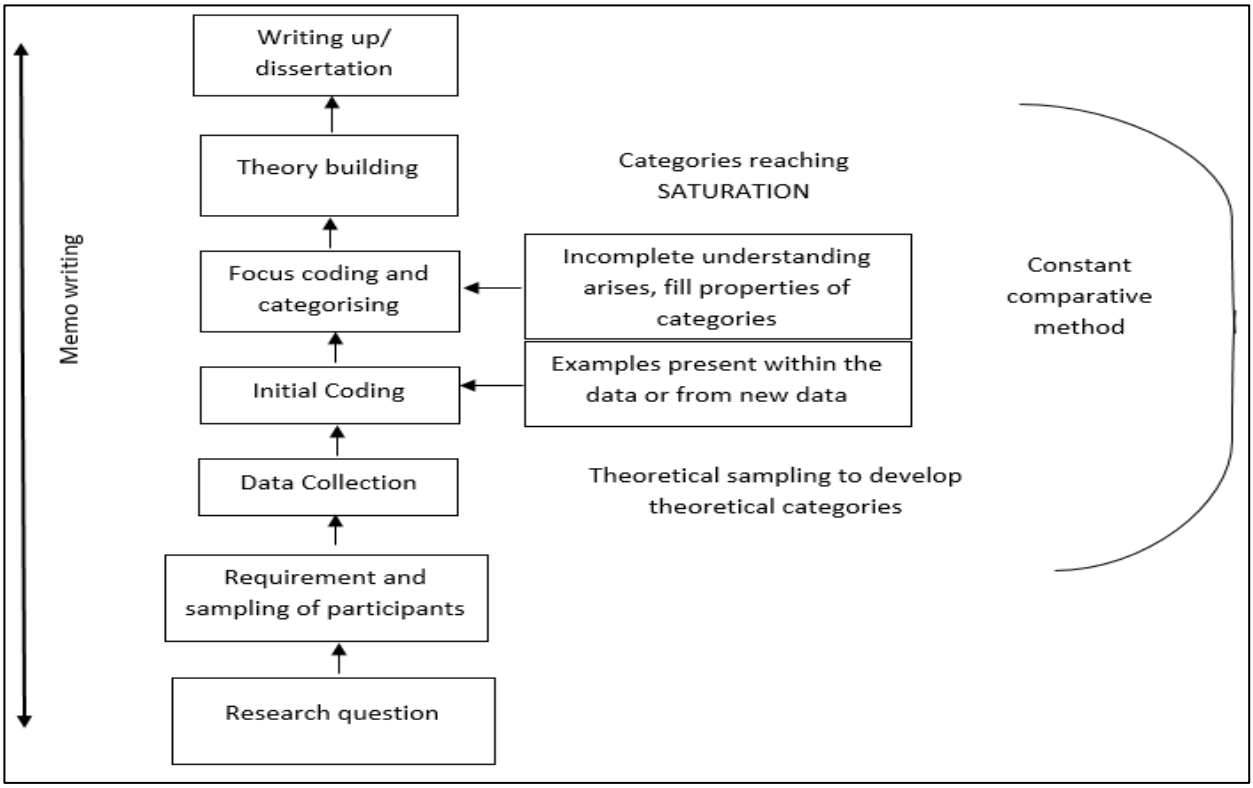


Figure 2-2: Grounded theory process (Silverman, 2016:349)

Birks and Mills (2015:13) visually express, as shown in Fig. 2-3, methods in the process of conducting thoroughly grounded theory research related to gathering qualitative data. The biggest wheel in this process is to collect data with the purpose of identifying key coding categories by initial coding and sampling of data collected while, concurrently, data are being collected, and lastly, identifying the appropriate categories from sampling data. This wheel is the easiest wheel to accomplish in the process. The second wheel focuses on the theoretical sensitivity of the data collected. The intermediate coding process is used to analyse data collected in depth, identifying the core categories associated with data and the theoretical saturation involved. The smallest wheel is where the researcher conducts advanced coding, if

needed, with theoretical integration into the study. While the researcher conducts the process shown in Fig. 2-3, it is important to take memos regarding each step for any additional questions that might arise later in the study.

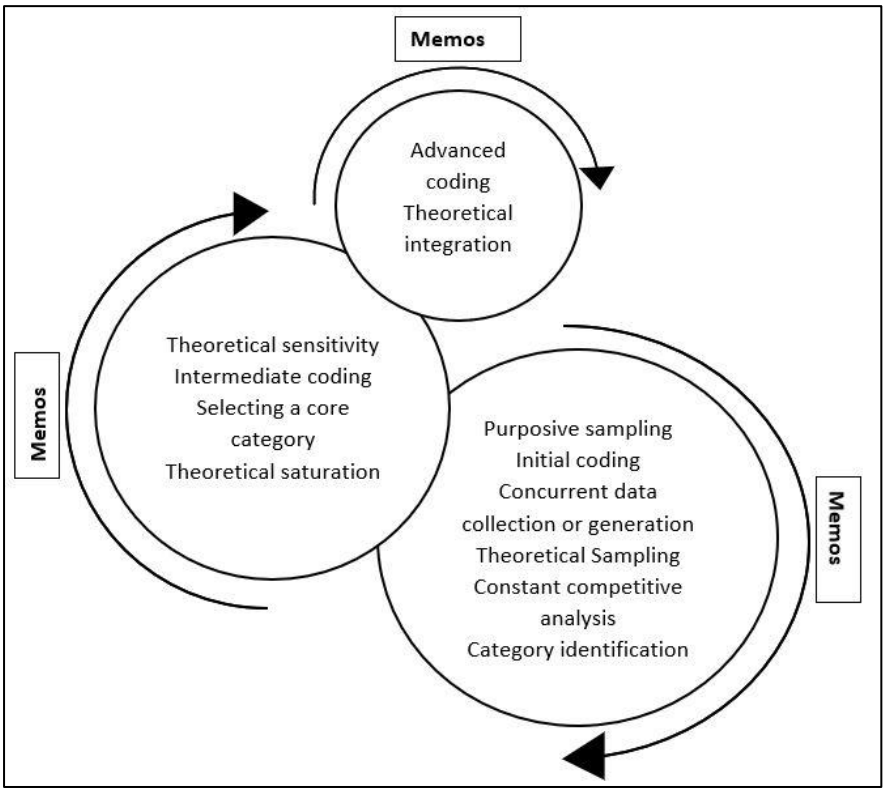


Figure 2-3: Grounded theory methods (Birks & Mills, 2015:13)

According to Silverman (2016:349), the main objectives of grounded theory research are for the researcher to conduct descriptive data analysis beyond boundaries, to identify data patterns associated with the research, and to develop new concepts rather than associating with current theories. Processes in the data are to be theorised. Grounded theory is commonly associated with the interpretive paradigm by developing a theory. This research strategy was not followed in this study, as the study aimed to construct guidelines for developing a BI architecture in a growing organisation as opposed to developing a theory.

2.3.4 Case study research

Another important research strategy when conducting a study associated with the IS field is case study research. Case study research can be defined as a research strategy of inquiry applied to understand how a social setting with complex relationships operates (Denscombe, 2014:4). An explanatory case study is a type of case study where the researcher aims to investigate and find an explanation for a singular natural phenomenon (Zainal, 2007:3).

When considering a case study, it might be necessary to complete a case study planning questionnaire (Hancock & Algozzine, 2016:6). The different criteria of a case study are as follows:

- Case study research reflects a single individual's representation of how a phenomenon impacted a group of participants.
- A natural context is needed for the research, and the research is bound by time and space.
- Various and descriptive resources of information are needed for thorough case study research to be conducted.

Thomas (2015:16), in agreement with Hancock and Algozzine (2016:15), states that extra precautionary steps can be taken when considering a case study research method. As shown in Fig. 2-4, Thomas (2015:16) explains that, to be able to conduct case study research, two parts need to work together. The first part is that the case study needs to take place at a time, at a place or by a person currently influenced by the study environment. Secondly, the appropriate analytical frame or object is needed in order to conduct the research. Without these two parts working together, the researcher cannot hope to conduct an effective case study.

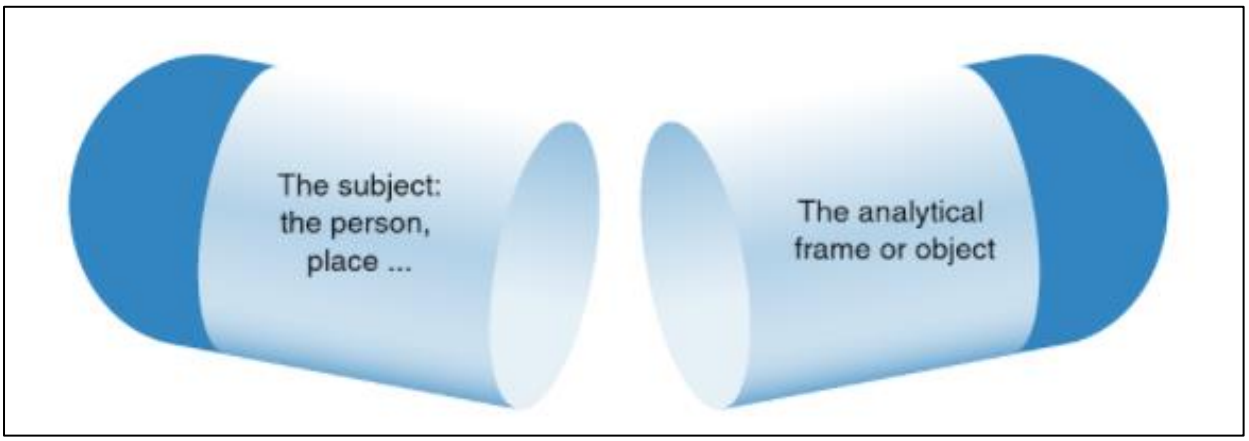


Figure 2-4: Parts needed for a case study research (Thomas, 2015:16)

Thomas (2015:26) later describes, as seen in Fig. 2-5, phases to consider when conducting a case study research:

- Purpose. The research being conducted must have a singular purpose.
- Questions. The research questions asked by the researcher need to be associated with the purpose of the study.

- Literature review. The researcher needs to conduct a complete literature review on the case study topic to identify what questions have been raised in the past regarding similar study environments.
- Approach to research. The researcher needs to select the correct research approach to conduct the case study.
- Design frame and methods. The researcher needs to develop the appropriate frame and design to conduct the study.
- Decision about kind and process of case study. The researcher decides what kind of case study will be conducted and what process will be followed.

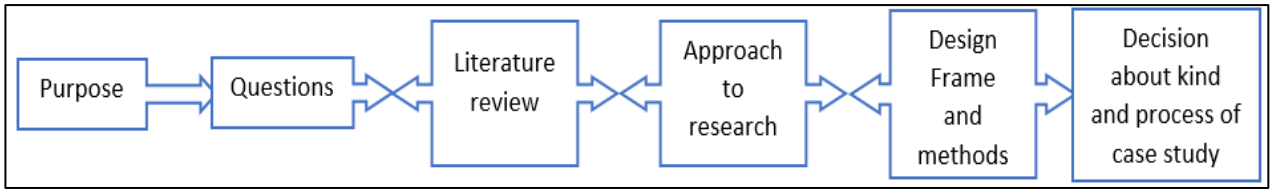


Figure 2-5: Features of designing a case study research (Thomas, 2015:26)

To perform a thorough case study, the most appropriate type of case study needs to be chosen. According to Simons (2014:459), the most common case study types are:

- Intrinsic case study. The case itself is studied by the researcher.
- Instrumental case study. The researcher tries to gain insight into a problem.
- Collective case study. A collective combination of different cases in one study.
- Theory-generated case study. The researcher wants to use a case to test a theory.
- Evaluation case study. The researcher tries to determine the value of the case study by collecting data.

An evaluation case study was used as strategy for the research described in this dissertation. Supporting an evaluation case study, the descripto-explanatory case study describes the study as a precursor to explaining the study (Saunders *et al.*, 2007:140). While description has a clear place in research in business, it is not done for the sake of describing the phenomenon one wants to collect data on, but is rather preparing for the explanatory part of the research. The descriptive part of the study, will be covered in Chapter 4, and the explanatory part, where qualitative data were collected, will be described in Chapter 5. The practitioner-researcher conducted the study at his workplace and the research object on which data were collected was the migration of a BI architecture in a developing organisation.

2.4 DATA COLLECTION

To be able to collect comprehensive and meaningful data from participants, a decision needs to be made whether to collect and analyse qualitative or quantitative data. Qualitative research concentrates on collecting data from participants in a form of verbal reports, while quantitative research focuses on gathering numerical data to create statistical reports (Smith, 2015:2). Hancock and Algozzine (2016:7) state that there is a great difference between qualitative research, which focuses on collecting meaningful information from participants, and quantitative research, which focuses on collecting numerical data only.

2.4.1 Quantitative research

Quantitative data collection concerns numbers and typical, theoretical connections made by variables being associated within the same population. Exploration and generalisation are some of the issues commonly associated with quantitative research, and data collection methods such as surveys are generally used (Brannen, 2017:8). According to Bryman *et al.* (2018:161), the steps to follow to conduct quantitative research are:

1. Elaborate theory. The researcher identifies and elaborates on a theory that needs to be proven.
2. Devise hypothesis. The researcher devises a hypothesis regarding the study environment.
3. Select research design. The research selects the appropriate research design associated with the study.
4. Devise measures of concepts. The researcher identifies concepts to be measured in the hypothesis defined for the study.
5. Select research sites. The researcher identifies the main research sites to be used in the hypothesis.
6. Select research subjects. The researcher identifies participants who are directly impacted by the study environment.
7. Administer research instruments. The researcher identifies and uses the relevant techniques and tools needed to collect data from participants.
8. Process data. The researcher processes the data collected.
9. Analyse data. The researcher uses analysing tools to analyse the quantitative data collected from participants.
10. Develop findings or conclusions. The researcher can then interpret the statistical findings to either prove or reject a hypothesis.

Data collection techniques normally associated with quantitative research are surveys using questionnaires.

2.4.2 Qualitative research

Qualitative research gathers, not numerical data, but meanings from the viewpoint of participants in the study. In this approach, the emotional state of the participant in the study environment is taken into consideration. Because of this, qualitative research is sometimes considered as unscientific (Brannen, 2017:5). According to LoBiondo-Wood and Haber (2017:89), qualitative research focuses on formulating an understanding of a specific phenomenon. Furthermore, qualitative research allows the researcher to collect data in an inductive manner that is descriptive and explanatory. Researchers can identify multiple realities associated with the study by spending time with each participant in the study. Mertens (2014:242) states that there are five types of commonly used qualitative research strategies, namely case study, grounded theory, ethnographic, action research and phenomenological research. It is essential for researchers to know what type of qualitative research strategy applies to their study.

According to LoBiondo-Wood and Haber (2017:90), in order to conduct a thorough qualitative study, the researcher should follow the following steps:

1. Literature review. The researcher needs to perform an in-depth literature review to identify what research has been conducted in the past.
2. Design the study. The researcher needs to design an outline of objectives that will be addressed in the study, according to the type of study. Afterwards, an appropriate research paradigm is selected.
3. Setting: Data collection and recruitment. According to the questions raised by the study, the researcher needs to recruit participants from the relevant study environment to collect data.
4. Data analysis. After the researcher has collected data from the participants, the data are analysed to find truths from the data collected.
5. Findings. The researcher needs to identify and finalise the findings.
6. Conclusion. A conclusion can then be drawn from the findings made through the data collected.

After the steps as described by LoBiondo-Wood and Haber (2017:90) have been completed, a thorough qualitative study may be expected.

2.5 METHODS

The data collection methods associated with the IS research environment that were considered to collect data from participants used in the study are discussed below.

2.5.1 Document analysis

Document analysis is a qualitative research method where the researcher analyses a set of documents on a specific topic to gather data associated with the study (Bowen, 2009:27).

There are some advantages to document analysis, as explained by Bowen (2009:31):

- Efficient method. Document analysis takes less time to complete than other qualitative data collection methods.
- Availability. Documents that are used for document analysis are more readily available and easier to access in organisations.
- Cost effective. Document analysis is more cost effective than other qualitative data collection methods. Researchers tend to select this data collection method if no new data collection is feasible.
- Stability. Collecting data from document analysis is a stable form of collecting data as documents are unlikely to change.
- Exactness. Documents analysed by the document analysis method are accurate and contain exact information of events and references needed.
- Coverage. Documents record entire and sometimes timely discussions and events, and research conducted by document analysis has a broad coverage of the entire topic being researched.

However, Bowen (2009:32) also explains that there are some limitations of using document analysis as data collection method:

- Insufficient detail. Some documents researched would not have all detail needed, as previous research might have only been associated with some parts of the study.
- Low retrievability. Documents can sometimes be very difficult to retrieve for various reasons, such as documents being deleted or lost to natural disasters.
- Biased selectivity. Documents can sometimes only contain biased data regarding topics because organisations are reluctant to record negative information.

It was not possible to use document analysis for this study as the organisation did not keep all documents from when the organisation started, when the BI architecture was implemented and when the data migration project was conducted. This data collection method would, however, be feasible for a study where all appropriate documents are available for the study.

2.5.2 Surveys

Surveys are another commonly used data collection method. Surveys can be defined as the approach to data collection where the researcher creates a structured list of questions, with sets of multiple predefined answers per question from which the participant can choose. Surveying is widely used in research and data collection throughout the world, especially when the researcher needs to gather data from a large population of participants. Surveys are primarily used for quantitative data collection, making data collection by using surveys an appropriate data collection technique for positivist research (Rea & Parker, 2014:3).

There are different types of surveys available to a researcher's disposal. According to Denscombe (2014:9), gathering facts from a social trend or phenomenon to prove a theory is the main objective of surveys. Different types of surveys are postal, face-to-face, telephonic, and using the internet or social platforms.

The type of data usually collected from surveys is numerical. Collecting data from surveys can be a cost-effective process for the researcher (Rea & Parker, 2014:9). This data collection technique was not appropriate for this study as this study aimed to gather qualitative data from a small group of participants. This data collection technique would, however, be beneficial for a study collecting quantitative data from a large set of participants.

2.5.3 Questionnaires

Questionnaires can be defined as a data collection method where a researcher creates a structured form using predefined questions that participants need to complete and return to the researcher for analysis after completion. Questionnaires are usually used to collect qualitative data with open-ended questions (Moser & Kalton, 2017:257).

When a researcher selects questionnaires as the primary data collection technique, it is necessary to select the most appropriate type of questionnaire. Sekaran and Bougie (2016:144) state that there are three types of questionnaire techniques that can be used, namely:

- Personally-administered questionnaires. The researcher personally administers the questionnaire with the participant. The disadvantage of this technique is that it takes extra time and effort to conduct from the researcher's perspective.
- Mail questionnaires. The researcher creates a questionnaire and sends the questionnaire to the participants through the postal services, the participants complete the questionnaire, and mail the answered questionnaire back to the researcher. The disadvantage of this technique is that the response rate from participants is quite low.
- Electronic questionnaires. The researcher creates an electronic questionnaire and then proceeds to send the questionnaire (or a link to the questionnaire situated online) to the participants via email. The participants then complete the questionnaire on an electronic device and send the questionnaire back to the researcher. The disadvantage of this technique is that the participant must be computer literate to receive, complete and return the questionnaire.

After the researcher has selected an appropriate questionnaire technique, a thorough data collection approach can be conducted to collect qualitative data from participants. Collecting data through questionnaires can be beneficial for studies that need to provide a large amount of participant data.

2.5.4 Focus groups

Focus groups can be defined as a data collection method where an individual researcher collects data from not only one participant but a group of participants at the same time. With this data collection technique, the participants can discuss the study among themselves, providing relevant aspects of concerns and ideas regarding the study and their grouped experiences (Parahoo, 2014:320).

The core elements of collecting data from focus groups need to be understood before a researcher can select this data collection method. According to Stewart and Shamdasani (2014:9), there are four core elements that drive focus group research, namely:

- Focused research. Focused research groups are associated with one objective.
- Group interactions. The researcher interacts not with one individual but with a group of individuals regarding the same topic, but with different perspectives regarding the study environment. The participants also interact with each other to share their individual perspectives.

- In-depth data. The researcher gets a deeper insight into the study environment through the participants' interactions.
- Humanistic interview. The researcher experiences the participants' reactions to questions as well as being able to sense the emotional manner in with a question is answered.

Collecting data from focus groups can be beneficial for the researcher as focus groups provide a general objective to a study environment, providing the researcher with a central representation of the study environment (Krueger & Casey, 2014:3). As this data collection technique is important to studies attempting to collect data from a group of individuals and get a general sense of emotion regarding the study environment, this data collection method was not suitable for collecting data from participants associated with this study as the study aimed to gather individual answers, emotions and reactions to questions asked to participants as opposed to a group of participants at the same time.

2.5.5 Observations

Observation can be defined as a data collection technique used where the researcher observes the participants in a study environment while collecting data from their actions, expressions and experiences inside the study environment (Jamshed, 2014:88). Observation is one of the most important techniques used for qualitative data collection. Observations usually occur in a naturally selected environment as opposed to interviews that are conducted in a designated area selected by the researcher (Merriam & Tisdell, 2015:137).

There are two types of observation techniques that a researcher can use to collect data. Participant observation is where the researcher is part of the participant group in the study environment and collects data not only from the participants' experience in the study environment but from his or her own experience. According to Musante (2014:258), when a researcher is a participative observer, the data collected for the study is much more accurate as it reflects the researcher's own experience of the study environment. The researcher might also, from interaction, identify new research questions to ask during the data collection process. In some cases, for the researcher to be a participant observant is the only option as no participants can be found or the researcher is involved in the progression of the study environment. The second observation technique is non-participant observation. This is when the researcher is only observing the participants from a perspective outside of the study environment (Laurier, 2016:169).

Collecting data with this data collection technique can be beneficial for a study that aims to collect data where the researcher is part of the study environment. However, this data collection

technique was not suitable for this study, as the researcher was not part of the study environment from the start of the organisation's existence.

2.5.6 Interviews

An interview, by definition, is the process during which the researcher, called the interviewer, collects information from a participant, formally known as the interviewee, regarding a specific study environment (Parahoo, 2014:308). According to Spradley (1979:3), to efficiently conduct an interview, the interviewer needs to be able to learn from interviewees instead of studying them. The interview process is considered one of the main sources of data collection for qualitative research. In a real-world environment, most studies conducted by collecting qualitative data from participants integrate the informal interview process with an observational aspect (Merriam & Tisdell, 2015:137). Recording devices are considered as a great way for the interviewer to keep track of all key points discussed with the participant during the interview process.

The interview process allows participants and the interviewer to exchange predefined dialogue. The responses from the participants are captured by the interviewer either on paper or by a recording device. After the interview, the interviewer reviews the recording of the interview to make sure that no important data from the participant were missed (Jamshed, 2014:88). There are two types of interviews that can be conducted. The first type is structured interviews, consisting of the interviewer constructing predefined questions to ask the interviewee and a selection of answers that are in line with what is exactly needed from the study. The second interview type, semi-structured interviews, give leeway to the participant to answer the predefined questions asked by the interviewer in an open dialogue manner (Smith, 2015:30). According to Bryman *et al.* (2018:214) and Smith (2015:30) the different types of interviews are as follows:

- Structured interview. In this interview process, the questions have been structured before the interview is conducted and cannot be altered throughout the interview process.
- Qualitative interview. The interviewer creates the questions dynamically as the interview process is being conducted, there are limited structured questions, and the interview is considered as an informal interview.
- Semi-structured interview. This is a structured interview process, but the interviewer has permission to ask questions outside of the pre-constructed question list.
- Intensive interview. This interview process is considered as an unstructured interview process where there are no questions prepared before the interview.

- Focused interview. The researcher asks questions regarding a focused topic of interest and nothing else.
- Group interview. The researcher interviews a group of participants at the same time and the information discussed is only partially focused.

Semi-structured interviews were selected as the primary data collection technique for the research. Additional questions may be asked during semi-structured and informal interviews. The interview as data collection technique is aligned with the principle of interaction between researcher and participant as stated by Klein and Myers (1999:72), where data are collected from the participants by face-to-face interaction.

2.6 DATA ANALYSIS

After data have been recorded from participants and text documents have been created, the next step for the researcher is to analyse the data collected. Data analysis, by definition, is when a researcher analyses data collected from participants in the study environment to find patterns and themes associated with questions asked by the researcher. Data collected from participants is imported into an analysis system such as ATLAS.ti and appropriate codes and themes are identified to gather true meaning from the data provided, which can then be used to make valuable conclusions and decisions (Silverman, 2016:332).

The researcher needs to use the maximum possible available resources and effort when gathering and analysing qualitative data to make sure that every meaning and expression from every study participant is included. Too often, researchers try to cut corners when it comes to data analysis, most of the time resulting in important data not being recorded (Caudle, 2004:419). According to Seers (2012:2), receiving piles upon piles of completed qualitative data back from participants can be overwhelming at first, but the researcher need not worry as themes in the data will emerge more quickly when the researcher spends more time in detailed analysis as opposed to trying to skip through all the qualitative data collected.

Distinguishing between inductive reasoning and deductive reasoning is what makes qualitative analytics possible (Thorne, 2000:68). To comprehensively analyse the data through inductive reasoning, the following criteria as mentioned by Hofstee (2006:117) were used:

- Content analysis. Behavioural and verbal data were categorised into interpretative and descriptive levels of data and summarised into categories of code.

- Thematic analysis. The codes were categorised into appropriate themes identified in the literature.

The researcher needs to make sure the information gathered is trustworthy and reliable, check the credibility of the data retrieved and ensure that the data are dependable (Bryman *et al.*, 2018:384). There are two levels of qualitative data analysis: the first is a descriptive level of analysis, where the researcher reads through the data and, using first-stage initial coding, identifies categories of descriptive meaningful information to be used in the study. The second level of qualitative data analysis is conceptual, when the researcher uses coding and analysis tools to investigate the data collected from a conceptual point of view to answer questions identified for the study (Friese, 2019:18).

When conducting thorough qualitative data analysis, the researcher needs to follow important rules. Silverman (2015:213) states that the rules to qualitative data analysis are as follows:

- Avoid the busywork and get down to the analysis as soon as possible.
- Try a variety of theoretical approaches to data analysis.
- See where the analysis development is leading before establishing a hypothesis.
- Focus on detailed small groups of data instead of the overall data collected.
- Analyse the data intensively and do not look for telling examples in the data collected.
- Try to be sequence-focused while analysing data collected.

Bernard *et al.* (2016:245) recommend the following steps for qualitative data analysis:

1. Using the existing theory, formulate a hypothesis or research question.
2. Select a section of text from the data retrieved to test the hypothesis or question.
3. Create codes and associate them with the selected questions or hypothesis.
4. Using the text selected, test associated variables accordingly.
5. Integrate the codes identified in step 3 to the rest of the text collected.
6. Create a case variable matrix using the codes and text.
7. By using analysis techniques, analyse the case variable matrix.

When following the steps as described by Bernard *et al.* (2016:245), coding the data are considered a very important step. Researchers commonly use coding methods to analyse large amounts of qualitative data collected from participants. Coding is the way that patterns and key

words are determined from analysing the data collected from participants. New ideas can be developed through coding (Richards, 2014:103).

In this study, qualitative data were collected through semi-structured interviews. In some instances, such as the data migration section of this study, open-ended interviews were also conducted.

2.7 STUDY PLAN

In this section, a complete study plan identified for this study is discussed. The study plan comprises two sections. Firstly, a graphical representation is provided of how the study is structured and what is discussed in each chapter. Secondly, the research plan is described. The features of a case study (section 2.3.4), as seen in Fig. 2-6, will be used as a guideline to describe the design of the study.

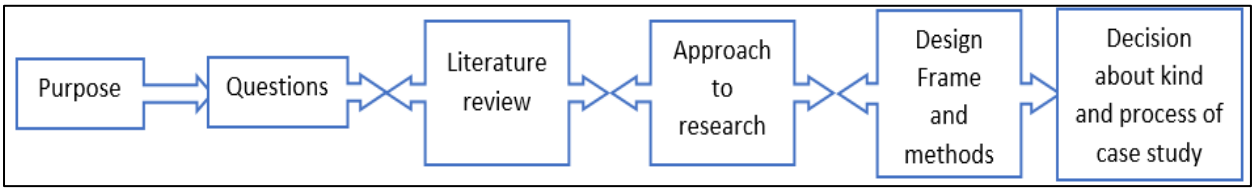


Figure 2-6: Features of designing a case study research (Thomas, 2015:26)

Purpose

The purpose of the case study is to determine guidelines to assist developing organisations to overcome challenges of data growth while successfully implementing a thorough BI architecture.

Questions

The main research question for the study was:

> *How can guidelines be compiled to assist developing organisations to overcome challenges of data growth while successfully implementing a BI architecture?*

The following sub-questions were used to guide the descriptive part of the case study research:

- How did the start-up organisation implement a BI architecture?
- How does the newly implemented BI architecture differ from the initial architecture?

- How could data from two DBMS platforms be integrated with minimal impact on the BI architecture implemented?

The following sub-question was used to guide the explanatory part of the case study research:

- What were the perceptions and experiences of stakeholders involved in the development of a new BI architecture?

Literature review

To motivate and support the study, the following concepts were defined and described from the literature:

- Data warehouse
- Business intelligence
- Data mining

These concepts were also used to justify the descriptive part of the research case study.

Approach

The interpretivist research paradigm was selected as the best possible research framework to use for the study. The appropriate qualitative strategy of inquiry identified for this study was to use a case study to investigate the phenomenon, gather qualitative data from participants and report on how the study environment originated and changed

Design frame and method

A descripto-explanatory case study was performed. The descriptive case study part was used to discuss the initial BI architecture, current BI architecture and data migration process.

The explanatory part of the case study discussed the outcome of the empirical research. The aim of the empirical research was to evaluate the perceptions and experiences of stakeholders throughout the BI migration phases. Semi-structured interviews were selected as the primary method to collect qualitative data from three levels of organisational participants.

The reason for this was that the researcher was able to structure a list of questions to ask the participants regarding the study environment but had the option to ask derived questions during the interview process. All the interviews were recorded by the researcher with the full consent of each participant. After the semi-structured interviews had been concluded, the researcher

transcribed the recorded data into documents. The transcribed documents for each participant were then loaded into ATLAS.ti where the appropriate codes and themes were developed. After concluding the analysis, the researcher was able to present the results and findings of the research.

Decision about kind and process of case study

A description of the study environment, the BI solution implemented in the organisation when the organisation was of start-up size, the BI development project plan to enhance the organisation's BI architecture and the currently implemented BI architecture will be provided in Chapter 4. The way data was integrated from two different DBMSs into one and how the additional data were introduced into the organisations BI architecture is also discussed in Chapter 4. In Chapter 5 the explanatory part of the case study (empirical research) is described. The participant selection process, data collection, data analysis and the results of the data collected are presented. From the results of the descripto-explanatory case study, guidelines for the migration of BI into a growing organisation were constructed.

A detailed chapter classification of the dissertation is presented in a graphical representation of the study. Figure 2-7 visually displays the study flow diagram for developing guidelines on how an organisation can sustain a fully functioning BI architecture while going through tremendous growth over a short number of years. In Chapter 1, the key objectives to the study, the problem statement, objectives of the study, research design, literature review, empirical study, ethical considerations and chapter classification were discussed. Chapter 2 provided an overview of the different research paradigms available in the IS environment, qualitative strategies of inquiry, data collection methods, data analysis techniques and a study plan. In Chapter 3, the literature review is discussed regarding concepts such as data warehouses, BI architecture and data migration.

Chapter 4 details the case study and associated concepts, such as the case study environment, BI architecture implemented when the organisation was of start-up size, the BI architecture currently implemented in the organisation, and migration of data from one organisation's database management system (DBMS) to another. Chapter 5 discusses the participant selection process, the data collected and analysed, the results of the research and the construction of guidelines for implementing BI architecture in a growing organisation. In Chapter 6, the study concludes on the study outcomes and provides recommendations for further study research.

After conducting the study as described in Fig. 2-7, a comprehensive understanding and conclusion to the study could be achieved.

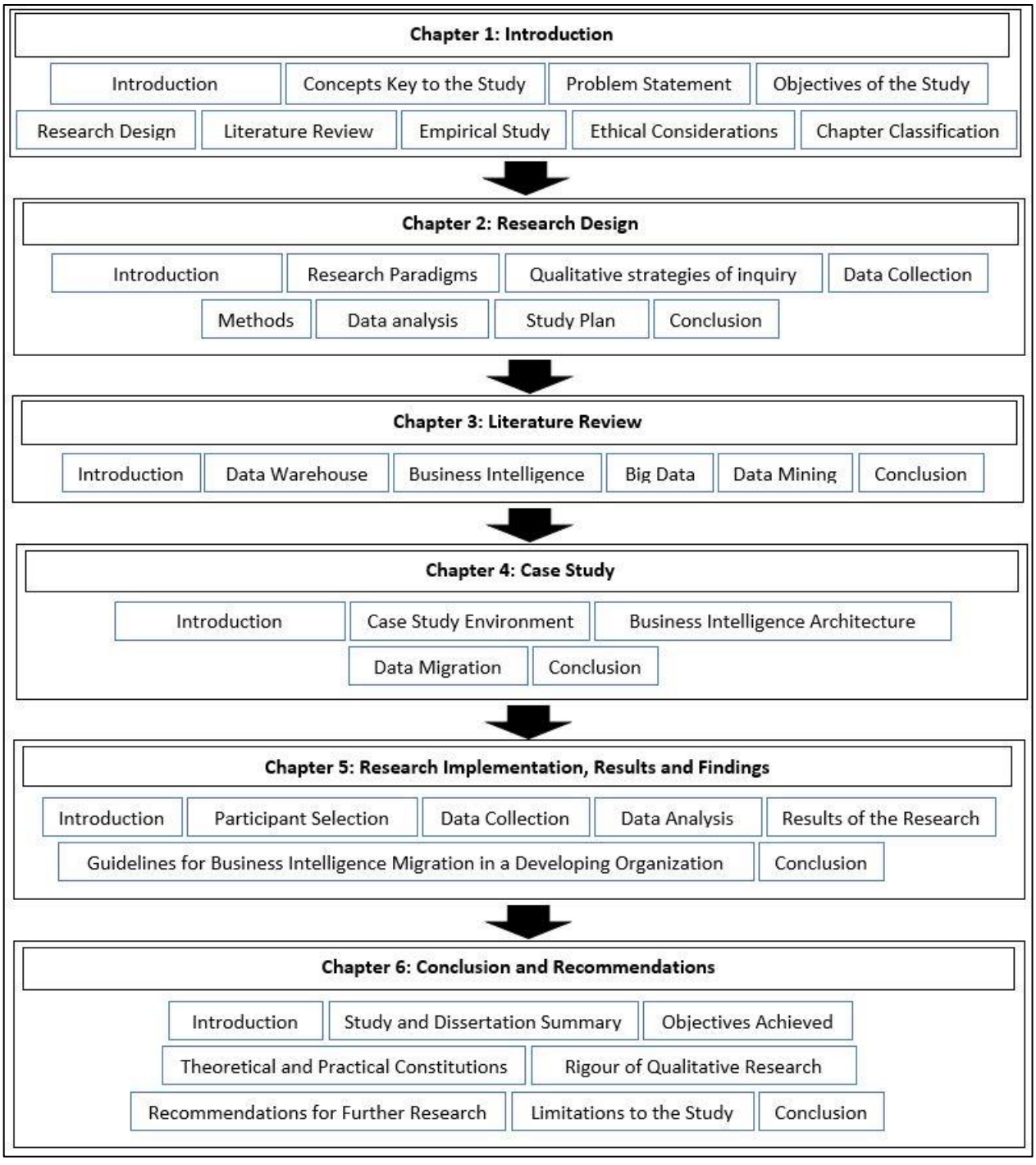


Figure 2-7: Study plan diagram

Rigour of the study

The interpretive research principles of Klein and Myers (1999:72) and criteria provided by Neergaard *et al.* (2009:3) as applied to this study, were discussed in section 2.2.4 and is listed in Table 2-1.

Table 2-1: Guidelines and principles directing the rigour of this study

Guideline or principle	Description
Authenticity guideline	The participants must be able to speak freely, their voices must be heard when spoken, and must be presented in an accurate way.
Credibility guideline	A true researcher's perspective must be captured and portrayed.
Integrity guideline	Bias must be reflected, informants must be validated, and peer review of data collected and analysed must be ensured.
The hermeneutic circle principle	The hermeneutic circle principle is primarily a representation of the human understanding of independent sections or parts of meaning from participants in the study. Human understanding is fundamental to all the other principles listed below. The researcher needs to consider and portray the participants' emotional state with regard to the study environment.
The principle of contextualisation	The principle of contextualisation requires the researcher to have a critical reflection of the historical and social background of the study environment in order to be able to understand how the current situation emerged. This principle requires that the researcher understands the BI architecture migration in its context, from its origin to its current state.
The principle of interaction between the researcher and the subjects	The principle of interaction between the researcher and the participants is reflected in the structuring of the data received from the participants. The researcher needs to be able to relate to the participant while conducting the research to truly understand what is being described by the participant. In this study, the researcher was exposed to the study environment during the BI architecture evolution and data migration from one DBMS platform to another.
The principle of abstraction and generalisation	The principle of abstraction and generalisation entails understanding the idiographic ideals relating to the social and cultural actions of the data collected from the participants. This principle needs to be correlated with the first two principles. The researcher needs to analyse data collected from participants by abstracting meaning in the context of the study, leading to the extraction of guidelines applicable in other contexts.
The principle of dialogical reasoning	The principle of dialogical reasoning revises the actual findings and research design, guided by sensitively identifying differences between preconceptions ("stories told through the data"). The researcher should be aware that the conclusion of the research design and research data, analysed from participants, might be different. In the case of this study, implementations regarding BI architecture and data migration might differ from the opinions of the participants directly impacted.
The principle of multiple interpretations	The principle of multiple interpretations involves different perspectives received from different participants in the same study environment or sequences of events expressed in multiple stories or narratives. Participants will describe the study environment as they experience it. The researcher should be aware that the opinions of participants impacted by the same study environment might differ.
The principle of suspicion	The principle of suspicion requires the researcher to be sensitive to possible distortions or biases in data collected from the participants. The researcher should not only rely on the words of the participants during the interview but should also look at behaviour such as facial expressions and body language when questions are answered.

The use of an analytical tool assisted in the analysis of qualitative data and peer coding of data ensured that the codes and themes were trustworthy and dependable (Oates, 2006:294).

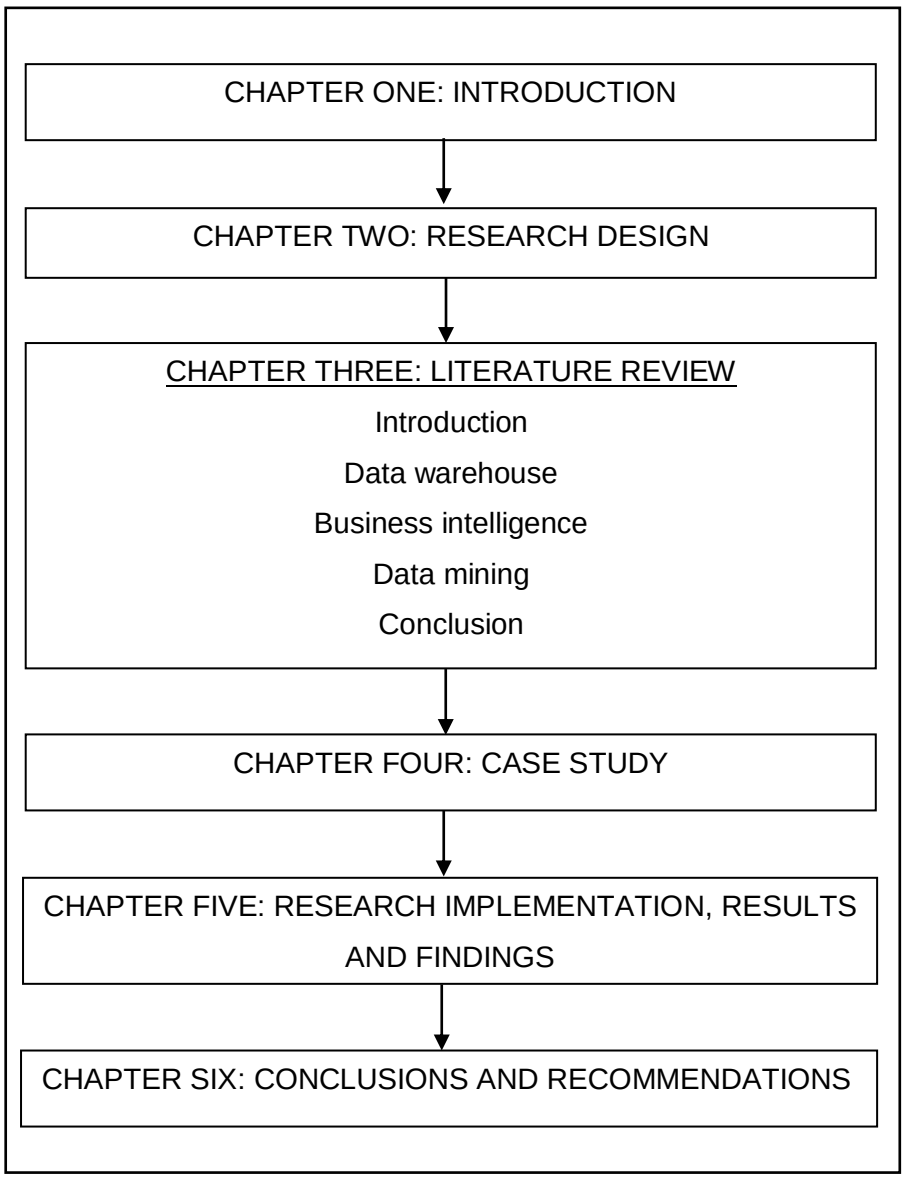
2.8 CONCLUSION

Discussed in this chapter were different research paradigms and their associated qualitative strategies of inquiry. The interpretivist paradigm was selected to be the main research paradigm

for this study, as the researcher wanted to understand how the organisation used in the study expanded and implemented a BI architecture to manage its exponential data growth. Case study research was selected to be the primary qualitative strategy of inquiry for this study, which reports on a singular phenomenon.

Different methods to collect data from participants and how to analyse the data collected were also discussed in this chapter. The study concluded that semi-structured interviews would be used as the primary data collection method for collecting qualitative data from participants. The data tool selected to analyse data received from participants was ATLAS.ti. In the next chapter, the literature review regarding BI architecture components such as the BI lifecycle, evolution of BI and BI maturity is presented. The literature on data migration, such as the data migration phases, extract, transform and load (ETL) process, data scientist and the difference between NoSQL DBMS and Relational DBMS is also discussed in the next chapter.

CHAPTER 3: CHAPTER FLOW



CHAPTER 3 LITERATURE REVIEW

3.1 INTRODUCTION

The study aims to provide meaningful knowledge to organisations experiencing exceptional data growth. The study can also be beneficial to organisations in need of improving their BI architecture or planning to implement data migration from one database management system (DBMS) to another. The literature is used to justify the reason for the research and provide an understanding of the concepts used to ground the research (Aveyard, 2014:2). The purpose of this chapter is to discuss the literature review conducted on past studies regarding data warehouse (DW), BI architecture, and data migration, correlating to theoretical objective T01.

A literature review is simply reviewing literature on past studies that are associated with the current study. There is no set of rules to conducting a literature review, as every study's literature review is unique (Hart, 2018:3). Business intelligence (BI) can be used in organisations to expand decision-making possibilities. New and insightful forms of analysis and forecast techniques can be performed on production data through the interpretations of a thoroughly designed BI architecture. With regard to the large data growth in organisations, Müller and Bundesbank (2019:2) state that most data are stored in unordered silos, forcing organisations to develop initiatives such as BI architectures and data warehouses to handle the large amounts of data ingested daily. With the expansion of data in organisations, new and innovative ways to clarify data growth in an organisational environment, as well as ways to present the increasing amount of data in a visual manner in order to make important decisions, are necessary to support users (Obeidat *et al.*, 2015:53).

Limited literature was found discussing theories or hypotheses regarding the migration of BI architecture in any sort of organisational environment in the last five years. However, literature from the past five years associated with BI architecture is included in this chapter. This study concentrates on constructing guidelines to develop an on-premise BI architecture in a growing organisational environment, as opposed to a cloud-based BI architecture.

In the subsequent sections, the following topics are addressed: the design and development of a DW, including the DW data flow, the DW maturity model, extract transform and load (ETL) process flow, dimensional modelling, the Kimball methodology and DW automation (section 3.2). The BI architecture design and development process are addressed next, including BI evolution, the BI maturity model, BI life cycle and analysis services (section 3.3). Data mining components such as data scientist, data mining life cycle and data mining phases depicting

detailed explanation of how and by whom the data are normally interpreted being imported into a data analysis system (section 3.4). Lastly, a conclusion to the chapter is provided (section 3.5).

3.2 DATA WAREHOUSE

The first step to understanding what a BI architecture entails is to understand what a DW is. According to Iyer and Lakhtaria (2017:61), a DW is a repository of data collected from heterogeneous sources, cleaned and developed to find one source of truth in organisational information. In agreement, Linstedt and Olschimke (2015:4) state that a DW system is a BI architecture tool that provides one platform of data consisting of multiple data sources, enabling operational and strategic decision-making. To be able to have a thoroughly developed BI architecture, a well-designed and well-implemented DW needs to be developed. Concepts to understand when implementing a DW, such as the DW data flow (section 3.2.1) and the DW maturity model (section 3.2.2) are discussed in this section. Furthermore, concepts to design and implement a DW, such as the ETL process (section 3.2.3), dimensional modelling process (section 3.2.4), the Kimball methodology (section 3.2.5) for implementing a DW, and DW automation (section 3.2.6) are discussed in this section.

3.2.1 Data warehouse data flow

The first step in developing a DW is to understand how data flows through a DW, and how to transform this data into meaningful information for the organisation to use. Data warehousing forms the key technology to BI solutions because of the ETL tool functionality inducted by the DW (Łęgowik-Świącik *et al.*, 2016:4). If implemented incorrectly, a DW can be a time-consuming and costly project (Rahman & Rutz, 2019:735).

To develop a thorough DW data flow process, Müller and Bundesbank (2019:2) explain that there are three basic steps to data integration into a DW, namely:

1. Heterogeneous data integration. This process is simply to store data from various locations into once central location. Common procedures can be used to complete this process.
2. Uniform data method. This is a process used to transform all data stored in one central location, into a standard structure in line with the organisation's data terms and policies.
3. Semantic harmonisation. This process is used to integrate relationships between the different data. This process can be very tricky to perform, as this step makes it possible for data to be linked together in a common dictionary.

Müller and Bundesbank (2019:2) further state that this process should formulate the basic thoughts of a DW developer before a DW is designed.

Sherman (2014:69) in contradiction to Müller and Bundesbank (2019:2), visually expresses a basic model that describes how an enterprise DW's data flows in Fig. 3-1. The systems used to keep records, such as enterprise applications, customer relationship management (CRM), cloud applications and databases, send and retrieve data from a central DW. An ETL process makes sure that the data are clean from any anomalies or faults and that there is one source of the true data that is stored in the DW. The data from the DW can then be retrieved by the BI architecture in place for analytical and forecasting purposes.

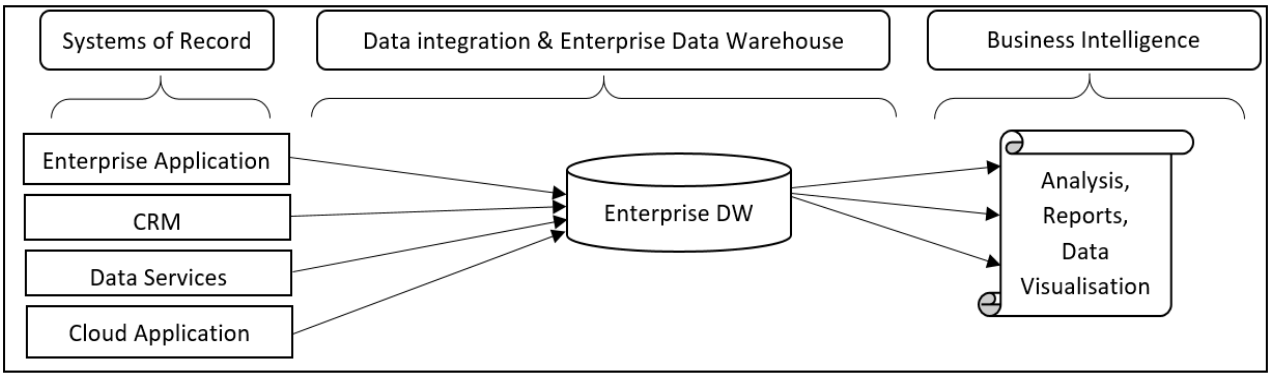


Figure 3-1: Enterprise DW as adapted from Sherman (2014:69)

This study believes that the description by Sherman (2014:69) in Fig. 3-1 is more appropriate to the study as this process of DW data flow is much more detailed than the process stated by (Müller & Bundesbank, 2019:2).

3.2.2 Data warehouse maturity model

Implementing a DW can be challenging and complex. There is no 'simple' solution for all organisations, as every organisation's data structure differs. Therefore, an organisation should be aware of its implemented DW environment and allow for continuous development of its DW. The organisation should not only focus on the technical aspects, but also on the organisational processes that sustain them. This is where a DW maturity capability matrix comes into play. Spruit and Sacu (2015:1512) visually express, in Fig. 3-2, the organisational and technical solutions steps that should be considered in order to create a DW maturity capability model:

1. Investigate and provide a set of points for the DW technical solutions implemented, such as the general architecture and infrastructure.

2. Analyse and provide a set of points for the data modelling techniques currently in place in the DW.
3. Inspect and provide a set of points for the ETL process currently implemented and the BI solution aligning to the DW.
4. Allocate a set of points for the DW organisational processes, such as the development and service processes identified.
5. With all the points accumulated, the organisation can perform a maturity assessment score on each of these processes to identify how mature each process is.

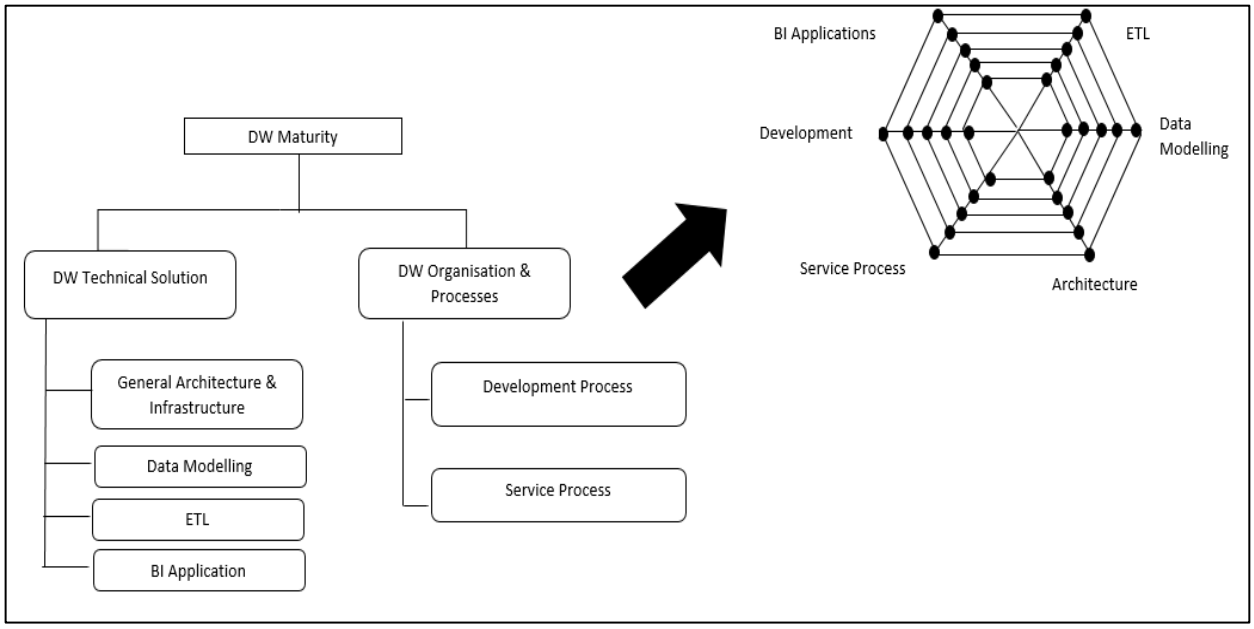


Figure 3-2: Data warehouse maturity capability model as adapted from Spruit and Sacu (2015:1512)

Spruit and Sacu (2015:1514) later explain that the DW maturity phases can be identified according to five levels, namely:

Phase 1 - The initial level. The DW has very limited maturity and no expertise.
Phase 2 - The repeatable level. This level has limited maturity and expertise.
Phase 3 - The defined level. This level has a medium level of maturity and expertise.
Phase 4 - Managed level. This level has an advanced level of expertise.
Phase 5 - The optimised level. This level has a highly advanced level of maturity and expertise.

After the organisation has made the assessment as explained by Spruit and Sacu (2015:1514), it will have the information needed to decide the maturity of the DW currently implemented.

3.2.3 Extract, transform and load flow

When an organisation decides to develop or upgrade a DW, one of the most important concepts to keep in mind is the ETL process. In this section, the ETL process is discussed, through which data are retrieved from the different CRM systems, loaded into a central database and transformed into meaningful information. To be able to design a good ETL process, the steps that must be taken should be thoroughly understood. In a similar study, performed by Microsoft, there are two globally known ways to ingest data from various sources into a DW, the first is by performing an ETL process where the data are extracted, transformed and then loaded into the DW (Anon, 2018b). The second is an ELT process where the data are extracted from various sources and then loaded into the DW. Only then is it transformed into the desired structure. Andersen *et al.* (2018:5) visually express, in Fig. 3-3, how the ETL process should be conducted in three phases:

Phase 1 - Initiation. The factual data are loaded to the staging environment while dimensional integrity is maintained.

Phase 2 - Processing. Data are retrieved from the staging environment, cleaned of any anomalies, transformed and moved to the fact tables.

Phase 3 - Data migration. Data are moved to the live environment, where types 1, 2 and 3 of slowly changing dimensional handling are implemented.

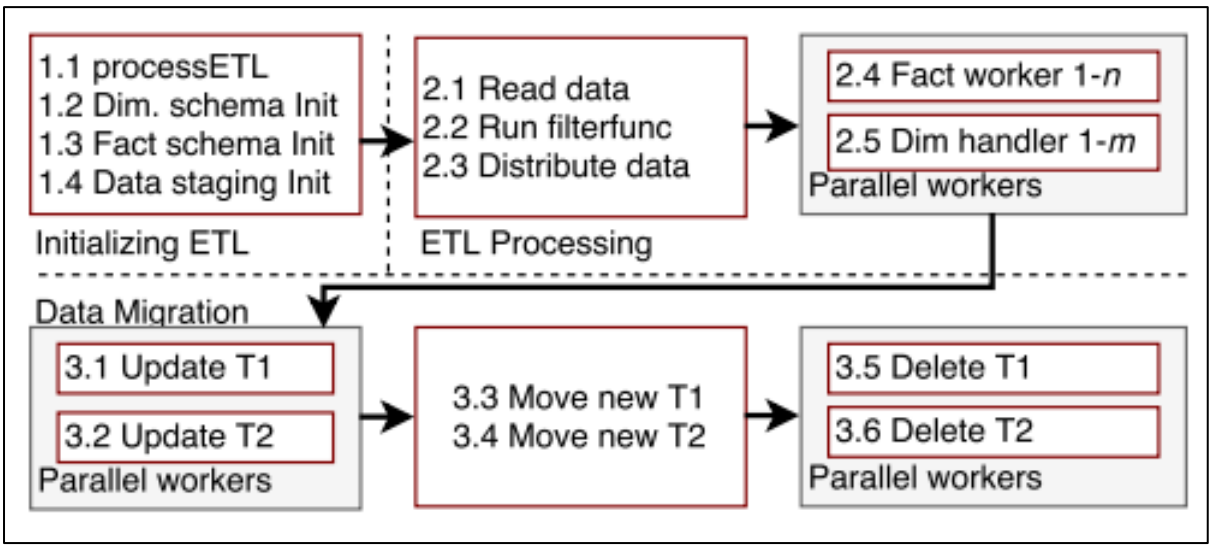


Figure 3-3: Extract, transform and load process flow as adapted from Andersen *et al.* (2018:5)

The process portrayed by Andersen *et al.* (2018:5) is easy to understand and to implement. This ETL process can also be adapted to work in most organisational BI architectures. Meehan

et al. (2017:2) differ from Andersen *et al.* (2018:5) when describing the ETL process as this approach to ETL design keeps the entire BI architecture in mind, as seen in Fig. 3-4, during the ETL process flow:

Phase 1 - Data collection. The ETL tool retrieves data from the multiple data sources.

Phase 2 - Streaming ETL. The data are transformed in a two-step solution, first cleaning the raw data by applying the appropriate data time convention to the data attributes, and then making sure the data to be added is unique and all key constraints are checked.

Phase 3 - Online analytical processing (OLAP) backend. The transformed data are loaded into the data warehouse.

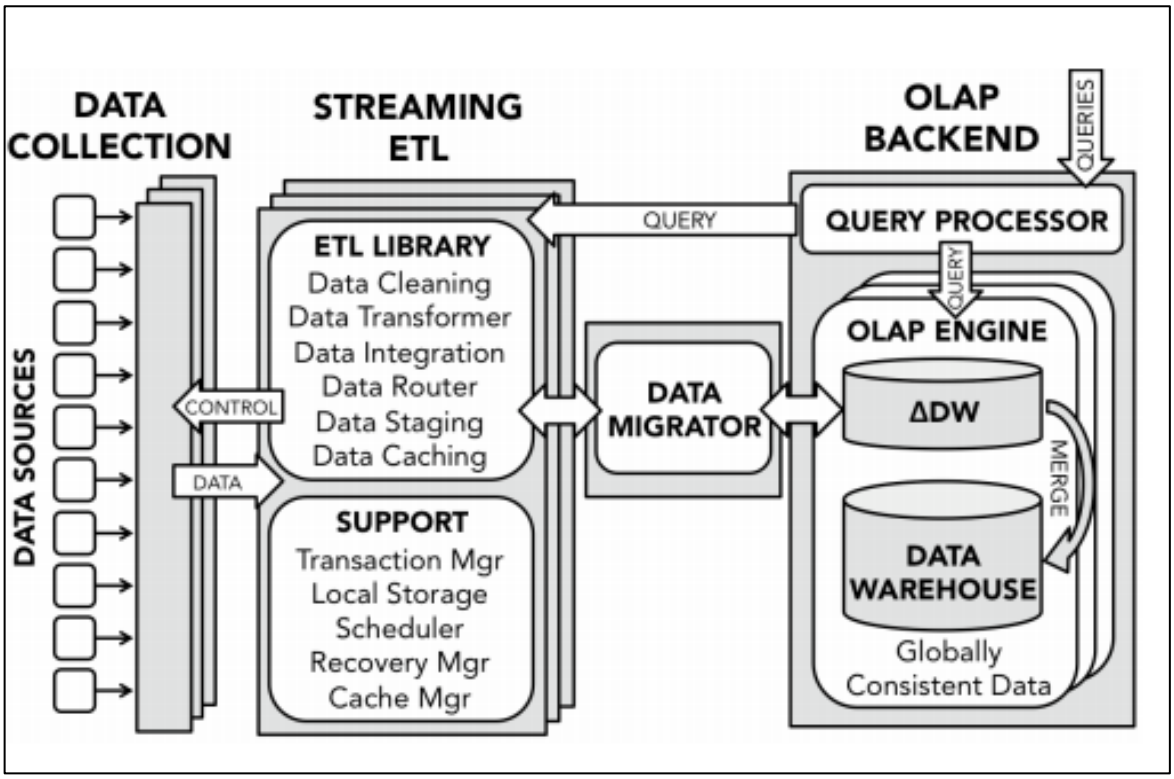


Figure 3-4: Extract, transform and load process flow as adapted from Meehan *et al.* (2017:2)

The approach in Fig. 3-4, as described by Meehan *et al.* (2017:2), is traditionally used in organisations. This technique keeps the entire BI architecture regarding development in mind and is more applicable to the purpose of this study.

3.2.4 Dimensional modelling

After the ETL process has been finalised, the next step to consider is the appropriate dimensional modelling design to be introduced into the DW structure. This structure is the general structure used by the relational database schema.

According to Kimball and Ross (2015:206), dimensional modelling has the following steps:

1. Choose the business process. Select a department of the organisation to base the dimensional model on, for example sales or commissions.
2. Declare the grain. The grain is the object inside the business process on which to focus.
3. Identify the dimensions associated with the grain. This will be descriptive information such as date or store.
4. Identify the fact. This will be the numerical values associated with the grain, such as price or fee percentage.

After identifying the business process and grain, a full list of dimensions associated with the business process needs to be identified. The fact table is the measurable event. The dimensional tables differ from fact tables in the sense that dimensional tables only consist of characteristic information such as client name, product name, date of birth and address to describe the attributes, with one primary key linked to a fact table. After associated dimensions to the business process have been identified, fact tables related to the business process should be determined, along with the content that needs to be included in them. According to Gosain and Singh (2015:305), the fact table is a numerical descriptive database table that is the primary table used in a star schema model. Fact tables consist of only numerical information such as amounts and many-to-one foreign keys indicating relationships between fact and dimensional tables. Fact tables do not have a primary key but at most have a surrogate key, which comprises a combination of many or all the foreign keys combined. After identifying all fact and dimensional tables, BI architects construct a bus matrix to associate the appropriate relationships needed between fact and dimensional tables before development of the DW can begin.

A bus matrix is considered by most BI architects as a blueprint to all business objectives required in developing a DW. Kimball and Ross (2015:16) explain that an important step to creating a DW, as shown in Fig. 3-5, is to create a bus matrix depicting a list of relevant facts (the business grain objectives) on the left down and mapping relevant dimensions (the business who, what, where, when and how objectives) associated with the respective fact.

	Date	Raw Material	Supplier	Plant	Product	Shipper	Warehouse	Customer	Sales Rep	Promotion Deal
Raw Material Purchasing		x	x	x						
Raw Material Delivery	x	x	x	x						
Raw Material Inventory	x	x	x	x						
Bill of Materials	x	x		x	x					
Manufacturing	x	x	x	x	x	x	x			
Shipping of Warehouse	x	x	x	x	x		x			
Finished Goods Inventory	x				x	x				
Customer Orders	x				x	x	x	x	x	x
Shipping to Customers	x				x		x	x	x	x
Invoicing	x				x		x	x	x	x
Payments	x				x	x	x	x	x	x

Figure 3-5: Bus matrix as adapted from Kimball *et al.* (2015:16)

After designing a bus matrix, it is important to design and complete a star schema model. The star schema model is used as a relational map of all dimensions and fact tables that will be implemented in the DW. Iyer and Lakhtaria (2017:64) visually present, in Fig. 3-6, an example of a simple star schema model that is developed before constructing the DW. The star schema model as shown in Fig. 3-6 consists of one primary fact table and all dimensional tables related in a one-to-many type relation. The '(SK) InvoiceID' is a surrogate key and column inside the 'Invoice_FACT' fact table that is comprised of all of the primary keys from the dimensional tables associated with the fact table, and is used to uniquely identify records inside the fact table.

After the construction of the star schema is complete, the next topic to understand in developing a DW is what to do when data in the DW tables are updated or replaced after the DW is implemented in a live environment, causing the dimensional table structure to change. Kimball and Ross (2015:25) describe an answer for this question by introducing a process called slowly changing dimension handling. While designing the DW structure, the DW architect needs to consider the process as to how slowly changing dimensions in the DW will be addressed. Kimball and Ross (2015:25) describe three types of slowly changing dimensions:

Type One - Overwrite. This is when the newly added information associated with the dimension overwrites the original information added to the dimension. The original information is replaced during this type of dimensional handling.

Type Two - Add a new row. This is when a new row of information is added below the original row of information, resulting in the original information not being replaced.

Type Three - Add new attribute. This is when a new attribute with data are created next to the original information and the original information stays as is.

The typical approach to cater for slowly changing dimensions differs between organisations.

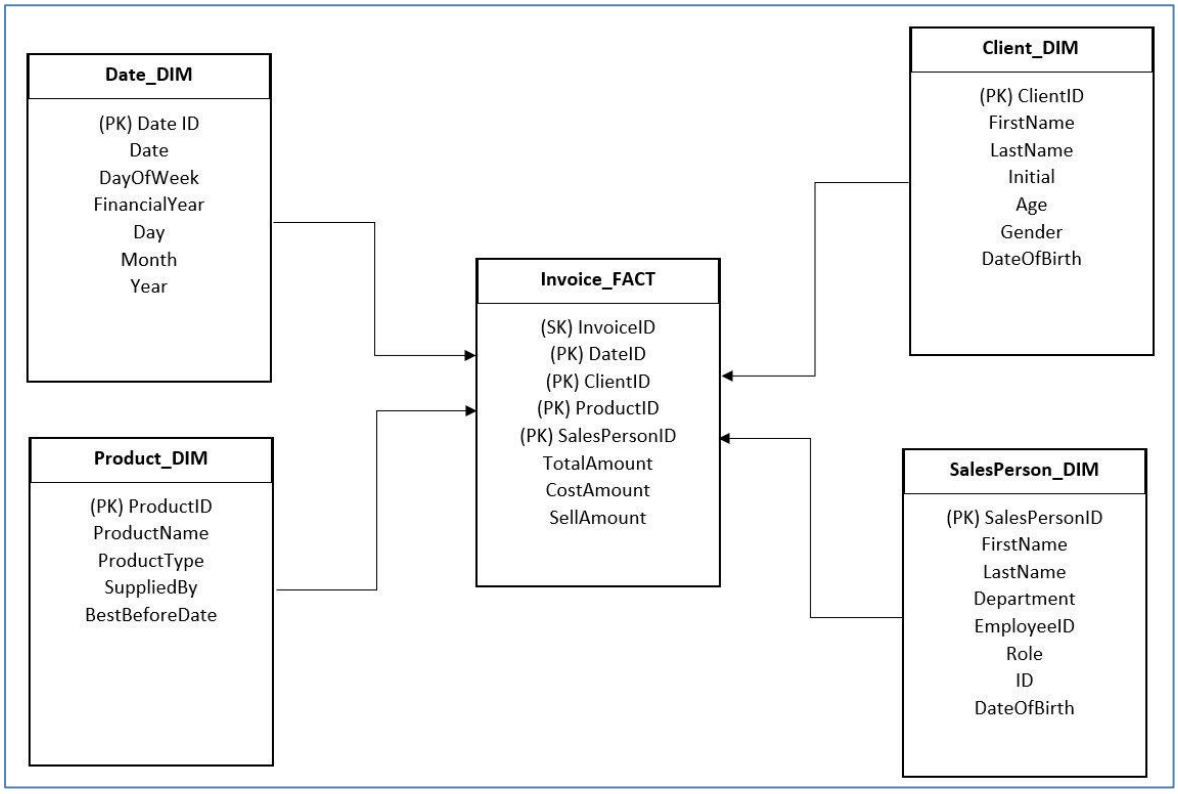


Figure 3-6: Star schema as adapted from Iyer and Lakhtaria (2017:64)

3.2.5 The Kimball methodology

The last step to consider before developing a DW, is to determine the DW dimensional design methodology to be used. There are two generally accepted modelling techniques in the field of DW development and design: the Kimball methodology (dimensional modelling) and the Inmon methodology (Entity Relational modelling). Kimball and Ross (2015:3) explain that the Kimball methodology for design and development of a DW starts with lower-level objectives and works up to higher-level objectives, beginning by identifying the grain, business rules, dimensions and facts. The Inmon methodology, however, focuses on designing the DW structure from the top

down, beginning with the highest-level objective and working towards the lower-level objectives. For this study, the Kimball methodology is focused on as the Kimball methodology is used in the study environment.

Kimball business intelligence lifecycle

The Kimball BI lifecycle needs to be taken into consideration at each phase of development of a DW when the Kimball methodology approach is used to design and development of the DW. Kimball and Ross (2015:33) visually express, as shown in Fig. 3-7, that a thoroughly developed DW begins with the programme and/or project plan drafted by the project manager and approved by the project owner and stakeholders. Next the project manager will get all business requirements needed to be implemented into the DW from the business users. The BI architect then conducts three phases of the BI lifecycle:

Phase 1 - The first step in this phase is to develop a technical architecture design to identify all components of the technical architecture needed and to map it out. The second step is to select the products that will be used to create and maintain the DW.

Phase 2 - The first step in this phase is for the BI architect to create the dimensional model associated with the DW, containing the dimensions and facts as prescribed by the business requirements defined. The second step in this process is for the BI architect to construct the physical DW model design. After the physical design, in the third step, the ETL design and development takes place. This maps out how the data coming into the DW will be cleaned and transformed into meaningful information that can be used by the DW.

Phase 3 - In the first step of this phase, the BI architect develops the appropriate BI design using the business requirements defined. The second step in this phase is to develop the BI application as designed by the BI architect.

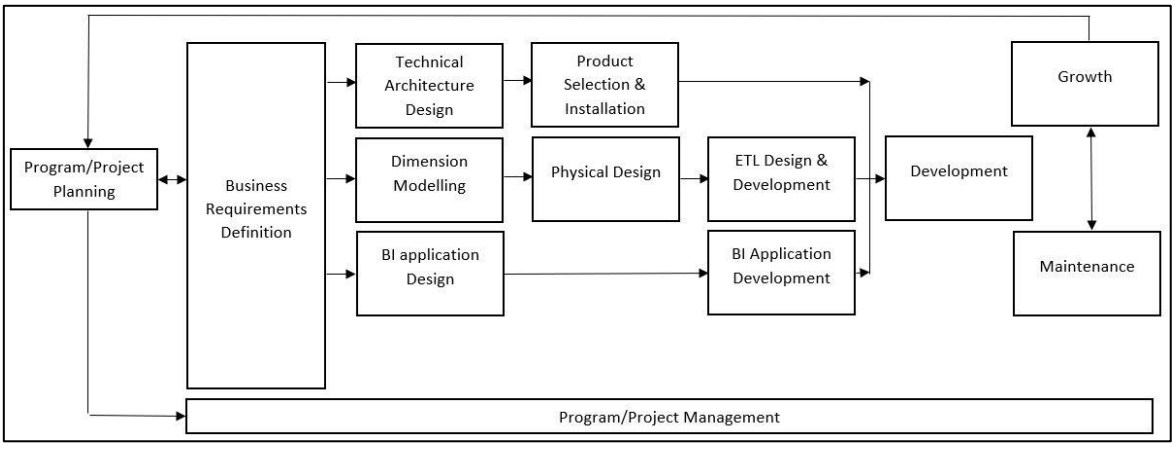


Figure 3-7: Kimball lifecycle as adapted from Kimball and Ross (2015:33)

After completing the three phases, the BI architect works with a BI developer to develop the DW. When development of the DW is completed, the BI developer is in charge of maintaining the DW from time to time. In addition, the project manager will make provisions for excessive data growth during the DW existence.

Gosain and Singh (2015:306) portray that designing a bus matrix is probably the most important step to perform when using the Kimball DW design. This will help to clarify any questions regarding the star schema design and identify all fact and dimensional tables associated with the DW design. After a DW is implemented into an organisational environment, the DW data flow process to keep data introduced into the DW up to date needs to be automated.

3.2.6 Data warehouse automation

Rahman and Rutz (2019:740) state that, after a thorough BI architecture has been implemented, an automation process needs to be developed to automatically reduce time for development and increase efficiency. Müller and Bundesbank (2019:2), in agreement with Rahman and Rutz (2019:740), describe that, due to an increase in data analytics, there is a need in organisations for data to be automated and available quickly. Introducing automation to DW processes eliminates the possibility of human error. During the DW design phase, senior DW staff may mistakenly forget to include the automation process. A typical automation process generally consists of five general concepts.

Instructions are then given by the automation process regarding how to change or digest the affected information inside the DW. Tomingas *et al.* (2015:8) explain that there are a few concepts to consider when identifying only changed data from the source databases, as follows:

1. Construct the appropriate data manipulation language (DML) statements to identify changed or added data from a singular data source.
2. Construct the appropriate singular mapping for all the different source data languages.
3. Identify and develop all possible scenarios to detect, analyse and load data into the DW.
4. Construct an automation solution based on all server query language (SQL) statement mappings.
5. Make sure the automation solution is easy to maintain and not too complex.

General indication key columns or values such as dates are used to identify and select only changed or added data. By taking note of these considerations, a thorough DW automation approach can be implemented.

3.3 BUSINESS INTELLIGENCE

According to Sauter (2014:2), BI architecture is a set of techniques and strategies used to transform data into meaningful information that will provide the organisation with accurate and insightful knowledge used for making business decisions. Business intelligence is a data optimisation solution used to make an organisation's data more accessible, resulting in more insightful business decisions being made (Azam & Tanweer, 2017:1324). Although most organisational managers make use of financial analytics, dynamic strategic decisions are normally missed by organisations that fail to prominently implement a BI architecture (Richards *et al.*, 2019:1). Enterprise decision-making of any organisation relies deeply on BI and analytics as a foundational cornerstone (Côrte-Real *et al.*, 2014:172). In agreement, Gupta *et al.* (2015:450) state that using BI tools such as data mining and predictive analytics, organisations can make more reliable business decisions.

Organisations all over the world are investing more resources in the design and development of comprehensive BI architectures. In a case study conducted regarding the effects of a BI system in start-up organisations, Azeroual and Theel (2019:38) state that BI solutions is an effective tool for successful start-up companies to gather every bit of increasing operational data and transform it into usable information. According to Giménez-Figueroa *et al.* (2018:113), enterprise environments are becoming much more dynamic with regard to the day-to-day production and commercial information being processed. Enterprise managers are looking more into BI to gain a competitive advantage and strengthen their customer outreach. Bodislav (2015:258) portrays that BI not only answers important business-oriented questions such as what the perspective is on long-term strategies and appropriate business decisions to gain a new level of understanding regarding day-to-day operations but can lower the costs of company operations and improve ease of access to a bigger area of operational data. Concepts relevant to the BI architecture, such as the BI evolution (section 3.3.1), BI maturity model (section 3.3.2) and BI lifecycle (section 3.3.3), are discussed in the subsequent sections. Furthermore, the analysis services (section 3.3.4) to develop a thorough analysis service associated with the BI architecture are discussed.

3.3.1 Business intelligence evolution

There is a big increase in organisations' desire for a deep analytic skillset in their employees and managers, which can only be met by upskilling current staff or investing in new employees with the needed skillset. Giménez-Figueroa *et al.* (2018:113) explain that organisations need to either upskill internal staff or employ additional staff to keep up to date with BI and analytic advancements and not to fall behind the competition.

Stefanovic (2015:912) states that, by using BI, users can perform advanced self-service analytical BI operations without the need of IS support. By 2016, most organisations already had access to self-service BI tools when analysing and manipulating information through the implementation of a BI solution (Oestreich, 2016:1). Then in 2017 it was said that BI has evolved to a state where it can be used as a framework in strategic business and management strategies (Alnoukari & Hanano, 2017:7). Business intelligence has also evolved to such an extent that it is being taught in universities all around the world. For any student who wants to follow a career associated with data mining, data analytics or business decision-making development of any kind, it is important to have up-to-date BI knowledge about the important fundamentals of DW design and the Kimball or Inmon methodology. Each certification stage has four areas of knowledge to consider, namely factual knowledge, conceptual knowledge, procedural knowledge and metacognitive knowledge. The most important skills to acquire are how to analyse, evaluate and develop BI.

Over the last few years, BI and analytics have grown to a large extent and the literature available on BI has increased. Côrte-Real *et al.* (2014:173) argue that the BI literature can be divided into four different diffusion states, namely adoption, implementation, use and impacts of use. While, in general, the knowledge on BI architecture is increasing in the educational environment, BI system knowledge might be implemented differently in organisations depending on their context and business needs. Most organisations use their BI architecture to integrate business strategies. In Fig. 3-8 below, Stockdale *et al.* (2015:263) visually express that BI architectures have evolved to the extent that different business strategies can form part of one BI architecture that leads to organisational effectiveness. Organisations can also have their own data structures that lead to organisational effectiveness. Through using this approach, organisations can make more effective decisions than before.

An example of this concept is that an organisation can have a strategy (called strategy A) where data from the ERP system is queried directly and displayed in a visualisation representation without data cleaning or sampling. This strategy will work if the information being retrieved back by the organisation does not require a high accuracy level, for example, number of customers in the organisation's database. The next strategy (called strategy B) an organisation can implement would be where data from the ERP system transfers through a BI architecture and is then displayed back to the organisation through visual representation, resulting in more accurate information being displayed back to the organisation. Strategy B performs better when a high accuracy level is needed in the information retrieved by the organisation, such as production or account information.

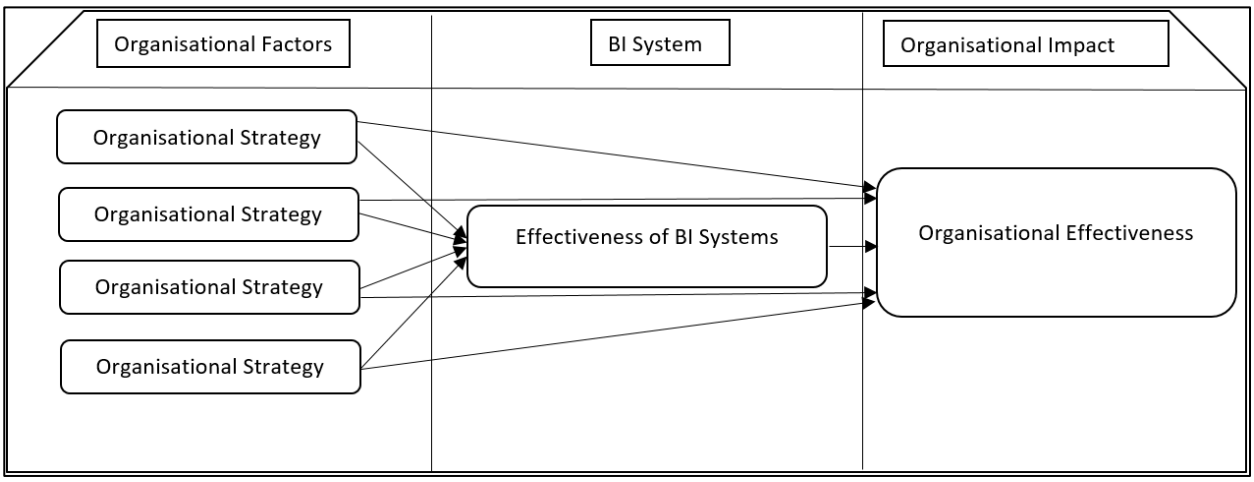


Figure 3-8: Business intelligence conceptual model as adapted from Stockdale *et al.* (2015:263)

While BI architecture is being developed in organisations and universities are upskilling students in BI concepts, the question remains: "What is the next logical approach to develop a BI platform on"? Theories regarding migrating an organisation's BI architecture with cloud-based applications have arisen during the last couple of years. According to Tole (2014:57), migration of a BI architecture into a cloud-based service has benefits such as cost and performance, but also has its limitations, the most concerning of which is privacy. While the organisation would be able to maintain a far less expensive BI architecture on a cloud-based platform than on a traditional on-premise server, the organisation's data might be in a higher sense of danger to privacy breaching or hacking.

3.3.2 Business intelligence maturity model

When an organisation decides to implement a BI architecture, it is best practice for the organisation to consider any future development and growth of the BI architecture as the organisation grows. This concept is called BI maturity. Business intelligence maturity is not only a technical model but a scientific model that impacts the organisation. According to Tavallaei *et al.* (2015:1006), developing an effective BI architecture or implementing a BI architecture can be quite complex, and even more difficult to successfully mature. Tavallaei *et al.* (2015:1015) visually present, as shown in Fig. 3-9, that a BI infrastructure can be measured in five levels of maturity percentages, namely:

Level 1 - Initial level (20%). The DW has very limited maturity and no expertise.
Level 2 - Immature level (40%). The DW has limited maturity and expertise.
Level 3 - Controlled level (60%). The DW has a medium level of maturity and expertise.
Level 4 - Managed level (80%). The DW has an advanced level of expertise.

Level 5 - The matured level (100%). The DW has highly advanced levels of maturity and expertise.

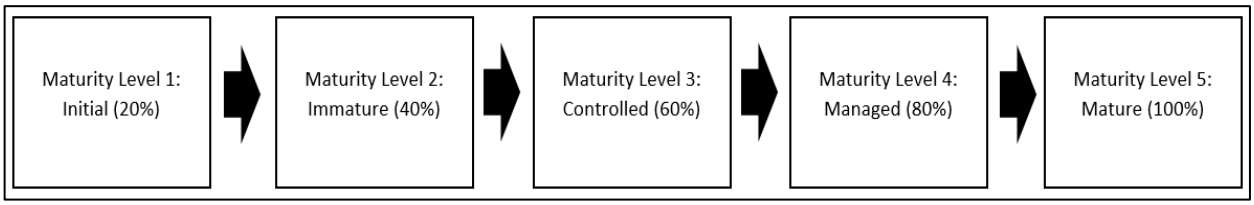


Figure 3-9: Business intelligence maturity model as adapted from Tavallaei *et al.* (2015:1015)

When an organisation has reached maturity level 5, the organisation can be comfortable in having a thoroughly designed and developed BI architecture.

3.3.3 Business intelligence lifecycle

After understanding BI maturity and how BI has evolved to its current state, the next logical step is to determine how the BI lifecycle flows from importing raw data from various data sources to presenting meaningful information back to the organisation (Kimball & Ross, 2015:729). The BI lifecycle can be defined as a lifecycle in which a BI architecture is implemented into an organisation, and data are transformed into meaningful information and presented back to the organisation. This includes, but is not limited to, retrieving data from a single CRM source, multiple CRM sources or from the production databases, implementing the ETL process to clean and transform the data, the data analysis process to analyse data, and then visually presenting the data from the DW via BI applications to the organisation for decision-making purposes. The characteristics of a well-designed Microsoft BI architecture include data sources from various locations ingested into a central DW, the data are then added placed through an ETL process where the data are cleaned from anomalies and structured in a DW method (Anon, 2018a). The structured data are then placed in an analytical platform where additional business logic, data sampling and forecasting occurs. The data are then placed on a visualization platform and presented back to the organisation.

In an organisational environment, the BI architecture needs to be thoroughly designed and developed to cater for any changes to business operations that might have an impact on the BI architecture. Corporate management practices are heavily influenced by how well a BI architecture is implemented, which intermittently has a large impact on production effectiveness (Richards *et al.*, 2019:3). To build a well-designed BI architecture, Sirin and Karacan (2017:210) visually present, as shown in Fig. 3-10, the basic BI lifecycle, which consists of data that is

received from numerous sources such as different DBMSs, Excel sheets, and text files. The data are put through an ETL process where it is cleaned of all anomalies and corrupted data. The cleaned data are then loaded into a single DBMS instance. This is followed by a DW stage, where the data are loaded into a DW and converted into a dimensional structure, prepared to be analysed, and business rules are developed and applied. Lastly, the data from the DW is analysed and presented back to the organisation via statistical reports.

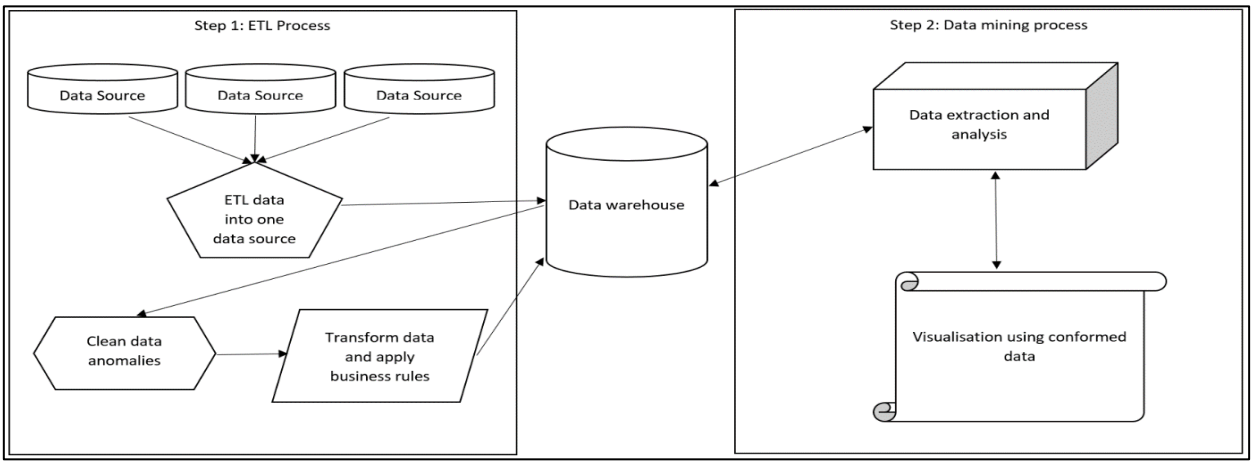


Figure 3-10: Data flow in a simple BI architecture as adapted from Sirin and Karacan (2017:210)

Stefanovic (2015:916) says that a BI lifecycle should be developed as shown in Fig. 3-11. The BI lifecycle begins by retrieving data from multiple data sources. An ETL process is implemented to clean and conform data gathered from multiple sources into one concrete state. After the ETL process is completed and the data are loaded into a DW, the data transfers through an analysis phase where data layers, business logic and data modelling are introduced. Finally, the completed analytical data are transferred into a BI web portal where the data are accessible for user self-service, research and employee collaboration.

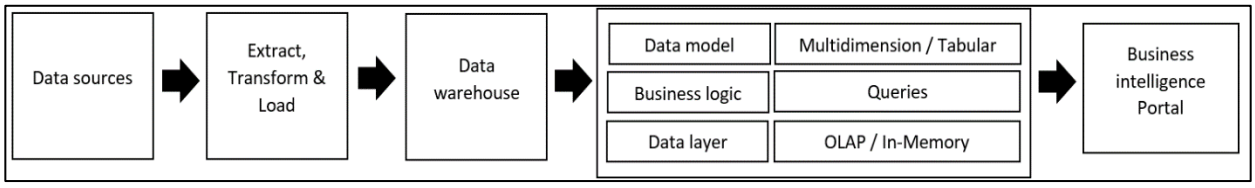


Figure 3-11: Business intelligence lifecycle as adapted from Stefanovic (2015:916)

Stefanovic (2015:917) describes that, through the BI lifecycle as seen in Fig. 3-11, BI-concentrated advantages emerge, such as:

- Data flexibility. In association with the analysis process of the BI solution, data in the BI solution is much more reliable and can adapt to change.
- Richness of data. By using key performance indicators (KPIs) and data mining techniques, business logic in the BI solution is better encapsulated.
- System performance. Data retrieval is more sufficive and scalable through the implemented BI solution. Organisations do not need to wait for organisational data to be retrieved. Business intelligence makes it possible for data to be retrieved 100 times faster for analysis purposes (Fields & Sheppard, 2012:2).
- Collaboration. Better collaboration between business and technical staff is gained.
- Usefulness. Alternative uses for business objectives can be identified through the BI solution's seamless self-service interface.

Łęgowik-Świącik *et al.* (2016:5), in contradiction to both Sirin and Karacan (2017:210) and Stefanovic (2015:916), visually express, as shown in Fig. 3-12, that the first stage of BI lifecycle starts with the business decisions that need to be made. Then, based on which business decision has been made, the ETL, DW, analysis and reporting stages can commence. After analysing the visualisations, if needed, the organisation can define the next business decision that needs to be implemented into the BI lifecycle.

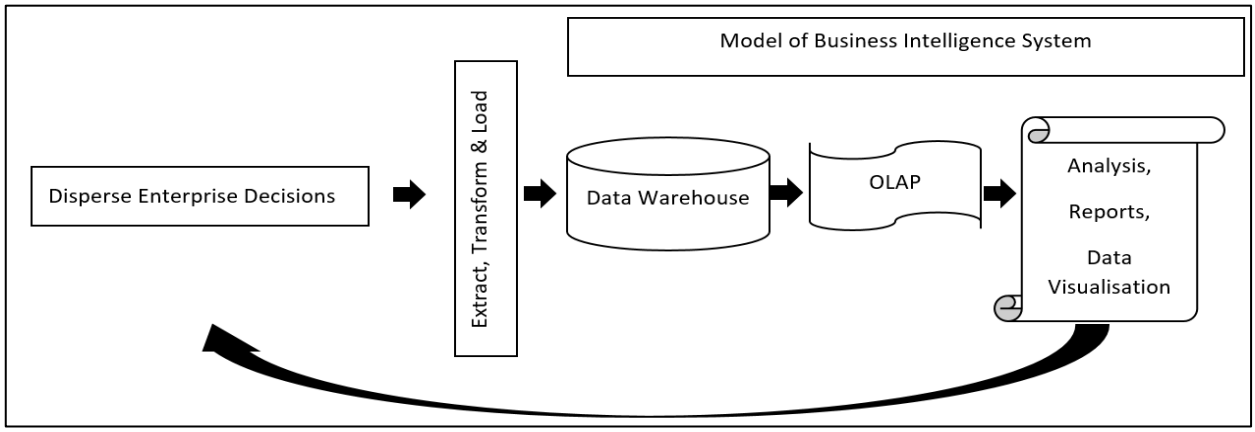


Figure 3-12: Business intelligence lifecycle as adapted from Łęgowik-Świącik *et al.* (2016:5)

Each BI lifecycle is well structured, but the most appropriate BI lifecycle of this study is depicted by Sirin and Karacan (2017:210) in Fig. 3-10.

3.3.4 Analysis services

The analysis services form an important part of the BI solution. The analysis services are used to define and create relationships, create additional business logic, create KPIs and create any additional measurements needed from the organisation. The analysis services are usually interlinked with the DW directly. Analysing data in an analysis service is normally the last step in the BI lifecycle, where data are transformed before it is displayed back to the organisation in a reporting platform (Gallinucci *et al.*, 2018:1).

When developing a Microsoft analysis service, there are generally two types of analysis service design methodology, namely multidimensional modelling using multidimensional expression (MDX) programming and tabular modelling that uses data analysis expression (DAX) programming. The organisation used in this study uses tabular modelling as an analysis services design methodology (Day, 2017:19).

Multidimensional modelling is the traditional analysis services design method used in organisations and focuses more on large amounts of data ingestion, such as transactional data (Day, 2017:20). The tabular modelling analysis services design methodology, on the other hand, is a newly developed method that is focused on processing smaller amounts of data. Over the last couple of years, organisations have tended to develop analysis services using tabular modelling as opposed to multidimensional modelling, due to the simplicity in design and development and because tabular modelling is compatible with a wider range of Microsoft products, including Microsoft Azure, Power BI and Microsoft Reporting Solutions (Ferrari & Russo, 2015:1).

When developing an analysis solution, the developer will first import the fact and dimensional tables, and then develop an associated cube. Rad (2014:1) visually expresses, as shown in Fig.3-13, an example of an analysis cube. The analysis cube can be seen in different layers, and the factual data can be sliced by dimensions such as date, product or store. Each layer in the cube depicts factual data in a specific date, store and product.

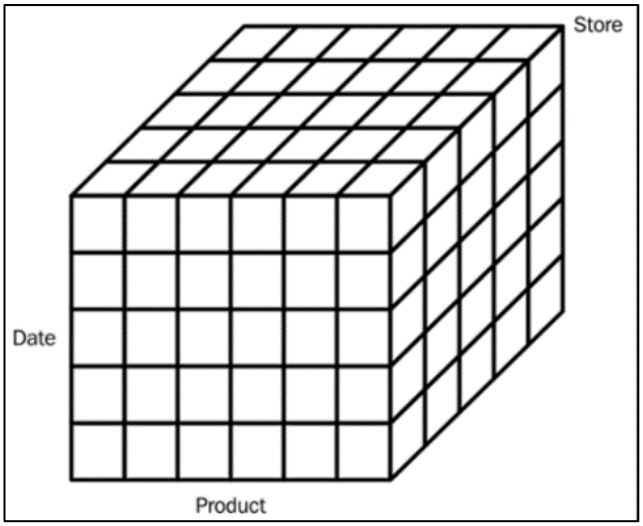


Figure 3-13: Multidimensional cube (Rad, 2014:1)

The multidimensional analysis services modelling development uses MDX code to retrieve and transform data situated inside the cube. Rad (2014:2) visually expresses, in Fig. 3-14, an example of MDX coding used in a multidimensional model in SSAS. The purpose of the measurements is to retrieve the “Sales Territory Countries” value from the “Sales Territory” table and “Order date year” value from the “Order Date” table associated with the “Internet Sales” cube.

```
select [Sales Territory].[Sales
Territory Country].members on
0,
[Order Date].[Calendar].
[Calendar Year].members on 1
from [Internet Sales]
```

Figure 3-14: MDX coding example (Rad, 2014:2)

Unlike multidimensional analysis modelling, tabular modelling uses DAX coding to retrieve and transform data. This form of coding is similar to the coding used in both Microsoft Excel and Power BI. Cuevas *et al.* (2017:3) visually present, in Fig. 3-15, an example of a DAX coding segment.

In the coding segment, a measurement is created to retrieve a count of all values in the “UnitStatus” column where the value in the “UnitStatus” column is either “Leased” or “Unknown”.

```
Sum of Occupied:=Calculate(
                        COUNT([UnitStatus]),
                        'LocationFact'[UnitStatus] = "Leased" || 'LocationFact'[UnitStatus] = "Unknown"
                )
```

Figure 3-15: DAX coding example (Cuevas, 2017:3)

More organisations are focusing on developing a tabular analysis services design methodology due to the simplicity of developing and easy-to-understand facilities associated.

3.4 DATA MINING

Data mining can be defined as when data that is already in the database is investigated to find meaningful and logical data patterns (Witten *et al.*, 2016:6). Roiger (2017:5), in agreement with Witten *et al.* (2016:6), portrays that data mining can be defined as the process of finding interesting and insightful patterns and trends in data structures. According to Rokach and Maimon (2008:2), data mining occurs when data scientists examine and converse through large amounts of data to find interesting patterns of truth. Data mining can occur on medium to large amounts of data for forecasting and analytical purposes. Data mining can be defined in four levels, namely facts, concepts, procedures and principles. Concepts to understand in association with data mining, such as data scientist (section 3.4.1), data mining lifecycle (section 3.4.2) and data mining phases (section 3.4.3), are discussed in the subsequent sections.

3.4.1 Data scientist

From the 1990s, organisations and educational services have looked to quantitative data for decision-making to strengthen organisational weaknesses. This is where a data scientist comes into play by looking at the organisation's data, analysing the data and building reports from the analytical data to present insightful information back to the organisation (Agasisti & Bowers, 2017:184). Granville (2014:9) states that, unlike statisticians or engineering analysists, a data scientist is someone who can implement an ETL tool on raw data, drill through the data and analyse the data to present meaningful information back to the organisation.

Parks (2017:9) portrays that the term 'data scientist' can be defined as an individual with the skillset to solve analytical problems, opposing the boundaries of statistics and disciplines. The skillset needed from a data scientist includes, but is not limited to:

- Data analysis functions. A data scientist should be able to analyse raw data and data structures to present analysis in a visual form, using analytical and mathematical methods.

- Tools and programming. A data analyst needs to be able to use analytical tools required by the organisation or develop an analytical tool as needed to analyse organisational data.
- Machine learning techniques and solutions. A data scientist should be able to build machine learning algorithms to analyse and present data as needed.
- Interdisciplinary knowledge. A data scientist should have mastered academic disciplines such as economics, accounting, maths and IS to correctly analyse organisational data.
- Critical abilities. A data scientist needs to be able to communicate and summarise organisational data for presentation.

When an organisation has staff members with the skillset as mentioned by Parks (2017:9), the organisation can be sure to conduct concentrated data mining techniques.

When a data scientist is approached to perform data mining, the data scientist should follow a set of phases to analyse the organisation's data. Agasisti and Bowers (2017:189) visually express, as shown in Fig. 3-16, how a data scientist goes about analysing an organisation's data in the following steps:

1. Collection and acquisition. The data scientist collects requirements such as business decisions that need to be answered, in addition to determining the business objectives that need to be investigated.
2. Storage. The data scientist acquires raw data from the organisation's DBMS.
3. Cleaning. The data are cleaned and prepared for use.
4. Integration. Multiple data sets are integrated into one source of truth.
5. Analysis. Multiple analysis tools and techniques are applied to the datasets to get clear insight into the data.
6. Representation and visualisation. The data scientist develops visualisation representations, such as queries or reports of the data analysed.
7. Action. The representations are analysed by the managerial staff to identify the appropriate actions to be taken.

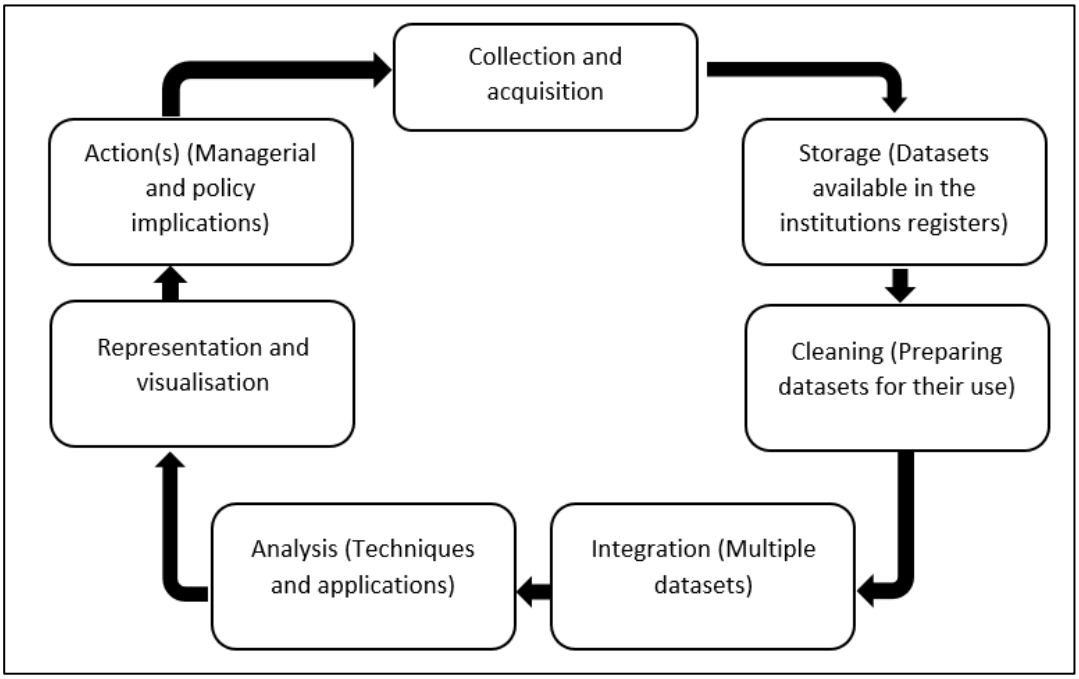


Figure 3-16: Data scientist work flow as adapted from Agasisti and Bowers (2017:189)

A data scientist needs experience dealing with big data, the ability to detect insightful knowledge from big data, and the ability to identify faults or to ask additional questions (Granville, 2014:3).

3.4.2 Data mining lifecycle

Data mining justifies a large section of a data scientist's skillset, but before being able to mine data for analytical and forecasting purposes, it is important for the data scientist to follow and complete a thoroughly developed data mining lifecycle plan. Witten *et al.* (2016:29) visually presents, as shown in Fig. 3-17, how the data mining lifecycle flows:

Phase 1 - Business understanding, by identifying the business objective you want to achieve.
Phase 2 - Data understanding, by understanding how the data flows in the DW and knowing what data will hold meaning.
Phase 3 - Data preparation, using machine learning algorithms to process the raw data identified in the data understanding phase.
Phase 4 - Modelling, by visual models developed to find new insight into the data.
Phase 5 - Evaluation, to identify whether the data mining is complete and complies with all business objectives requested.

If the evaluation is a success, the data mining models are deployed. If not, the whole process starts over.

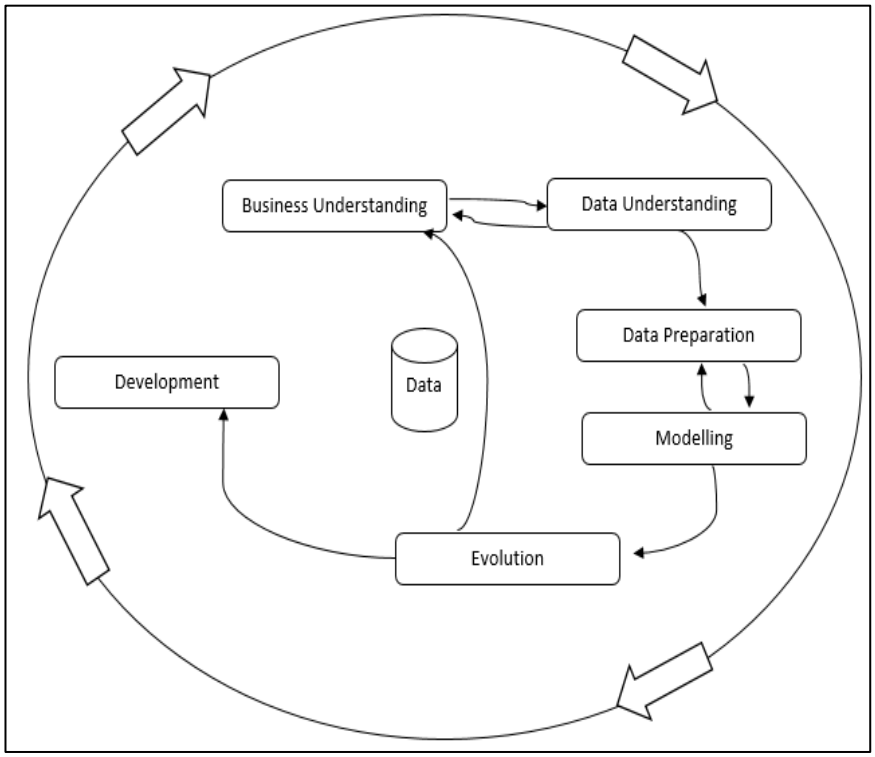


Figure 3-17: Data mining lifecycle as adapted from Witten *et al.* (2016:29)

By following the data mining lifecycle as mentioned by Witten *et al.* (2016:29) in Fig. 3-17, the data scientist can perform data mining tasks such as forecasting and analytics.

3.4.3 Data mining phases

After knowing what the data mining lifecycle consists of and before the data scientist can deliver a thorough forecasting or analytical report to the organisation, the data scientist needs to follow data mining phases. Stefanovic (2015:918) visually expresses, as shown in Fig. 3-18, the data mining phases as follows:

Phase 1 - The data are manipulated using a data mining management system as follows:

1. Data analysis chooses the right algorithm to be able to build the appropriate model.
2. The data are trained over the data model to find any anomalies in the data.
3. By using the data model, the selected data are tested using the selected algorithm. If the model test is successful, predictions regarding the data can be performed.

Phase 2 - All the predictions are then displayed in a visualisation technique as reports.

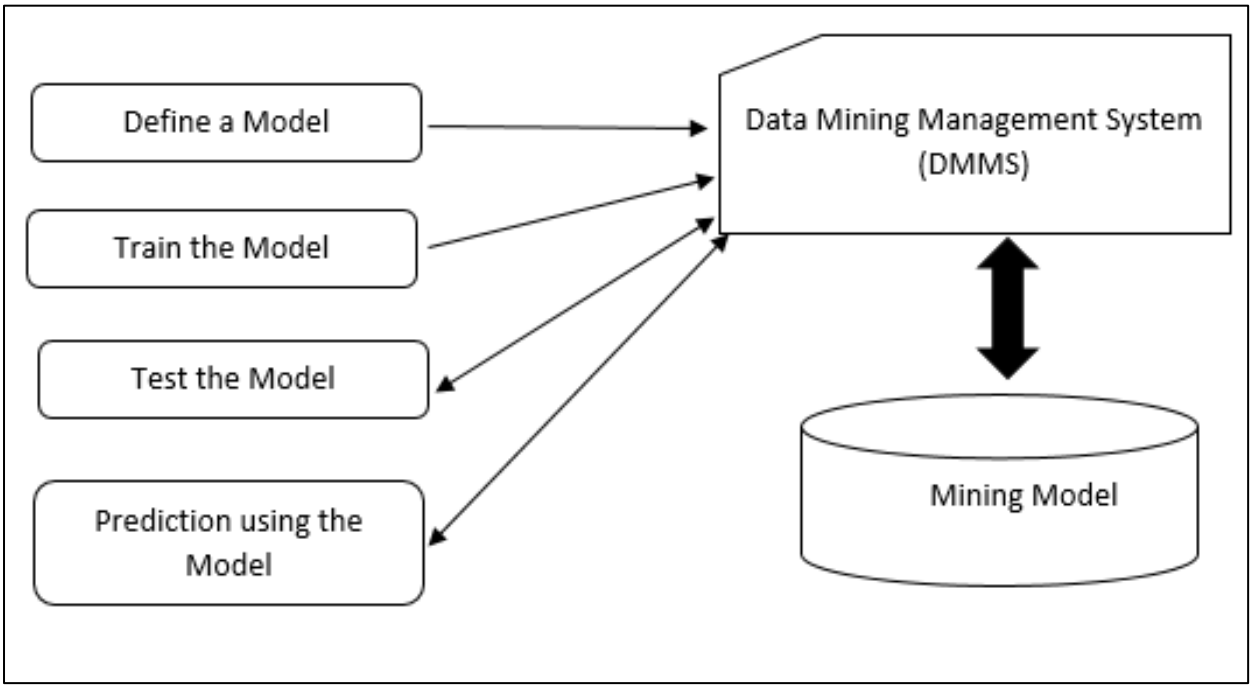


Figure 3-18: Data mining phases as adapted from Stefanovic (2015:918)

After the data mining phases have been identified, the next logical step is to know what level of data mining to conduct. Rokach and Maimon (2008:3) visually present, as shown in Fig. 3-19, the four phases that data mining can be used for:

Phase 1 - An appropriate learning algorithm is selected to use in the data mining process.

Phase 2 - The algorithm is then used with the raw data to develop an appropriate model, such as a decision tree, artificial neural network or partitioning model.

Phase 3 - The appropriate tasks, such as classification, regression or clustering tasks, are developed.

Phase 4 - After the tasks have been developed, they can be applied to answer business objectives such as fraud detection, demand planning or credit scoring.

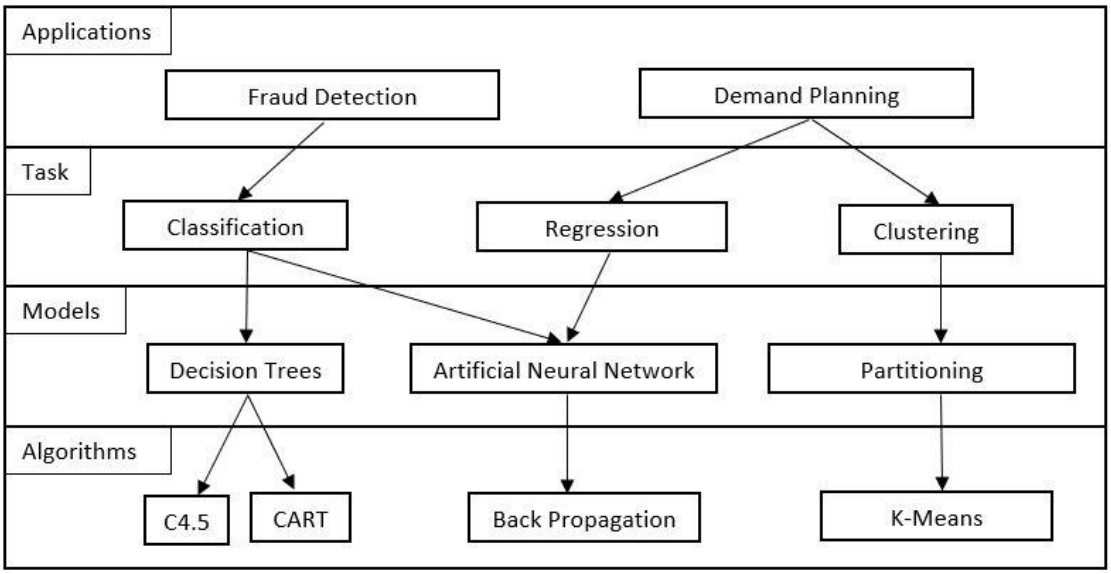


Figure 3-19: Phases of data mining as adapted from Rokach and Maimon (2008:3)

Rokach and Maimon (2008:5) later describe that the steps needed to be performed to develop an appropriate data mining application, as seen in Fig. 3-20, are the following:

1. Identify the business objectives and knowledge discovery in databases (KDD) goals needed from the organisation.
2. Select raw data from the databases to use in the data mining process.
3. Clean the selected data sample of any data anomalies and faults.
4. Process the selected data sample using a machine learning algorithm.
5. Develop a model from the processed data.
6. Interpret and evaluate the model to find answers to questions raised in the beginning.

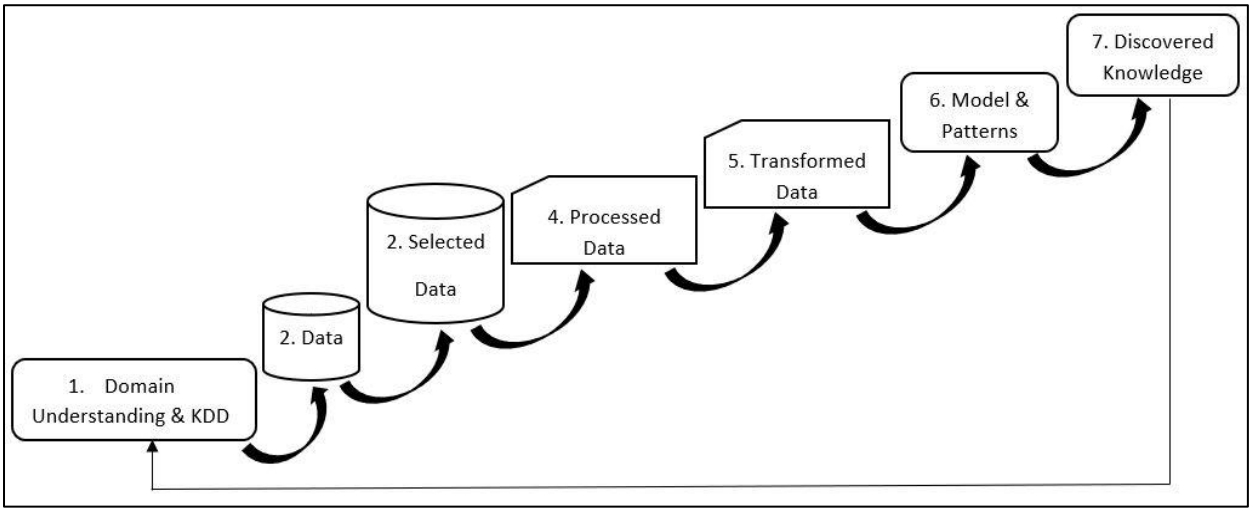


Figure 3-20: Data mining phases as adapted from Rokach and Maimon (2008:5)

The explanations of data modelling phases as presented by Rokach and Maimon (2008:5) in Fig. 3-19 and Fig. 3-20 are much more detailed than that of Stefanovic (2015:918), shown in Fig. 3-18. As a result, Fig. 3-20 is a more detailed approach to develop a well-structured data modelling lifecycle.

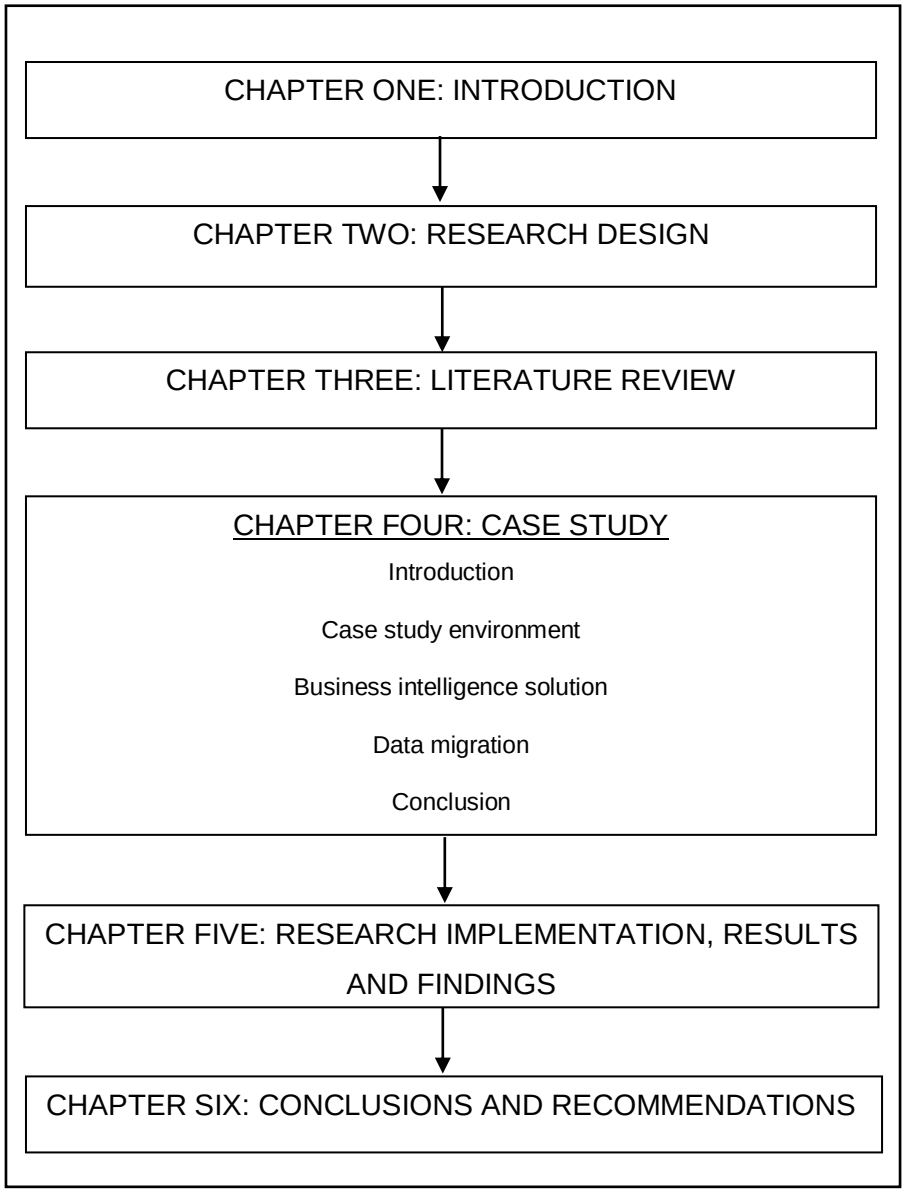
3.5 CONCLUSION

Business intelligence and analytics play a large part in decision-making in organisations. If a BI architecture is correctly implemented into a growing organisational environment, the organisation is expected to benefit from the analytical decisions made possible by the BI architecture. If an organisation introduces a BI architecture without an appropriate framework or implementation plan, it could result in misleading information, leading to a quest for additional resources such as costs, development time and staff needed. Theoretical objective T01 was presented in this chapter by means of a supporting literature review to describe all concepts key to the research, namely the evolution of a BI architecture, defining a DW, and research conducted about BI architecture in growing organisations.

This chapter discussed the literature associated with DW design and development. Components of DW design and development such as the DW maturity model, ETL process flow, dimensional modelling and the Kimball methodology were also discussed in this chapter. Business intelligence components such as the evolution of BI, BI maturity model and BI lifecycle were also included in this chapter's discussion. Lastly, data migration concepts such as the data migration phases, ETL process and data mining were also discussed.

In the next chapter, the case study is discussed in more detail. The BI architecture implemented into the study environment when the company was of start-up capacity, the BI development project plan, the current BI architecture and data migration from one organisation's DBMS to another are discussed in the next chapter.

CHAPTER 4: CHAPTER FLOW



CHAPTER 4 CASE STUDY

4.1 INTRODUCTION

The main objective of the study was to compile guidelines for growing organisations to overcome challenges of data growth while implementing a thorough BI architecture. A descripto-explanatory case study was conducted to address this objective. The focus in this chapter is to discuss the descriptive part of the case study. The researcher aims to accurately describe a phenomenon. A technical specification document, developed by the organisation that is the subject of this study, and personal experience of the researcher in the organisation were used to guide this chapter (Anon, 2016). Interview excerpts, obtained from participant interviews, and applicable to the discussion in this chapter, were utilised as empirical evidence.

Empirical objectives addressed are describing the BI architecture of the small organisation in its first year of existence – E01; describing the newly implemented BI architecture – E02; and determining how data from two different database platforms (Microsoft and Oracle) were migrated to one BI architecture – E03.

The contextualisation principle of Klein and Myers (1999:72) is adhered to in this chapter by describing the case study environment, the BI architecture that was used when the organisation was of start-up capacity, and how the BI architecture implementation currently functions.

In the subsequent sections, the following topics are addressed: the case study environment (section 4.2), BI architecture (section 4.3), data migration (section 4.4) and conclusion (section 4.5).

4.2 CASE STUDY ENVIRONMENT

The organisation in which the research was conducted was started early in the 21st century, and quickly grew to become one of the largest competitors in the value added tax (VAT) recovery industry. Because of the organisation's growth and success in its ventures, it was possible for the organisation to acquire divisions of its direct competition, resulting in the organisation becoming one of the largest global leaders in the VAT recovery industry. The organisation is now providing VAT recovery services to more than 16 000 clients worldwide. Along with the increasing growth in the organisation's client base, came exponential growth in its data. The organisation soon realised that exponential data consumption would occur, and exclusive BI development was needed to support its growing business endeavour.

The organisation's daily operations include the retrieval of numerous invoices from clients traveling around the world for business purposes. After receiving these invoices, the organisation receives the funds from the associated VAT authority. The organisation procures a percentage of the funds for services and pays the remaining funds back to the client. The organisation receives large quantities of invoices from a variety of clients in different parts of the world daily. These invoices then need to be managed, checked for eligibility and then captured into the organisation's enterprise resource planning (ERP) system. Invoice reclaim eligibility differs between countries' VAT authorities.

The organisation developed an in-house ERP system. This ERP application was a desktop system designed by an in-house developer using c-sharp (C#) coding in association with object-oriented programming. Through the evolution of the organisation, a new version of the in-house ERP system was developed and implemented. The redesigned ERP system was developed by a consulting firm around 10 years after the original in-house ERP system had been implemented. The redesigned ERP system was designed and developed as a web-based application, using hypertext mark-up language (HTML) and JavaScript. Both the original ERP system and newly redesigned ERP system are still used in the organisation and read/write information straight to a live database in Microsoft DBMS. In addition to the organisation's ERP system, the organisation also has a number of additional applications being used by various departments inside the organisation that read/write information to their own respective DBMSs.

4.3 BUSINESS INTELLIGENCE ARCHITECTURE

The purpose of this section is to explain the BI architecture development and upgrading process of the organisation that was guided by an organisational compiled technical document (Anon, 2016). As seen in Fig. 4.1, the case study event flow of the BI architecture in the organisation that is described in this chapter is portrayed as follows:

1. Initial BI architecture. The first event that transpired was the BI architecture that was implemented when the organisation was of start-up capacity. This section addresses empirical objective E01 – to describe the BI architecture of the small organisation in its first year of existence. The subsequent concepts to be discussed in this section are the initial server structure and initial reporting structure implemented when the organisation was of start-up size.
2. Business intelligence development project plan. This was designed and implemented to upgrade the BI architecture in the organisation.

3. Newly implemented BI architecture solution. This section addresses empirical objective E02 – describing the current BI architecture implemented in the organisation. The subsequent concepts to be discussed in this section are the currently implemented BI server structure, data warehouse (DW), analysis services, reporting solution and automation process.
4. Data migration. This section deals with the migration of large data quantities from one organisation's DBMS database structure to another and how the organisation incorporated this large amount of data into the organisational BI architecture. This section addresses empirical objective E03 – to determine how data from two different database platforms (Microsoft and Oracle) may be integrated.

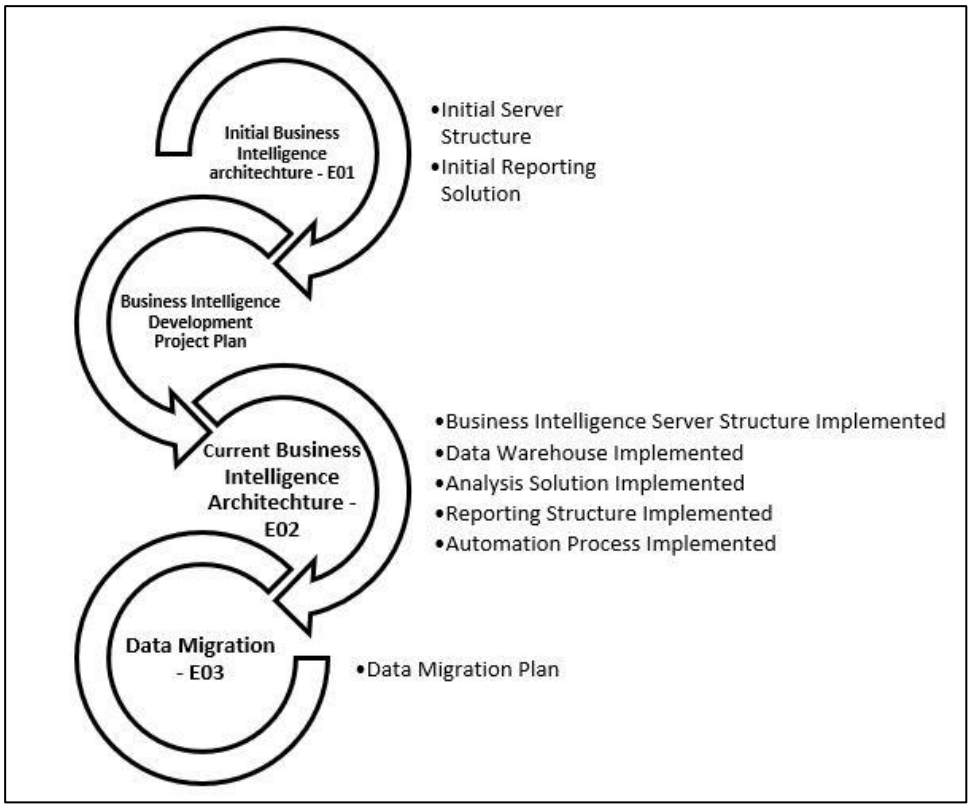


Figure 4-1: Case study event flow

In the subsequent sections, the researcher discusses the initial BI architecture (section 4.3.1) that describes the BI architecture that the organisation used when it was of start-up capacity; BI development project plan (section 4.3.2) developed to improve the BI architecture; currently implemented BI architecture (section 4.3.3) providing knowledge of the final BI architecture that was implemented by the organisation; and the automation process (section 4.3.4) to keep the BI architectures data up to date.

4.3.1 Initial business intelligence architecture

In this section, the study aims to address E01 by describing the BI architecture integrated into the organisation when the organisation was still considered to be of start-up size.

4.3.1.1 Initial server structure

In the first three years of the organisation's existence, the amount of data collected and transformed daily was of medium-size capacity and, as such, the organisation had a simple, yet hard-to-configure, server structure. The organisational server structure, as depicted in Fig. 4-2, used an ERP system to capture invoice information from clients. The information was then stored in a live database server. Afterwards, information was distributed via replication to three servers: the replicated reporting European server for reports used by the organisation's Europe-based employees; the reporting internal server used by the employees based in South Africa; and the MySQL server that was used by clients all over the world for cloud-based reporting.

The organisation used a Microsoft DBMS application called Microsoft SQL Replication. The purpose of this software was to replicate an exact up-to-date snapshot of the organisation's entire operations database structure and data from the host server to another. This was a continuous process, and if the replication process was interrupted by any means, the entire replication process would fail and would have to start over again.

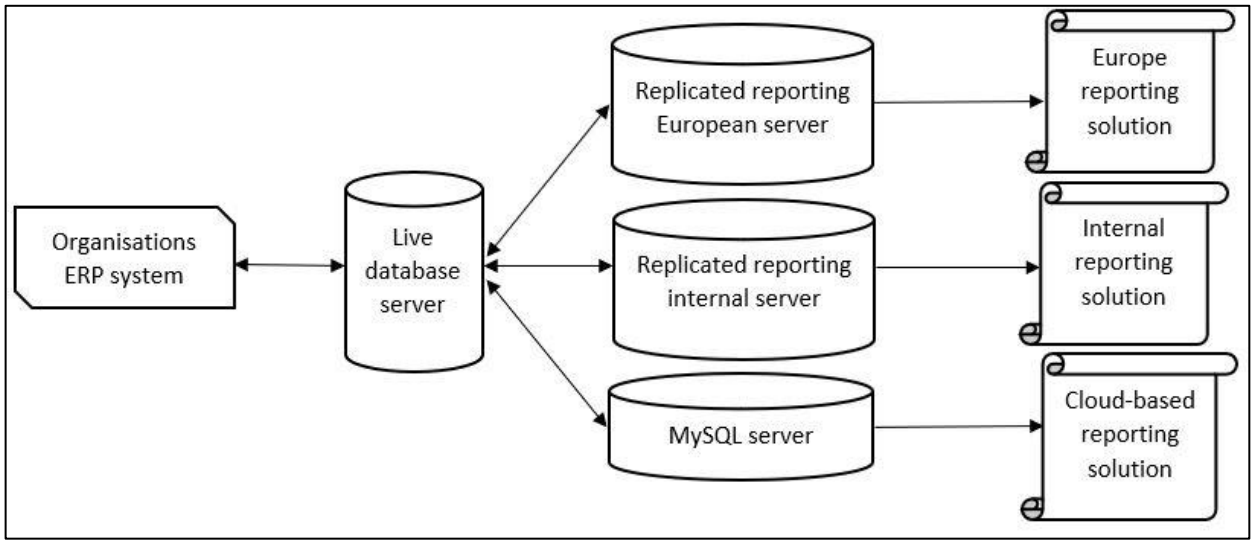


Figure 4-2: Old server structure

As the organisation grew, the amount of data being passed through to the live database server, as depicted in Fig. 4-2, increased, which led to complications. Some of the complications experienced by the organisation were that various reporting solutions tended to take longer to

generate and ERP application malfunctions or network failures occurred. Furthermore, inconsistent and incorrect critical information was displayed back to the users on the different reporting solutions as a result of reporting solutions retrieving information from different database servers. The organisation urgently needed to develop and enhance the BI architecture implemented to overcome the complications.

Table 4-1 explains that the BI solution that was implemented when the organisation was of start-up capacity, utilised a Microsoft Access DBMS and then upgraded to Microsoft SQL DBMS.

Table 4-1: Management – Old BI solution

Participant:	Yeah so, everything was running off an Access database and I think they used a, I forgot the name, they used the standard access front-end, its got a name but I forgot it. That was just for the first couple of year that the business has been around but pretty soon after that the size of the data grew beyond the scope of what Access can handle, and that when SQL server was implemented. SSRS reports was then generated from the SQL server database.

4.3.1.2 Initial reporting solution

When the organisation was of start-up capacity, the reporting structure of the organisation was straightforward and easy to understand. The reporting solutions were directly linked to the individual database servers, using transactional server query language (T-SQL) to directly query data from the database servers and generate various individual reports.

The organisation mainly used Microsoft SQL Server Reporting Solutions (SSRS) and Microsoft Excel to generate both the internal reports and the European client reports (Table 4-1). The organisation also used an internal cloud-based reporting solution that was generated by a uniquely built website, developed with tools such as HTML and JavaScript.

The organisation soon realised that, with the growth of invoiced transactional data being integrated from the ERP system into the live Microsoft DBMS, critical complications such as reporting data inconsistency and redundancy occurred. An additional concern to the organisation was that the reporting solution would run for significantly longer, and on most occasions would time out, causing the organisation to have to re-run reports. These concerns urged management to implement an up-to-date operational process. Senior management decided that the BI architecture needed to be improved to manage the additional amount of transactional information captured daily.

4.3.2 Business intelligence development project plan

The extension of the BI architecture needed to accommodate central up-to-date, accurate, clean data without any anomalies, and furthermore, to cater for any or all data growth that would occur in future. When this BI architecture was planned and developed, the organisation decided to use internal staff that had the required skillset to develop and integrate a BI architecture into the organisation's current BI architecture.

The organisation used the software development lifecycle, an agile project management technique. An internal project manager was appointed for the project, along with internal BI architecture and development staff. Figure 4-3 below depicts the timeline in which the BI architecture project was completed, with all associated phases:

Phase 1 - The initial understanding of project scope, project timelines and costs.
Phase 2 - Business requirements and reporting needs were collected from the organisation.
Phase 3 - The new BI architecture was designed and developed by the BI development staff.
Phase 4 - The developed BI architecture was deployed to a testing environment where selected business staff performed data quality checks.
Phase 5 - The BI development team provided feedback on data inequality issues and prepared the version of the BI architecture for final deployment into the organisation's live production environment.
Phase 6 - The final version of the BI architecture was deployed to the live operational environment of the organisation. The BI staff then provided for additional support and report changes as and when needed from the organisation.

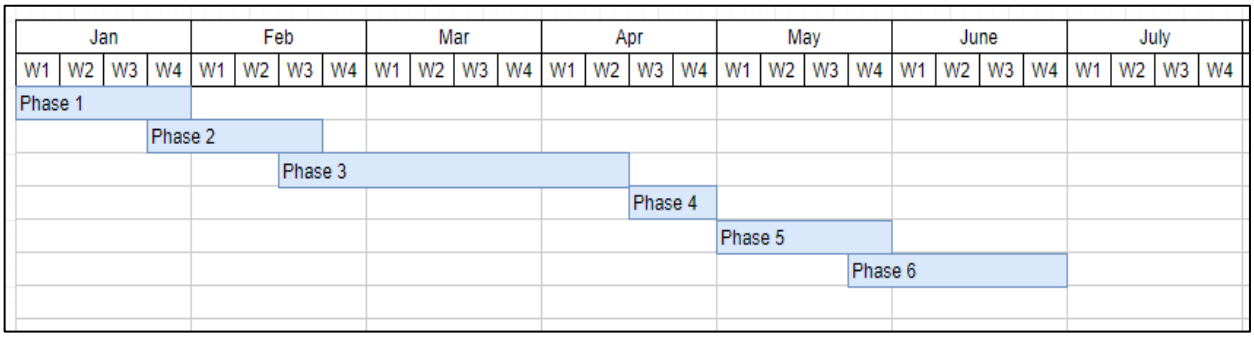


Figure 4-3: Business intelligence architecture Gantt chart

The project management team believed that, with the BI architecture progress timeline as set out in Fig. 4-3, the BI architecture project could be completed within the six-month completion requirement.

4.3.3 Current business intelligence architecture

In this section, the study aims to address E02 by identifying the currently implemented BI architecture. Subsections on the newly implemented BI architecture, such as the BI server structure, DW, analysis solution, reporting solution and automation solution implemented are presented.

4.3.3.1 Business intelligence server structure implemented

The BI server structure includes the entire BI architecture and all associated components of the organisation's implemented BI architecture. Fig. 4.4 describes the BI lifecycle (section 3.3.3). This is in accordance with the technical specifications document as implemented by the organisation A (Anon, 2016:8). All organisational ERP systems read and wrote from the live database server. The data were then transferred into a central DW server where the ETL process took place (section 3.2.3). The ETL process cleaned the data inserted by the ERP system from all anomalies and duplications and transformed it into meaningful information. The information was then stored inside the allocated dimensional table structure. The information was transferred into an analysis services server where relationships between dimensional and fact tables were formulated. Business rules were applied to the data according to each division's requirements in the organisation. Additional measurements and calculations were added where needed by the organisation. The information could then be retrieved by internal, Europe-based and cloud-based reporting solutions.

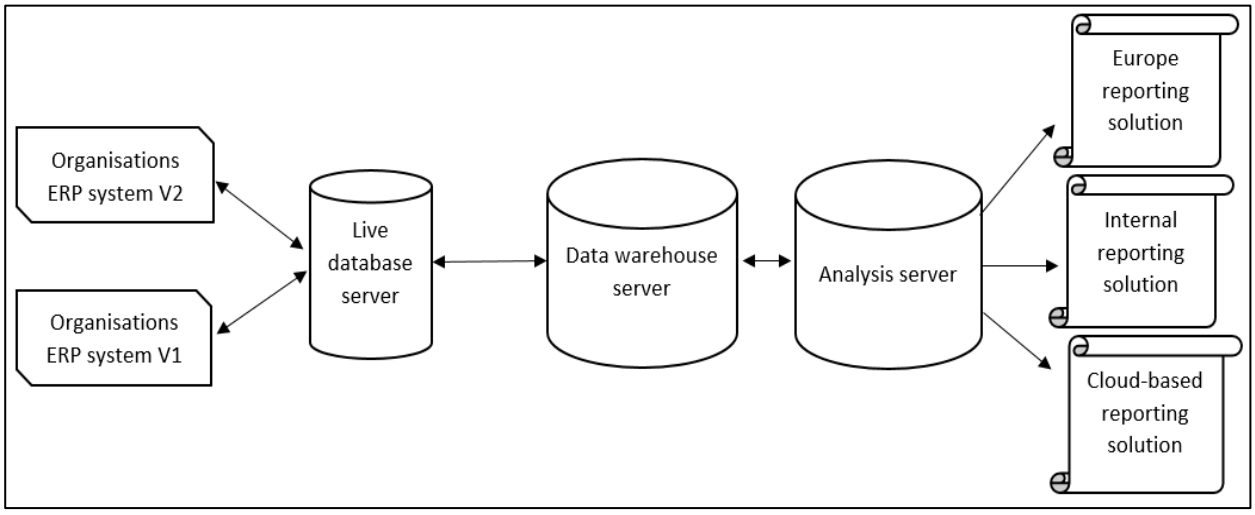


Figure 4-4: New server structure

With the newly implemented server structure, as shown in Fig. 4-4, the reporting solutions generated accurate information back to the organisation faster, with one source of the operational information shown across all reporting solutions.

4.3.3.2 Data warehouse implemented

The DW was developed according to the Kimball methodology approach to DW development (Kimball & Ross, 2015:33), as described in Chapter 3 of this dissertation. In Fig. 4-5, the currently implemented DW, in accordance with technical specifications for organisation A (Anon, 2016:8), is depicted with the data flow steps:

1. The data were transferred straight into the landing environment with the original table structure from the associated data source environment.
2. The data were transformed into a dimensional table structure and transferred to a staging environment. The data inefficiency and impurity were also kept, as the focus of this phase was to re-structure the originally normalised data structure into a dimensional data structure.
3. In this step, the data were cleaned, integrity checks were implemented, all data redundancies were removed, business rules were applied, and any data discrepancies were removed. The transformed information was then introduced into the live database structure.

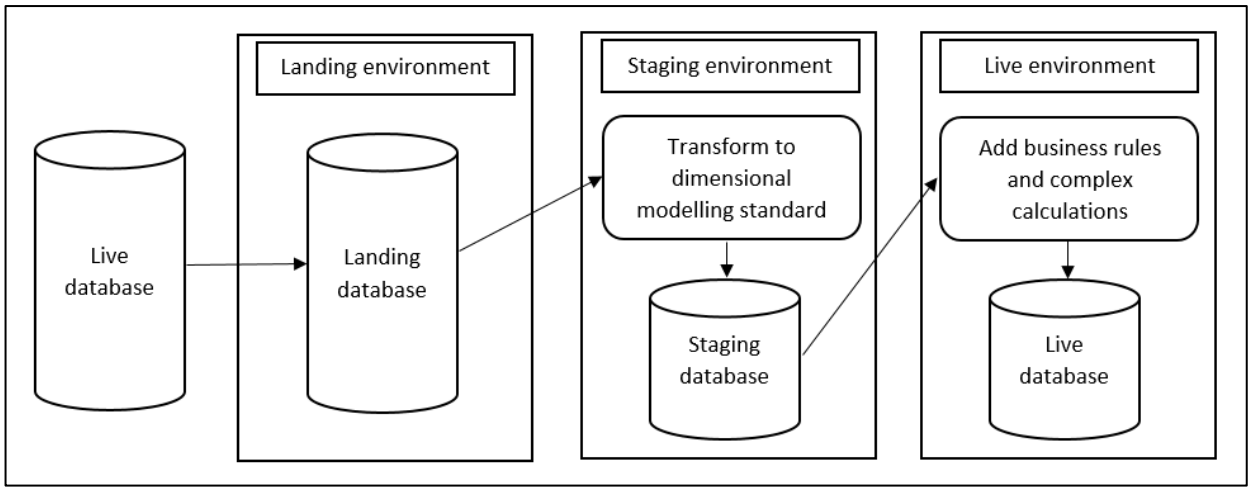


Figure 4-5: New data warehouse data flow structure

After the data flow process had been completed, the data were considered by the organisation as meaningful, accurate and reliable information to further be used in the analysis structure section of the BI architecture.

The above claims are supported by empirical data (Table 4-2). A participant stated that the newly implemented BI architecture implemented consisted of a DW solution, that all data sources are kept in one central location. The data from the DW was then transferred to an

analysis server cube were additional business requirements are implemented. The data was then retrieved by various reporting solutions.

Table 4-2: Management – New BI solution

Participant:	Well, I believe it was through the use of a data warehouse, the reason behind this is to have one central place to keep all source data such as Excel reports. An analysis services cube was placed on-top of the data warehouse for our business logic to be build into. And then the reports were built off of the cube in the relevant sectors of the business.

4.3.3.3 Analysis solution implemented

After the data had been transformed and integrated into the DW structure, the next step was to introduce the data to an analysis process. The analysis solution was designed and developed (section 3.3.4) in accordance with the organisation's technical specifications document as implemented by the organisation (Anon, 2016). A tabular analysis solution method was identified as the appropriate analysis solution method for the study environment.

The analysis solution introduced relationships between the dimension tables (that presented characteristic information) and factual table (that presented numerical data). Additional measurements, business rules and key performance measurements were added to the dimensional and factual structure by using data analyst expression (DAX) coding as expressed by Cuevas (2017:3) (section 3.3.4). The analysis solution was then processed, and the information was inserted into an analysis database. The analysis process is depicted in Fig. 4-6, with the following steps:

1. The first step in the analysis data process was to create relationships between the dimensional and fact tables in the DW's live database. This structure was then introduced into the analysis solution as a new analysis structure.
4. All table names and attributes in the analysis structure were cleaned and made user-friendly. This process was important to hide all key indication fields from business users, control the security of the data available to different priorities of business users, and provide fields available to business users that were easy-to-read and could be understood from a non-technical viewpoint.
5. Additional business-oriented measurements required from various divisions of the organisation were added to the analysis structured data. Normally, all the complex business rules were created in the DW data flow process, whereas the measurements created in the

analysis's solution would include simple calculations associated with the organisation's business requirements, such as aggregations and running total calculations.

6. Lastly, the analysis structure was integrated into an analysis database and was ready to be used for analytical and reporting requirements.

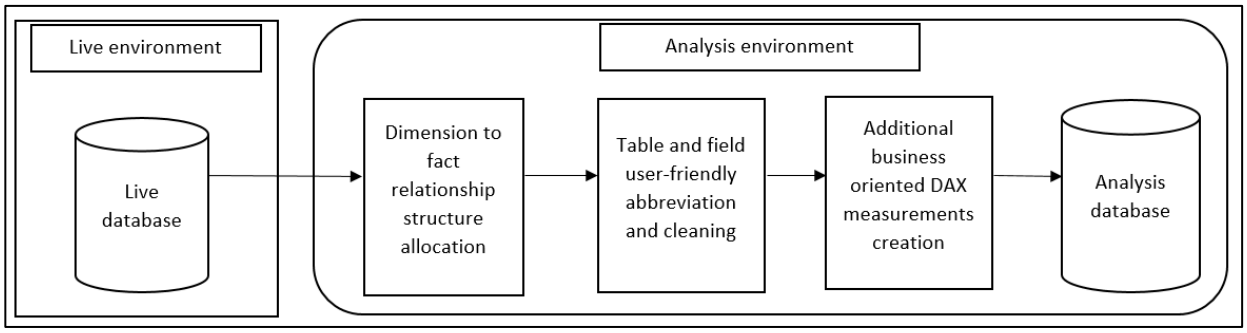


Figure 4-6: Analysis data flow

After the analysis data process had been completed, the analysis structure was implemented as depicted in Appendix A. As seen in the analysis structure portrayed in appendix A, the finalised analysis structure was amended with the key indication attributes greyed out, the non-key column names altered to be displayed in a user-friendly manner, and relationships between interactive fact and dimensional tables associated with the analysis solution presented accordingly.

The final analysis solution should present data to the user in a format as depicted in Appendix B. Key indicator fields are shown to be greyed out, attributes are displayed with user-friendly naming conventions and measurements are compiled as needed in the correct format. With the analysis structure integrated into the BI architecture, the organisation was confident that accurate, valued information was retrieved by the reporting solutions.

4.3.3.4 Reporting structure implemented

The reporting structure implemented by the organisation was used to retrieve information from the analysis solution and return it to the organisation's business users in a graphical and insightful manner in accordance with the organisation's technical specifications document as implemented by the organisation (Anon, 2016).The reporting data process is shown in Fig. 4-7. The information was generated in one of the following four reporting solutions, namely:

- Microsoft SQL Server Reporting Solutions. This type of reporting solution re-generated a line-by-line form like the representative report that the organisation uses for in-depth

analytical information analysis. The SSRS reporting structure could not be changed by users, only by the developer of the SSRS report.

- Client reporting website. This reporting solution was used by clients of the organisation across the world to view organisational information from an online platform. The client logged into a secure website using a secure username and password. Once logged in, the client was able to view analytical reports which displayed only information regarding the client's associated sections of business. This type of report was generated in a visual form, depicting user-friendly analytical graphs, generated using HTML and JavaScript. Client reporting website reports could only be changed by developers and could only be viewed by clients.
- Excel reports. This type of report was generated using Microsoft Excel. This type of reporting solution was used to generate either detailed row-by-row information or graphical information. The business user viewed or adjusts the report in Microsoft Excel according to the informational concepts of the business at the user's disposal. This reporting solution offered the business user more variety of information and ways in which to view the information.
- Power BI dashboard. This reporting solution was used to generate graphical analytics, displayed on a wide range of platforms such as desktop monitors, televisions and Android devices. Production users' viewed the progress of day-to-day production or financial processes. When viewing the report through a desktop monitor or Android device, the business user had the ability to create custom reports with a wide range of available information to compare and introduced to their unique reports or dashboards. When the business user viewed this reporting solution on a television or static platform, the business user could not alter the report in any way and only had viewing privileges.

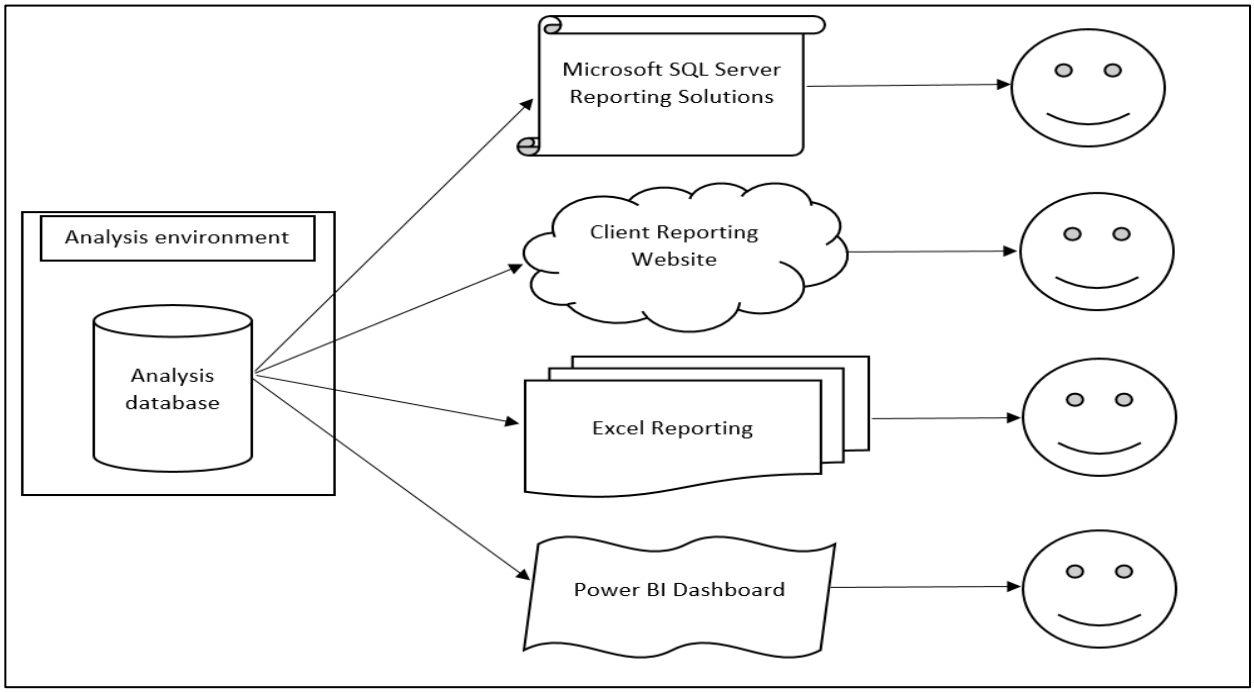


Figure 4-7: Reporting data flow structure

With the reporting data flow structure completed as shown in Fig. 4-7 and with the information displayed back to the business users as required, the data flow of the BI architecture was completed.

4.3.3.5 Automation process implemented

The automation process was implemented to make sure that the entire BI architecture solution could run automatically and with minimal or no user interaction needed in accordance with the organisation's technical specifications document as implemented by the organisation (Anon, 2016). This means that the BI architecture is always up to date with the latest organisational information. The concepts discussed in section 3.2.6 were used to develop the most optimal BI automation solution for the organisation.

The automated data flow process enables the organisation to keep information stored and controlled by the BI architecture up to date, and it comprises two automation structures. The first is an incremental automation process that occurs every 15 minutes daily, and the second is a historical automation process that occurs once daily. Figure 4-8 portrays the automation data flow process of the BI architecture that was developed through the following steps:

1. The incremental automation processes. In this process, only data that changes frequently is processed. The incremental automation process identifies data from the live production databases that has changed in the last 15 minutes. The incremental automation process

then executes the ETL process section of the BI architecture according to the changed data and inserts the data into an incremental category of the live DW database structure. The analysis structure processes incremental data only and stores the data in an incremental section of the DW analysis database structure. The reporting solutions structure then retrieves information from the incremental analysis database structure and displays the incremental information back to business users.

2. The historical automation process executes once daily and only after business hours, when there is less strain on the system performance of the operational server to which the BI architecture is deployed. The historical automation process runs the incremental ETL process first, receiving the incremental information loaded throughout the operational day into the DW live database environment. Thereafter, the historical automation process is executed to retrieve all other operational data captured throughout the day (that was not included in the incremental ETL process). The historical ETL process integrates the data received from both incremental and historical data flows into the central live DW environment while making sure that the data integrity and business requirements are implemented. Once the historical ETL process is complete, the historical analytical structure processes the historical data collected and incorporates the added or changed data into the historical section of the analytical database structure. Business users can then generate historical reporting solutions, which include the latest data changes that occurred a day prior.

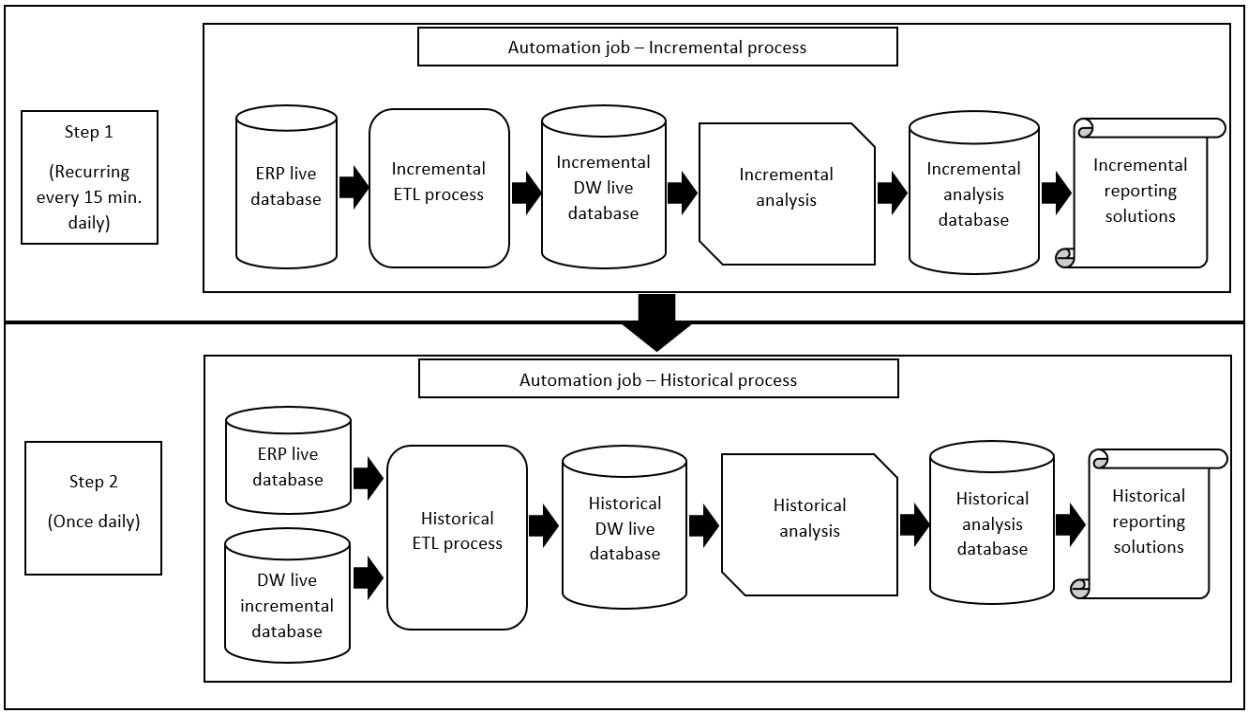


Figure 4-8: Automation data flow

With the automation process of the BI architecture completed and information generated through the BI architecture up to date, the organisation was satisfied with the upgraded BI architecture implemented to retrieve accurate, updated and reliable data for business decision-making purposes.

4.4 DATA MIGRATION

With a thoroughly designed and developed BI architecture implemented, the organisation embarked on another challenge. Organisation B was taken over by organisation A, and the data from both organisation A and organisation B was integrated into organisation A's DBMS and processed through a central BI architecture solution (Anon, 2016). This section, in correspondence to E03, details how the data migration plan was developed by organisation A.

Data migration plan

Organisation A was assigned the task of migrating data from one DBMS to another. Before organisation A performed this task, data migration factors were considered (section 3.4):

- The different DBMSs used – organisation A stored ERP data captured by business users in a Microsoft DBMS, whereas organisation B stored ERP data from business users in an Oracle DBMS.

- The differences in table structure and logical data flow between organisation A and organisation B.
- The quantity and complexity of data being migrated from organisation B into organisation A.
- The need for additional migration software to migrate data from organisation B's DBMS structure to organisation A's DBMS structure.
- The impact of the large quantity of additional data on organisation A's newly-implemented BI architecture.

With the associated factors considered, the data migration plan to import data from organisation B into organisation A's DBMS commenced, using the steps (section 3.4.1) depicted in Fig. 4-9:

1. Requirements gathering. The project management team, along with the database administrator (DBA) of organisation A, visited organisation B to analyse business processes, the Oracle DBMS structured data and organisation B's logical data flow.
2. This step consisted of two parts. During the first part, the data structure from organisation B's Oracle DBMS was converted to convey the Microsoft DBMS structure of organisation A by means of a third-party DBMS conversion tool. During the second part, the converted data structure was imported into a temporary Microsoft SQL DBMS database with the table structure logic equivalent to that of organisation B.
3. The data types and columns from organisation B were mapped to the temporary Microsoft DBMS database structure to comply with the data types and columns of organisation A. The data were then imported from the Microsoft temporary DBMS structure into organisation A's DBMS structure.
4. After the mapping was completed, the data from organisation B's DBMS database structure was imported into the mapped temporary Microsoft DBMS database structure.
5. The data from the temporary Microsoft DBMS database structure was then analysed and data checks were conducted to make sure that all data anomalies were corrected, avoided or removed.
6. A project signoff was done by organisation A's management staff after a full data import from organisation B to organisation A had been completed.

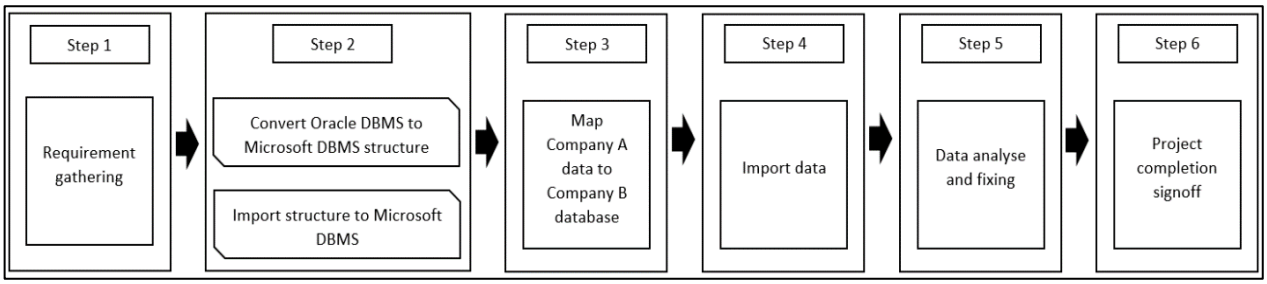


Figure 4-9: Data migration plan

The data migration project plan was initially scoped to not take any longer than six months. Figure 4-10 depicts the data migration project timeline that was completed with all associated steps, as follows:

1. The project management staff gathered all the requirements from the organisation from which the data were imported.
2. The DBA converted the Oracle DBMS data structure to a Microsoft DBMS data structure using a third-party DBMS conversion tool. The DBA then imported the converted Microsoft DBMS data structure retrieved from the third-party conversion tool into a temporary Microsoft DBMS database structure.
3. The DBA mapped the data and data types from organisation B's Oracle DBMS database structure to the temporary Microsoft DBMS database structure.
4. The data loaded into the temporary Microsoft DBMS database was imported into organisation A's production Microsoft DBMS database.
5. The data scientist staff analysed the data integrity of the imported data and amended any anomalies that occurred.
6. The final data fixes and support were conducted, and stakeholders signed off the project as completed.

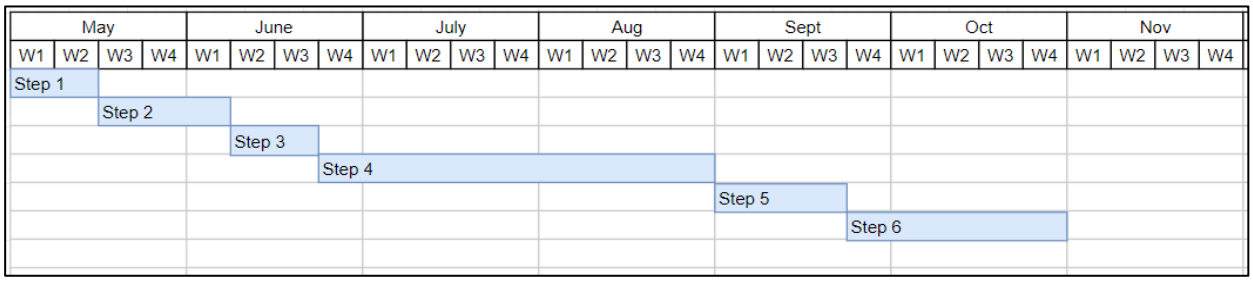


Figure 4-10: Data migration project Gantt chart

After the data migration process had been completed, organisation A's BI development staff designed a specialised ETL process to import the large amount of additional data into

organisation A's BI architecture. Organisation A's normal incremental and historical ETL automated BI architecture process was halted, and a specialised ETL process was created. The additional data were loaded into organisation A's BI architecture. After the specialised ETL process had been completed, organisation A's normal automated incremental and historical ETL process was restarted.

Empirical data listed in Table 4-3 supports the data migration process related above, from a developer's point of view. It relates the process as to how the data was integrated from the DBMS of organisation B into the DBMS of organisation A. The development team went to organisation B for a week to gather information surrounding the business rules of data to be imported, the mapping of the data, and the scope of work in order to integrate the data.

Table 4-3: Developer – Data Migration Preparation

Participant:	So, we went to the client for about a week before the whole project started to understand their business rules, to get the mappings, tried to understand what column A means on their side to what A means on our side. Also, just to get a list of how their business works.

In Table 4-4, the developer participant continued to describe the way data was integrated into the DBMS of organisation A. Organisation B used an Oracle DBMS, whereas organisation A used a Microsoft DBMS, thus making the table structure between the two DBMSs significantly different to understand and to map data between the two DBMSs. Once the data was mapped to the DBMS of organisation A, the data was then imported into Microsoft Excel files to be accurately mapped and balanced. The data was then integrated from the Excel sheets to the DBMS of organisation A.

Table 4-4: Developer – Data Migration

Participant:	Ok, so their database is Oracle, from there we got the Oracle DB into a normal SQL DB, according to their structure. From there we used Excel to do the actual mapping and export their data into a flat file where we did the actual mapping and balancing, and additional checks I would say. From Excel then we used SQL scripting to import.

A data scientist inspected the DW and BI solution of organisation A, in accordance with the phases of data mining (section 3.4.3), to identify and correct any mismatched or skewed data incorporated into the DW. After completion of the data integrity evaluation of the newly integrated data of organisation A, the organisation was able to confidently sign off on the completion of the data migration project.

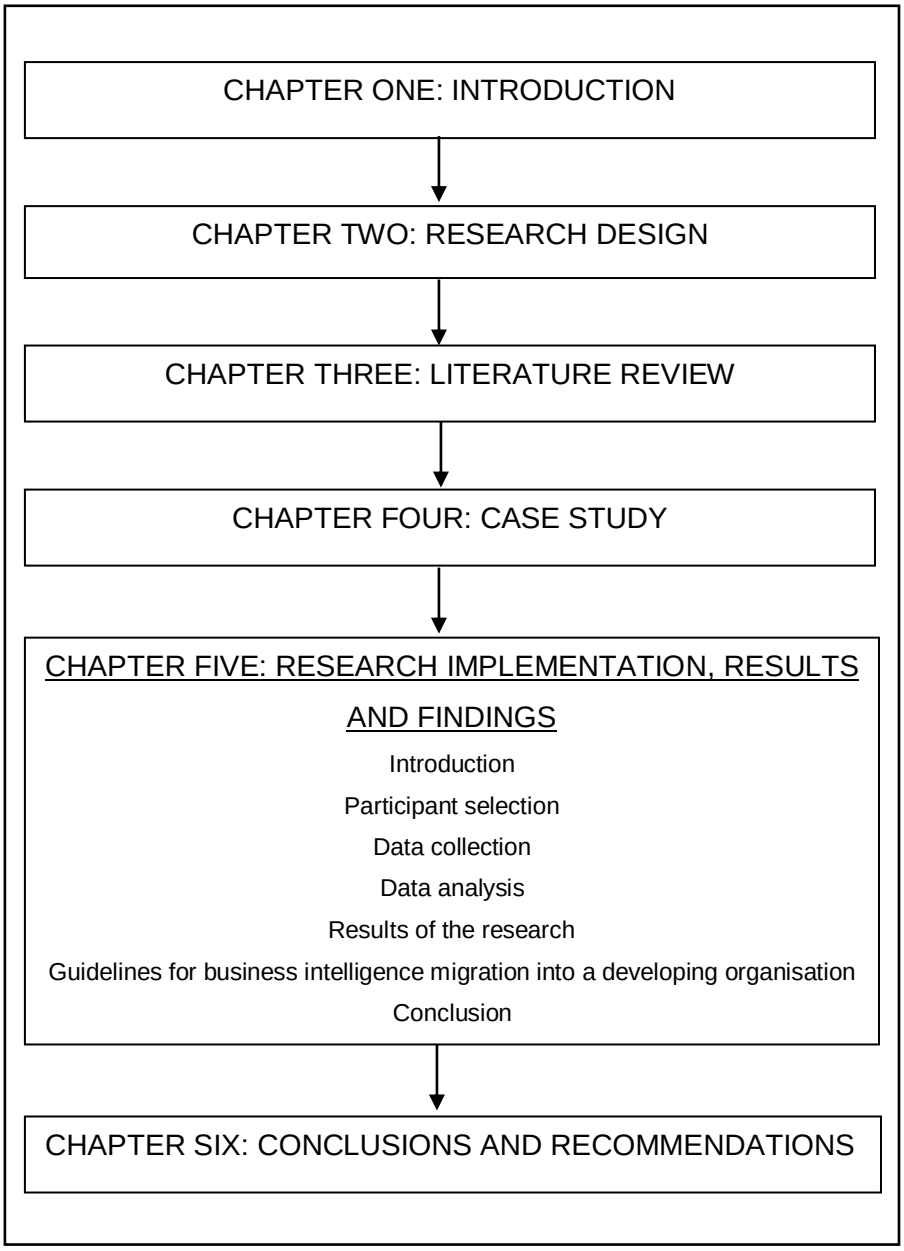
4.5 CONCLUSION

The case study environment used in this research was a singular phenomenon of a global leading VAT reclaim organisation in the industrial sector. The organisation grew from a small start-up business to a corporate organisation in just 10 years. With the growth of the production data, the organisation was forced to upgrade its BI architecture to a newer and more robust BI architecture. With the newly-developed BI architecture implemented, the organisation decided to purchase a competing organisation and was tasked to migrate the data from both organisations into one BI architecture.

In this chapter, empirical objective E01, the BI architecture implementation when the organisation was a start-up, was discussed. Thereafter, components of the current BI architecture, such as the BI architecture project plan, current BI architecture, current BI server structure, current BI analysis solution, current reporting solution and current automation structure, aligned with empirical objective E02, were discussed. Empirical objective E03 was also discussed, namely the components of the data migration project, where large amounts of data from the newly acquired organisation B were integrated with the original start-up DBMS of organisation A. How the large amount of data imported from two different organisations was migrated to the BI solution of organisation A, was also discussed.

In the next chapter, empirical objective E04 is discussed in more detail. The chapter discusses the participant selection process, data collection techniques, results of the data gathered and analysed from participants, and guidelines developed for BI migration in a developing organisation.

CHAPTER 5: CHAPTER FLOW



CHAPTER 5 RESEARCH IMPLEMENTATION, RESULTS AND FINDINGS

5.1 INTRODUCTION

In Chapter 4, the descriptive case study part was presented and therefore research objectives E01 to E03 were addressed. In this chapter the explanatory case study part will be addressed. The empirical study performed to describe the personal experiences and involvement of stakeholders in the course of migrating a newly developed business intelligence (BI) architecture into the organisation (E04), is discussed in this chapter. Participants' personal experiences and involvement throughout the BI architecture and data migration project were acquired through a qualitative data collection process. The literature review, case study outcomes and results of the empirical study were used to compile the guidelines for the migration of a BI architecture in a developing organisation.

In the subsequent sections, topics addressed are the participant selection process (section 5.2), the techniques that were used to collect qualitative data from participants (section 5.3), data analysis (section 5.4), the results of the research (section 5.5) and the guidelines for BI migration in a developing organisation (section 5.6). Finally, a conclusion to the chapter is presented (section 5.7).

5.2 PARTICIPANT SELECTION

The technique of purposive sampling, where participants are chosen because of their knowledge and capability to advance the purpose of the research, was used (Rule & John, 2011:64). The selection of participants directly involved and impacted by the study environment for this study was made from three categories, namely senior management staff, technical staff and data capturing staff. The reason for the three categories of participants was to gather information from a variety of different levels of participant experience and management, as well as to ensure that applicable data were collected.

5.2.1 Senior management

The senior management category portrayed personnel with several years of experience and knowledge of how the BI architecture was implemented when the organisation was of start-up capacity. Two senior management participants were selected to participate in the study. The first had joined the organisation when it started. The second manager selected for the study had joined the organisation just before the newly implemented BI architecture was developed.

Participants in this category had the most applicable high-level knowledge of business decisions made, the costs and scope involved with the project, as well as senior management processes involved in the implementation of a new BI architecture. The participants in this category normally generated high-level reports used for critical organisational decisions, such as financial reports.

5.2.2 Technical staff

The participants in this category were involved with technical aspects of the BI architecture and data migration project such as project management and database administration (DBA). There was a total of five technical staff participants used in this study. Three technical staff members selected for this study had joined the organisation when it was of start-up size. The other two technical staff participants selected for the study had joined the organisation soon before the BI architecture was upgraded. Participants associated with this category were staff members that designed, developed and implemented either the BI architecture or data migration from one organisation's database management system (DBMS) to another. Participants selected had in-depth technical knowledge of both the organisations' DBMS and data, as well as knowledge of BI principles.

5.2.3 Data capturers

The participants in this category engaged with the organisation's Enterprise Resource Planning (ERP) system. There was a total of five data capturing staff participants used in this study. All participants in this category had been employed by the organisation from the beginning. Participants in this category captured data into the organisation's ERP system and generated low-level reports such as production reports. Participants in this category had low-level operational experience.

5.3 DATA COLLECTION

Qualitative data were collected from participants through semi-structured interviews. Semi-structured interview questions were formulated based on the literature review associated with BI and data migration. The data collection approach is aligned with the principle of interaction between researcher and participant as stated by Klein and Myers (1999:72), where the researcher was exposed to the study environment and could relate to the participants' experiences by asking appropriate questions during the interview process.

5.3.1 Semi-structured interviews

An appointment was made each participant and a consent form explaining the research was send to the participant prior to the interview. Interviews were recorded with consent from participants. After each interview, the researcher compiled a summary of the interview, with added notes on behaviour.

Appendix C depicts the questions that were asked to the senior management participants regarding the BI architecture when the organisation was of start-up size. Further questions included asking them about the development of the BI architecture and the current BI architecture implemented. Appendix D portrays the questions that were asked to the senior participants regarding the data migration from one organisation to another.

Interview questions that were asked to the five technical staff participants were focused on technical concepts related to the BI architecture solution and data migration projects. Data gathered from these participants included technical aspects such as the planning, design, development and implementation of the BI architecture when the organisation was of start-up capacity, the BI architecture development project, the BI architecture implemented and the data migration from one organisation's DBMS structure to another. Appendix E describes the questions that were asked to these participants regarding the BI architecture. Appendix F portrays the questions that were asked to the technical participants regarding the data migration from one organisation to another.

Interview questions asked to data capturing staff focused on standard production concepts related to the BI architecture and data migration projects such as capturing and editing invoice information in the organisations ERP system. Data gathered from this category of participants was about generating and creating day-to-day operational reporting solutions. Appendix G lists questions asked to the data capturing participants.

5.4 DATA ANALYSIS

Data analysis conducted in this study followed the steps portrayed by Silverman and Bernard (Section 2.6). By using the literature reviewed (Chapter 3), the researcher prepared questions that were asked to the participants (as seen in Appendix C, D, E, F, G). The recorded interviews were transcribed and saved as text documents. A qualitative data analytical tool called ATLAS.ti was used to analyse the transcribed data. The objectives of the research as well as concepts related to the research and identified from the literature were used to identify codes and themes

(Kimball & Ross, 2015). Hofstee (2006:117) criteria to analyse data were implemented (section 2.6).

Content analysis was used to classify the transcribed interview data into code categories (section 1.6.3). Codes were identified from the transcripts as words or phrases that were best applicable to the study. Thematic analysis was used to identify themes from the code categories, which then were transformed into meaningful information to enable a conclusion of the study.

The relationship between themes allowed the researcher to understand, clarify and interpret the case (Rule & John, 2011:78). After all codes and themes had been established, a code network diagram for each category of participant (senior managers – Appendix H, technical staff – Appendix I and data capturers – Appendix J) was drafted, depicting the relevant codes and themes in a visual manner. The collection of codes and themes was exported into a Microsoft Excel spreadsheet to show the total responses for the different participant categories. This is aligned with the principle of the hermeneutic circle as stated by Klein and Myers (1999:72), which describes the manner that was used by the researcher to identify codes and themes from the data collected from participants.

The codes and themes as analysed from the transcripts are portrayed in Tables 5-1 to 5-5 below. The following themes were identified:

- Data migration: Comments received from participants relating to data migration from one organisation's DBMS to another. The participant categories that commented on this theme were senior managers and technical staff.
- New BI: Comments received from participants relating to the new BI architecture that was implemented in the organisation.
- Old BI: Comments received from participants relating to the old BI architecture that was implemented when the organisation was of start-up size. The participant categories that commented on this theme were senior managers and technical staff.
- New reporting: Comments received from participants relating to the new reporting solution in association with the newly implemented BI architecture used by business users. The participant categories that commented on this theme were senior managers, technical staff and data capturers.

- Old reporting: Comments received from participants relating to the old reporting solution implemented when the organisation was of start-up size. The participant categories that commented on this theme were senior managers, technical staff and data capturers.

Table 5-1: Theme identified: Old BI

Theme: Category and/or code	Management staff	Technical staff	Data capturer	Total
Old BI: Used an access database.	1	-	-	1
Old BI: Complications – Needed to build duplicate SSRS reports of the same data given different filters.	1	-	-	1
Old BI: Complications – Business rules differed between departments.	-	5	-	5
Old BI: Data received from reporting solution by direct query from database.	2	-	-	2
Old BI: Microsoft stacked solution was implemented.	1	-	-	1
Old BI: Reporting used – Excel.	1	-	-	1
Old BI: Reporting used – SSRS.	2	-	-	2
Old BI: Reporting did not reflect correct business data.	1	-	-	1
Old BI: Small technical team at beginning of organisation's existence.	1	-	-	1
Old BI: Used SSRS before BI solution was implemented.	-	3	-	3
Grand Total	10	8	0	18

Table 5-2: Theme identified: Old reporting

Theme: Category and/or code	Management staff	Technical staff	Data capturer	Total
Old report: Complications – Ineffective.	-	-	2	2
Old report: Complications – Not frequently used.	-	-	2	2
Old report: Complications – Not user friendly.	-	-	1	1
Old report: Complications – Oversee critical information.	-	5	-	5
Old report: Complications – System malfunctions.	-	-	1	1
Old report: Complications – Tedious.	-	-	2	2
Old report: Complications – Duplicated data.	-	-	3	3
Old report: Complications – Difficult reading.	-	-	4	4
Old report: Complications – Low productivity awareness.	-	-	1	1
Old report: Complications – Manual report needed and had to be generated multiple times.	-	-	2	2
Old report: Complications – Time-consuming.	-	-	2	2
Old report: Complications – Would sometimes not update.	-	-	1	1
Grand Total	0	5	21	26

Table 5-3: Theme identified: New BI

Theme: Category and/or code	Management staff	Technical staff	Data capturer	Total
New BI: Was developed by developers.	1	-	-	1

Theme: Category and/or code	Management staff	Technical staff	Data capturer	Total
New BI: Better graphical visibility of data in user reporting.	-	2	-	2
New BI: Central place to store data.	-	4	-	4
New BI: Clean data introduced to BI solution.	-	1	-	1
New BI: Complications – Business rules not very clear.	-	2	-	2
New BI: Complications – Build a BI maturity model to identify general problems.	1	-	-	1
New BI: Complications – Data quality check was a concern before implementing new BI solution.	1	-	-	1
New BI: Complications – Gathering stakeholder buy-in into BI solution project.	1	-	-	1
New BI: Complications – Understanding of BI in general.	1	-	-	1
New BI: Data are consistent.	-	2	-	2
New BI: Data are stored in one central location.	2	1	-	3
New BI: Development had help from external consultants.	2	-	-	2
New BI: Development by mostly internal staff.	3	-	-	3
New BI: Development phase 1 was to develop DW and some reports.	1	-	-	1
New BI: Development phase 2 was to design Power BI reports.	1	-	-	1
New BI: Development staff was originally DBA and data control staff.	2	-	-	2
New BI: Project development took six months.	2	1	-	3
New BI: Ease in reporting.	-	8	-	8
New BI: Extracting data are easier.	-	3	-	3
New BI: Failure notification implemented if ETL fails to populate data in DW.	-	1	-	1
New BI: Project took five months' development.	-	2	-	2
New BI: New BI is an improvement over the old BI solution.	-	1	-	1
New BI: One month to implement BI solution.	-	1	-	1
New BI: Positive business feedback.	1	-	-	1
New BI: Positive response from users.	-	5	-	5
New BI: Provides more accurate data.	1	-	-	1
New BI: Provides more information.	-	2	-	2
New BI: Provides more insight into data.	2	-	-	2
New BI: Reason for – Business relies heavily on data.	1	-	-	1
New BI: Reason for – Key business deliverable for year.	1	-	-	1
New BI: Reason for – Need for accurate data.	2	-	-	2
New BI: Reason for – Deeper insight into organisational data.	2	-	-	2
New BI: Reason for – To make better business decisions.	2	-	-	2
New BI: Reason for – To view data visually.	1	-	-	1
New BI: Provides more reliable, accurate data.	-	2	-	2
New BI: Reporting used – Power BI.	2	-	-	2
New BI: Reporting used is user-friendly and it is quick to build Power BI reports.	1	-	-	1
New BI: Reporting in a visualisation approach.	2	-	-	2
New BI: Roles – Decision makers and stakeholders.	1	-	-	1
New BI: Seamless integration into business.	-	4	-	4

Theme: Category and/or code	Management staff	Technical staff	Data capturer	Total
New BI: Simplified stable structure and logic.	-	1	-	1
New BI: Reporting used – SSRS.	2	-	-	2
New BI: Type 2 slowly changing dimension used for incremental data change.	-	1	-	1
New BI: Use of incremental ETL loads.	-	2	-	2
New BI: User ease of use to develop reports.	-	4	-	4
New BI: Users built unique reports on Power BI.	-	3	-	3
New BI: Works very well.	-	1	-	1
Grand Total	39	54	0	93

Table 5-4: Theme identified: New reporting

Theme: Category and/or code	Management staff	Technical staff	Data capturer	Total
New report: Convenient in new BI solution.	-	-	2	2
New report: Efficient.	-	-	2	2
New report: Higher production awareness.	-	-	6	6
New report: Higher production motivation.	-	-	3	3
New report: Higher user interaction with reporting solutions.	-	-	3	3
New report: Improvement over old reporting solution.	-	-	1	1
New report: Irregular information updating.	-	-	4	4
New report: Users “love it”.	-	-	4	4
New report: One central view of information.	-	-	7	7
New report: Precise data received in reporting solutions.	-	-	1	1
New report: Preference over old BI solution.	-	-	1	1
New report: Complication – Some reports had too busy design layout.	-	-	1	1
New report: Data updates frequently.	-	-	3	3
New report: User-friendly.	-	-	6	6
New report: Variety of information available.	-	-	1	1
Grand Total	0	0	45	45

Table 5-5: Theme identified: Data migration

Theme: Category and/or code	Management staff	Technical staff	Data capturer	Total
Data migration: Lack of preparation before data migration.	-	6	-	6
Data migration: Acquired external staff to help with data import.	2	-	-	2
Data migration: Actual project completion time took one year.	-	3	-	3
Data migration: Challenge in balancing financial data imported.	1	-	-	1
Data migration: Complications – Bad quality data.	1	-	-	1
Data migration: Complications – Data inequality.	1	-	-	1
Data migration: Complications – Different data structures.	2	-	-	2
Data migration: Complications – External staff not experienced enough to accomplish tasks.	2	-	-	2
Data migration: Complications – Implementation team under immense pressure during data migration.	1	-	-	1

Theme: Category and/or code	Management staff	Technical staff	Data capturer	Total
Data migration: Complications – Managing expectations to senior management.	1	-	-	1
Data migration: Complications – Manual way of bringing in data.	2	-	-	2
Data migration: Complications – Only one staff member in technical staff participant category is an expert.	1	-	-	1
Data migration: Complications – Staff needed to take shortcuts to meet deadlines.	1	-	-	1
Data migration: Complications – Understanding new business rules of other organisation's source data.	1	-	-	1
Data migration: Project completion time estimated at six months.	-	3	-	3
Data migration: Initial completion timeframe was less than a year.	2	-	-	2
Data migration: Initial timeframe for completion was not met.	2	-	-	2
Data migration: Skills needed to perform data migration – experience in both DBMS systems.	1	-	-	1
Data migration: Skills needed to perform data migration – data-oriented skills.	2	-	-	2
Data migration: The reason for migration was that organisation bought out another organisation.	2	-	-	2
Data migration: Project completion took longer than six months.	-	1	-	1
Data migration: Would include different data migration equipment and tools if migrating data ever again.	1	-	-	1
Data migration: Would allocate more time to plan and analyse data structures before data migration, if given the chance to migrate data again.	3	-	-	3
Data migration: Cardinality was different between the two DBMS structures.	-	4	-	4
Data migration: Created a custom ETL package to digest data into organisation's BI solution.	-	1	-	1
Data migration: Database structures differed between different organisations' DBMS solutions.	-	7	-	7
Data migration: Estimated development time – Six months.	-	3	-	3
Data migration: Help from external consultants was used.	-	3	-	3
Data migration: Complication – Kept project owners inside the loop.	-	1	-	1
Data migration: Complication – Limited project timeframe.	-	2	-	2
Data migration: Data digested into BI solution had little impact on organisation's reporting solution.	-	1	-	1
Data migration: One week of planning conducted before data migration began.	-	1	-	1
Data migration: Oracle to Microsoft conversion tool used for mapping.	-	1	-	1
Data migration: Complication – Performance issues when importing data into the systems.	-	1	-	1
Data migration: Had to prefix data to identify different source data.	-	3	-	3
Data migration: Complication – Source data was dirty.	-	4	-	4
Data migration: Used Excel to map data structures.	-	3	-	3
Grand Total	28	48	0	76

5.5 RESULTS OF THE RESEARCH

Results from the data collected from the senior management participants, technical staff as well as the data capture staff are discussed in more detail in this section.
The following principles of interpretivism stated by Klein and Myers (1999:72) were adhered to in this section:

- The principle of dialogical reasoning where the researcher revises the actual findings guided by sensitively identifying differences between preconceptions from participants.
- The principle of multiple interpretations where the researcher portrays different perspectives of participant in the same study in different forms such as senior management participants, technical staff participants and data capture participants.
- The principle of suspicion that required the researcher to be sensitive to possible distortions or biases in data collected from the participants.

5.5.1 Senior management participant results

This section focuses on the results received back from senior management participants regarding the old BI architecture, the new BI architecture implemented, and the data migration conducted from one DBMS to another. The senior management participants were asked more business-oriented questions as opposed to technical-oriented questions.

The first set of questions related to empirical objective E01. The intention of the researcher was to gain perspective on the experience and knowledge of the managers in interacting with the old BI architecture that was implemented when the organisation was of start-up capacity. Table 5-6 below provides the question and most direct answer from one participant. The participant described that the BI architecture implemented was initially hosted on a Microsoft-only framework, with a very limited amount of data being loaded into the SQL server database. An SQL server reporting solution (SSRS) reporting solution generated reports.

Table 5-6: Management – Old business intelligence solution that was implemented

Interviewer:	Ok, so then that leads me to question 1.2, what BI solution was implemented in the new organisation, how did you implement the BI solution?
Participant:	So initially, the staff compliment that was responsible for technology was very small and because of that limitation or restriction on resources, I believe, it was before my time but I believe that the decision was made to run with the Microsoft framework which at the time included SQL server with SSRS being the front end reporting platform that the users would engage with.

The next set of questions (associated with objective E02) focused on the reason for implementing a new BI architecture. The reasons for the reporting solution to be updated were various factors, such as the version of reporting solution being out of date and reporting software being incompatible with newer platforms displaying the reports, as stated in Table 5-7. The execution time to generate reports was long and the same version of a report generated in different reporting platforms, such as SSRS and the internet reporting solution, would present different information that did not balance. Lastly, the business users generating reports on the old reporting solution found the reports difficult to read and interpret.

Table 5-7: Management – Reason for new business intelligence solution implementation

Interviewer:	Ok, so then question 1.4 then leading up to the new BI architecture being implemented, what was the reason behind implementing a new BI solution in 2018?
Participant:	So, the reporting solution being used by the business is very out-of-date and takes long to open reports. The information displayed on reports differs on the different reporting platforms in place. The types of reports we had in place was very difficult to read as well.

The next set of questions discussed with the participants was about the newly implemented BI architecture, as shown in Table 5-8. One participant commented that the new BI architecture resulted in accurate data. The organisation was more confident in the type of information being generated in the reporting solution and the clean and reliable data. The participant also stated that, with the new reporting solution, business users could simply create new reports online interactively instead of having only a static set of reports available.

Table 5-8: Management – Contributions of the new business intelligence solution

Interviewer:	So with the newly implemented reporting solution will it provide more accurate information for decision making purposes?
Participant:	Yeah absolutely, the issue with the SSRS is that you just simply, instead of consolidating reports and building on them, there is a tendency to create new reports as new people join the business and as new processes come online and instead of consolidating the old reports, new reports gets created. So, with BI, because BI is kind of like a living organism the development time is spend on maturing a model rather than replacing it entirely. So yes, definitely more accurate, more consisted, with much less let's say distrust in the raw data.

The next topic discussed with participants in this category related to empirical objective E03, associated with the data migration of organisation B's DBMS into organisation A's current DBMS system. Table 5-9 portrays that the participant experienced many complications with the

data migration from one organisation's DBMS to another. One of the complications was to manage the expectations of stakeholders associated with the data migration project. The stakeholders had high expectations regarding time and deliverables, which put pressure on the development team to conduct the data migration process. Due to the large amount of pressure to deliver, development staff overlooked critical underlying data structures and business rules associated with the data, which caused the development phase of the project to be prolonged. The shortcuts that the development team thought they had accomplished turned out to prolong the process.

Table 5-9: Management – Complications with the data migration project

Interviewer:	Ok, then on question 2.4, if any, what complications were there regarding the data migrating project?
Participant:	There were many complications, I think the main one I would say was managing of expectations. So, there was an expectation by the senior management at the time that the project will last a certain amount of time, and so then those responsible for implementing it were under immense pressure to deliver. Because they were under so much pressure they had to look at a very manual way of bringing the data in instead of having the time to map out and understand that obviously there organisation worked differently then ours so there was a process of understanding there data and understanding there process which I think because of the pressure put on that, put on those people, they had to, lets say that they had to take some shortcuts or seemingly they were seemingly shortcuts but in fact they turned out to be long cuts.

The researcher then proceeded to questions regarding empirical objective E03 as described in Table 5-10 below. From the comments received, it appeared that it was very important that the large amount of data being introduced into the organisation's BI architecture should have minimal on impact the organisations production. Furthermore, the participants commented that, because of the robust data migration extract, transform and load (ETL) package developed, there was minimal impact on the organisation's production server and that the data migration into the BI architecture was completed without complications.

Table 5-10: Management – Stakeholder impression regarding the data migration into the organisation's business intelligence solution

Interviewer:	Ok, so then when the data was ready and merged into the organisation's BI solution through a custom ETL package, what was the overall impression and impact that the data load had on the organisation's BI architecture?
Participant:	So, as I said the business relies exclusively on the information provided by the BI framework, the business is extremely depended on it and is extremely sensitive to it so the impact was, even just with the vision of having a single version of truth consisted data, with a robust data migration ETL package, lucky there was limited strain on our BI server and the data was integrated into the BI solution seamlessly. I think that part of the data migration went very smooth and according to plan.

The last questions asked to the participants in this category were linked to empirical object E04 (the overall stakeholder experience and impression surrounding the BI architecture implemented), as depicted in Table 5-11 below. From the comments made by the participants, the overall impression from stakeholders regarding the implemented BI architecture was found to be positive. The organisation's management team and users tended to interact more with the reporting solutions associated with the BI architecture as the information was received in a visual and user-friendly manner.

Table 5-11: Management – Stakeholder impression surrounding the business intelligence solution)

Interviewer:	What would you say was the overall impression and impact that the BI architecture had on the organisation?
Participant:	I think it was a positive impression, and I think that users and management like to have visibility of the data and they excepted and there, there very positive about it

Microsoft Power BI was used to visually display the results for the technical staff and data capturers (section 5.5.2 and section 5.5.3) as the larger number of participants deemed it possible and made it easier to communicate, whereas statements made by the management staff (section 5.5.1) could be reflected as there were only two management participants.

5.5.2 Technical staff participant results

This section is focused on the results received from technical staff participants regarding the old BI architecture, the new BI architecture implemented, and the data migration project conducted.

The first results from the participants addressed empirical objective E01, as displayed in Fig. 5-1. Five participants stated that the old BI reporting solution had different business rules across different departments in the organisation, making it difficult for the participants to keep track of all business rules associated with reporting development processes. Five participants commented that information presented to users generating reports from the old reporting solution would sometimes be duplicated. This was caused by the complexity of and outdated transactional structured query language (T-SQL) attached to the old reporting solution. A large amount of time was used by participants to correct each report as there were different business rules associated with different departments generating reports on the old reporting solution. The last comment made by three technical staff participants was that the organisation mostly made use of SSRS in the old BI architecture. This resulted in data being displayed in a non-graphical manner, which made it difficult for users to read and interact with the data.

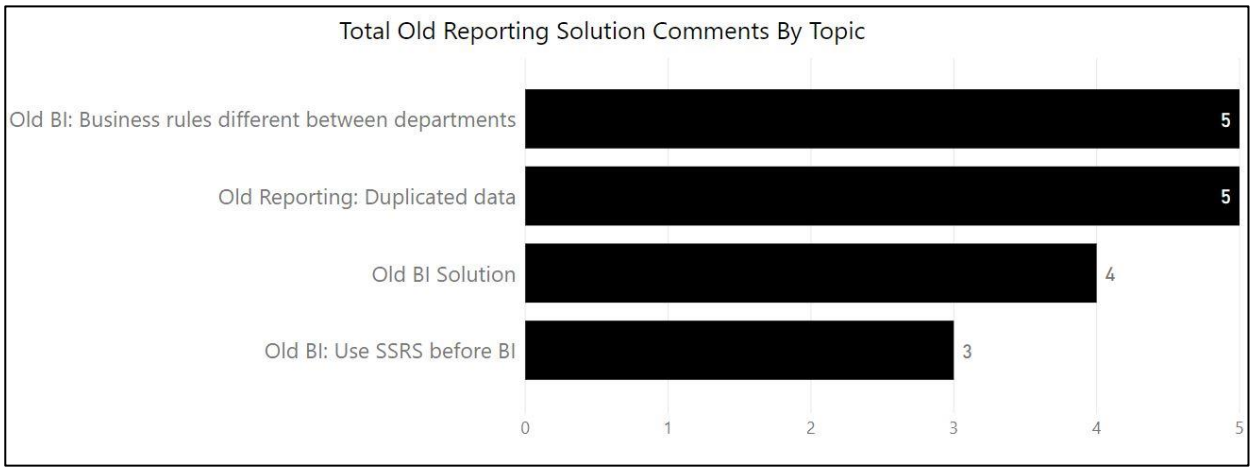


Figure 5-1: Technical staff total number of comments regarding the old business intelligence solution

The next results collected from participants related to empirical objective E02 – participants' experiences with using the newly implemented BI architecture. As presented in Fig. 5-2, the main response from participants was that reports associated with the new BI architecture were easier to design, develop and understand from a developer's point of view. The second most common comment received from participants was that only positive feedback was received from various business users throughout the organisation using the new reporting solution. Further feedback received by participants was a combination of codes, such as that the new BI reporting system was integrated seamlessly into the organisation, the business users were able to interact well with the new functionality of the BI reporting solution and that the data being retrieved from the organisations ERP system was stored in one central location, making development of new reports on different platforms much more efficient. Furthermore, as

presented in Fig. 5-2, the rest of the comments from participants included codes such as information being introduced to business users being consistent, business users feeling positive regarding information displayed in a graphical manner, reliable, accurate data being shown to business users generating reports, better data extraction from source systems, data being managed in one central location, and the new BI architecture simply being an improvement over the old BI architecture.

The results of data collected from technical staff participants, as depicted by Fig. 5-2, also portrayed that the new BI architecture implementation took five months to be developed. This result aligned with the new BI architecture development project development time forecasted and introduced by the project management staff. Participants also commented that data extraction using hourly incremental ETL processes kept information displayed in reports to business users up to date. Furthermore, any changes that might have occurred on the pre-existing information already integrated into the new BI architecture were amended by using a type two slowly changing dimension ETL approach. Lastly, participants confirmed that, with the table structure and logical flow of data, architecture new BI architecture was significantly easier to understand and develop.

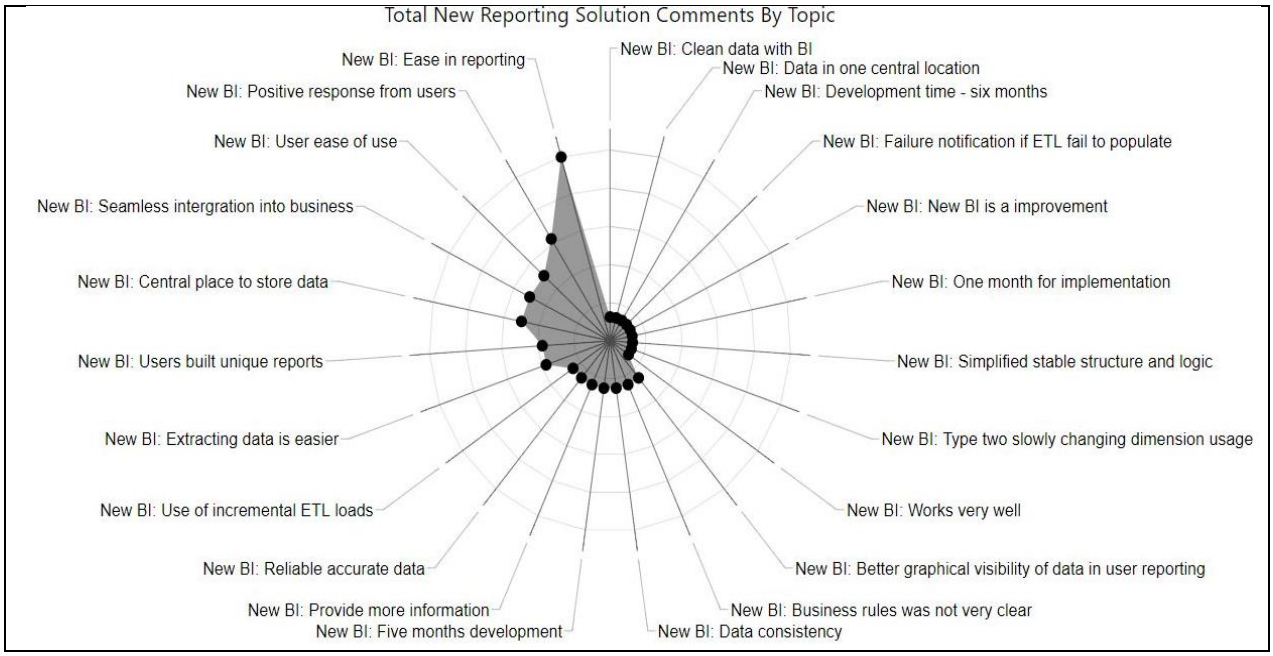


Figure 5-2: Technical staff total number of comments regarding the new business intelligence solution

The next set of comments received from participants concerned empirical objective E03, responding to the researcher's questions associated with the data migration from one DBMS to

another. In Fig 5-3, the results of the data collected from participants regarding the data migration from one organisation DBMS to another are presented. The comment most raised was that there was a big difference between the two organisation's data structures, which made the data mapping functionality extremely hard to comprehend and manage. The second most mentioned comment from participants was a lack of preparation before the data migration occurred, which impacted the time and effort that was needed to conduct the migration process. The next most common comments from participants were that the high level of cardinality of data imported between organisations made the data migration take significantly longer, and that it was difficult to map data between the organisations' DBMS structures.

From the results of data collected from technical staff participants, it was also shown that the actual completion time of the data migration project was more than a year. This was well beyond the six-month completion timeframe initially forecasted by the project management staff. The organisation made use of outside consultants to prepare and import data into the organisation's DBMS. This resulted in the organisation having to allocate more funds to the data migration project than originally planned. Participants also commented that, when the integrated data were ready to be introduced to the new BI architecture, a new data migration ETL package was created, the organisation's daily ETL package was halted and the newly-created data migration ETL package was executed to import the integrated data into the new BI architecture. Once the new data migration ETL package had been completed, the organisation's daily ETL package was restarted and continued as normal, with minimal impact on the organisation.

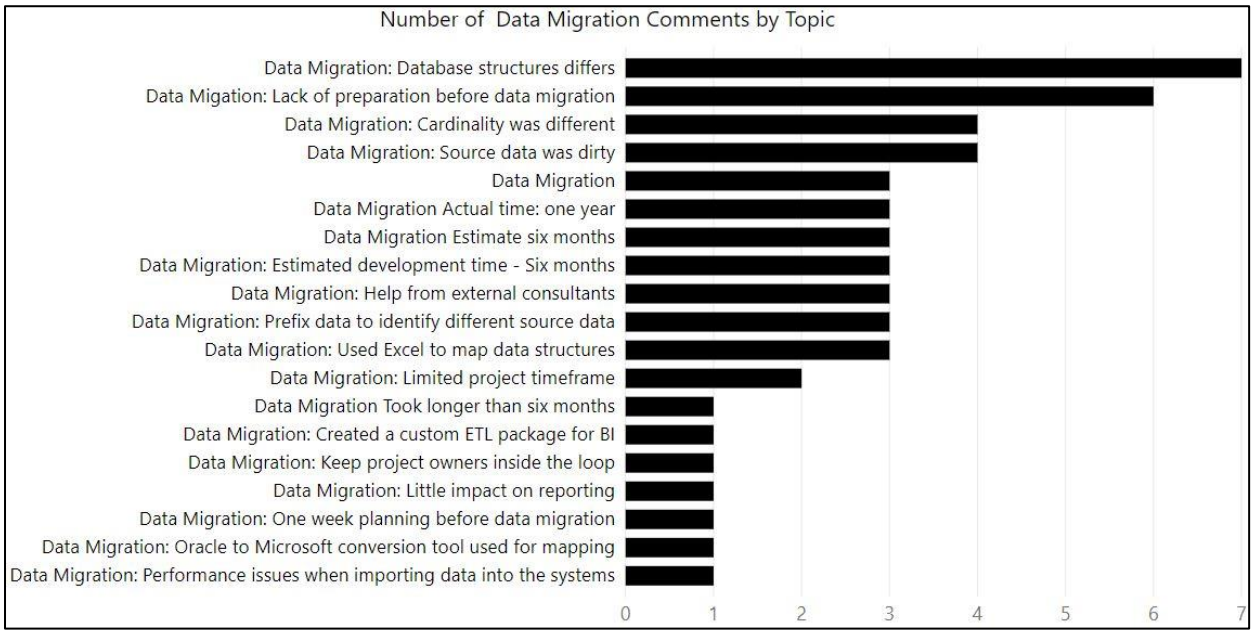


Figure 5-3: Technical staff total number of comments regarding the data migration project

In conclusion, the results of data collected from participants regarding the data migration process and concerning empirical objective E03 revealed that there were frustrations associated with the minimal time allocated to the planning stage of the data migration project. This left the technical staff no with choice but to prolong the actual data import phase of the data migration project, resulting in extra consulting resources and additional funding being needed. However, when the data were ready to be introduced into the new BI architecture environment, the process was seamless and had minimal impact on the organisation.

5.5.3 Data capture participant outcomes

In this section, the results from the data capturing category of participants associated with empirical objective E01 are discussed first. In Fig. 5-4, the feedback to the set of questions about the participants' experience in interacting with the old BI architecture in the organisation are provided. The most negative comments were made about the old BI reporting solution being too time-consuming when data capturers attempted to generate reports. The second most negative comment was that data capturers needed to manually re-generate the same report multiple times during the day to retrieve up to date information from the system.

The third most negative response was a combination of four comments, namely:

- the reporting solution was not user-friendly and was sometimes complex to work with;
- the static reporting solution would, at times, not include critical production data needed for business decisions;
- the old reporting solution would often malfunction after being manually generated; and
- due to the re-generation of reports, the old reporting solution would not motivate staff to work effectively on production tasks.

Lastly, a combination of five comments received from data capturing participants indicated that the old reporting solution was ineffective, not frequently used by data capturers and too tedious to use; furthermore, data capturing participants found it difficult to read complex report; and reports would sometimes not update when these reports were most needed.

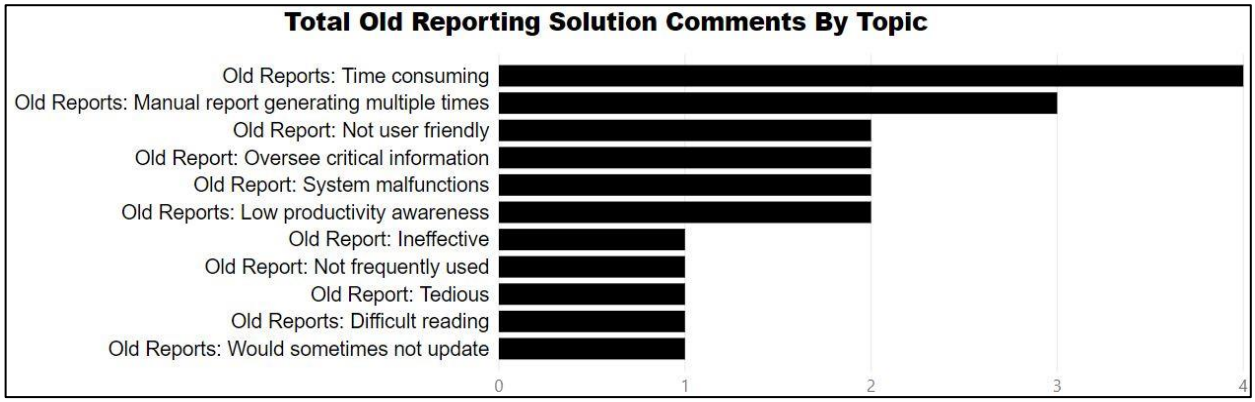


Figure 5-4: Data capturer total number of comments regarding the old business intelligence solution

The next set of questions addressed the participants' experience and feedback after using the new BI architecture (E02). Figure 5-5 presents the feedback from data capturing participants that interacted with the new implemented BI reporting solution. The comment made most frequently was that the information displayed back to the participants using the new reporting solution was constant, with information being distributed in the same manner across all reporting platforms in the organisation. The second most frequent combination of comments was that the new BI reporting solution was very user-friendly to interact with as the information was displayed visually to the data capturers. This led to higher user productivity as the data capturers' interest in interacting with the new BI reporting solution increased.

From the results, as depicted in Fig 5-5, participants also mentioned that information processed through the new BI reporting solution was updated on a more frequent basis than the old BI reporting solution, resulting in participants interacting more with the new BI reporting solution and therefore being more motivated and productive. Other comments received from data capturing participants included codes such as the new BI reporting solution being more convenient to use, effective, precise and displaying a wider variety of information to data capturers, and that the participants preferred interacting with the new BI reporting solution as opposed to the old BI reporting solution.

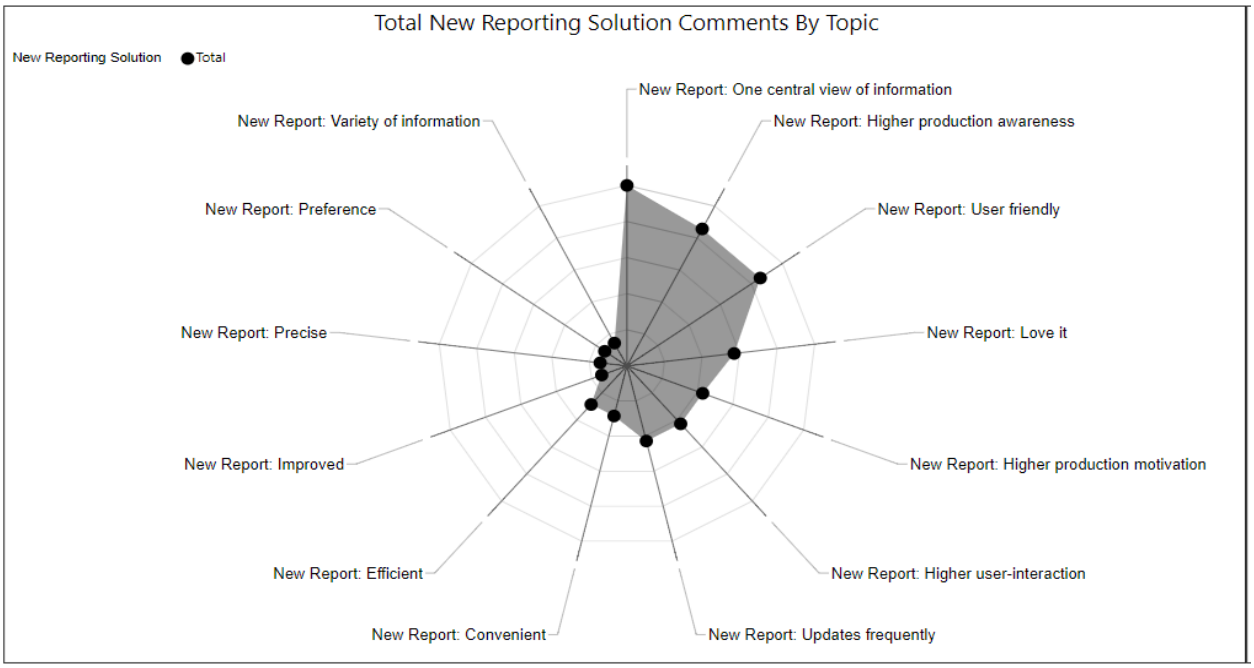


Figure 5-5: Data capturer total number of comments regarding the new business intelligence solution

The new BI architecture provided one central location of information across multiple reporting platforms, was more user-friendly, and encouraged data capturers to be more productive as the information was up to date and readily available.

5.5.4 Conclusion of the results

In this section, a summary of the results of the senior management staff, the technical staff and the data capturing staff is provided.

5.5.4.1 Senior management participants

Results of data collected from the senior management participants regarding the old BI architecture are as follows:

- The old BI architecture in the beginning of the organisation's existence used a Microsoft framework structure, with SSRS reporting solutions using T-SQL queries to retrieve the data straight out of the Microsoft DBMS structure. As the data growth inside the organisation increased, there was frustration with the quality and integrity of the data generated by die old reporting solution. The old reporting solution presented duplicate data and insufficient information back to the organisation for decision-making purposes (empirical objective E01 was partially achieved).

- The new BI architecture generated positive feedback with minimal concerns. There were some initial concerns with senior management and stakeholder investments to development the new BI architecture. When the new BI architecture was implemented, the organisation was quite pleased with the quality of information available to users, the large amount of information available and the fact that data from all ERP systems was being stored in one central location (empirical objective E02 was partially achieved).
- There were concerns about the minimal time allocated for the planning phase during the data migration project. The planning and completion of the data migration took six months longer than anticipated by stakeholders. In addition, external staff was acquired to help with the workload, map data structures and load the data into the organisation's DBMS, costing the organisation more than anticipated. Another concern was to make sure the data integrated into the organisation's DBMS balanced with the general ledger acquired from the other organisation. Once the data were ready to be imported into the organisation's BI architecture, there was very little processing strain on the organisation's production BI architecture services during the additional data import, and the process was completed seamlessly (empirical objective E03 was achieved).
- The senior management staff was impressed with the ability to visually present production data with the Power BI reporting solution. It was easy for business users to create custom reports (empirical objective E04 was achieved).

5.5.4.2 Technical staff participants

Results of data collected from the technical staff are as follows:

- The old BI architecture was prone to issues such as data being duplicated on various reports. Data in the old BI architecture was situated in various locations, which made retrieving data from these locations an impossible task. Furthermore, the T-SQL queries linked to the reports of the old BI architecture were outdated and very hard to read and amend (empirical objective E01 was partially achieved).
- The new BI architecture was easy to understand and to build on from a development point of view. The data integration from the organisation's ERP systems was much more efficient, with higher data quality and better integrity than the old BI architecture. The new ETL automation packages allowed information being processed by the organisation's BI architecture to always be up to date for reporting use (empirical objective E02 was partially achieved).
- The most critical complication of the data migration project was that there was not enough time allocated in the project to plan how the data would be migrated. The data structures of

the two organisations were very different and the data to be integrated was problematic and dirty. When the data were ready to be imported into the new BI architecture, the data migration ETL package ran seamlessly and without any problems (empirical objective E03 was achieved).

- The development of reports using the new BI architecture was much easier and more user-friendly, with positive feedback from business users (empirical objective E04 was achieved).

5.5.4.3 Data capture participants

Results of data collected from the data capturing participants:

- More benefits were realised from interacting with the new reporting solution as opposed to the old reporting solution (empirical objective E04 was achieved).
- There was more frequent interaction with the new BI reporting solution (empirical objective E04 was achieved).
- The data were up-to-date, allocated to one central location to be accessed from multiple reporting solutions and report development was user-friendly and displayed information in a graphical manner. Data capturing participants were more motivated to work and thus produced more daily (empirical objective E04 was achieved).

5.6 GUIDELINES FOR BUSINESS INTELLIGENCE MIGRATION IN A DEVELOPING ORGANISATION

In this section, the guidelines for BI migration in a developing organisation are discussed. The literature review conducted in this study (Chapter 3), design principles extracted from the case study's newly implemented BI architecture (descriptive case study part – Chapter 4), the results gathered from participants at various levels of the organisation through the empirical study and data interpretation (explanatory case study part – Chapter 5) were used to compile these guidelines. This section is aligned with the principle of abstraction and generalisation as stated by Klein and Myers (1999:72), where the researcher construct guidelines for the migration of BI in a developing organisation by data analysed from participants by abstracting meaning in the context of the study that can be applicable to other contexts.

In the subsequent sections, the researcher indicates how the literature reviewed, the case study and the results from the empirical research were applied to compile the guidelines for the BI migration in a developing organisation. In the subsequent sections, topics associated with the

guidelines for BI migration in a developing organisation are discussed, namely the BI architecture (section 5.6.1), data warehouse (DW) design (section 5.6.2), analysis design (section 5.6.3), reporting design and BI automation design (section 5.6.4).

5.6.1 Business intelligence architecture design

The BI architecture is the blueprint of the entire BI architecture and how all the components, such as the ETL, data warehouse, analysis solution and reporting solution, are associated with one another.

From the BI architecture design as described in Section (4.3.3.1). The BI guidelines proposed by this study, as depicted in Fig. 5-6, show how data are handled through different phases:

Phase 1 - The data capturers enter or update production data on several user input devices, for example an ERP system (section 3.2.1 and section 4.3.1.1).

Phase 2 - The data are then inserted or updated in the organisation's live DBMS structure.

Phase 3 - The data are transferred (in its original state) into a "landing environment" database inside the DW through an ETL process. The ETL process cleans and transforms the data, converts the data structure from its original data structure into a dimensional structure and stores the information in a "staging environment". The data from the "staging environment" is cleansed from anomalies such as nulls, and complex business rules are applied. The data are then transferred into a "live environment" inside the DW. Once this is done, the data (in a dimensional structure) is stored in one central location and can be retrieved for further processing (section 3.2.3 and section 4.3.3.1).

Phase 4 - The data are imported into an analysis service where the appropriate relationships between dimensions and facts are created, and additional business logic and key performance indicators (KPIs) are applied. The data column headers are also converted into a user-friendly naming convention to be used by the reporting platform later (section 3.3.4 and section 4.3.3.3).

Phase 5 - Once an appropriate analysis cube has been created, the data are imported and stored in an analysis database for further processing.

Phase 6 - The information is retrieved from the analysis database by the various reporting platforms and displayed back to the organisation for decision-making and other reporting needs (section 4.3.3.4).

Phase 7 - An automation process is then executed on a timely basis to re-run phase 3, phase 4 and phase 5 to make sure that the data being displayed by the reporting solution is kept up to date (section 3.2.6 and section 4.3.3.5).

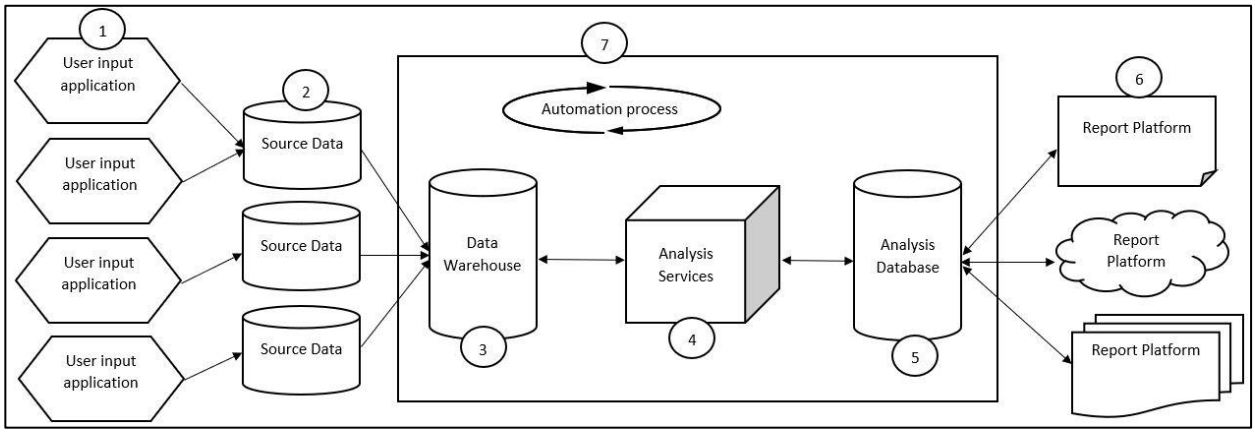


Figure 5-6: Business intelligence architecture guideline

If the process, as depicted in Fig. 5-6, has been successfully implemented, central, clean and structured information without anomalies or discrepancy can be expected by the organisation. The organisation may also rest assured that the information received is current.

5.6.2 Data warehouse design

The DW is a collection of data stored in one central location. The data goes through an ETL process to be cleaned from anomalies and discrepancies and is transformed into a dimensional structure for further use by an analysis solution or reporting solution.

From the DW design (section 3.2.1 and section 4.3.3.2) of the case study and in association with the Kimball BI lifecycle (section 3.2.5.1), Fig. 5-7 presents a guideline of how a thoroughly structured DW environment should be constructed and processed in three environment phases, namely:

Phase 1 - Landing environment. Data from various production sources, files and tables is digested into one central location. It is important that the data imported into the landing environment is implemented with the same data structure of the file or table from which the data are being imported. After the data have been transferred to the next environment, a full truncation of all tables in the landing environment is necessary to have a clean structure for the next data to be imported.

Phase 2 - Staging environment. The data are transferred from the landing environment into this environment. The data are then converted into a preliminary dimensional structure. While converting the data into this dimensional structure, the focus is on getting the data from its original data structure (as introduced in the staging environment) to be converted into a dimensional data structure with fact tables, incorporating numerical data only, and dimension tables, incorporating characteristic data. Any additional complexed measurements needed from the organisation are also constructed in this stage. It is important to keep in mind that, in this environment, data are not cleaned but the data structure is simply converted to a dimensional form. If applicable, a staging audit functionality can be added to the staging environment to cater for any error handling needed.

Phase 3 - Live environment. The data, in its newly structured dimensional form, is transferred into this live environment. The data are then cleaned of any anomalies, an unknown record is added in the dimensional tables to cater for any missing or corrupt values imported from the source tables, and the data are converted to the applicable data types needed. When the data are processed through this stage, all information should be cleaned of anomalies and should be ready to present back as meaningful information. A live auditing functionality is added to this environment to track and eliminate any issues that might arise during processing of data through this stage.

Once the data have been transferred through the live environment and transformed information is ready to be presented back to the organisation, the information is ready to be introduced into an analysis solution or reporting solution as needed.

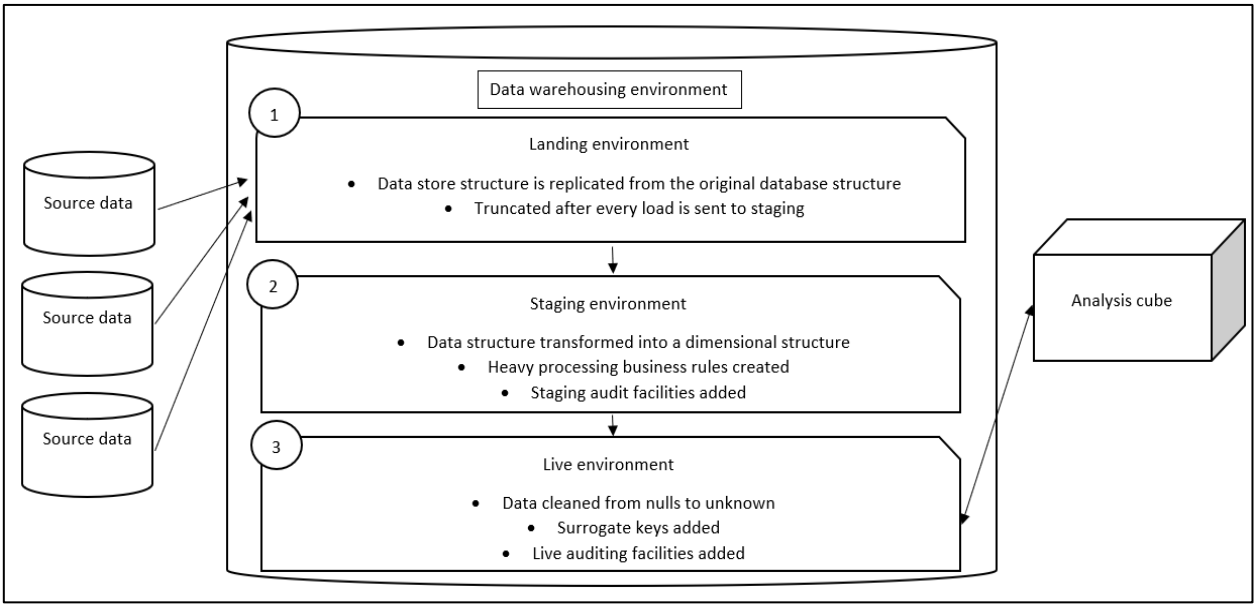


Figure 5-7: Data warehouse structure guideline

With the thoroughly designed DW structure as portrayed in Fig 5-7, the organisation can be confident that data gathered from various sources of business has been cleaned and stored in one central location to be processed through either an analysis solution or reporting solution.

It is important to be cognisant of the following considerations to address concerns that might arise when migrating large amounts of data from an external source, such as the data migration from the DBMS of one organisation to another:

- Allocate enough time for the planning and the scoring phase of the data migration project to identify all relevant sources and connections between data from various locations (Sections 5.5.4.1 and 5.5.4.2) and (The most critical complication of the data migration project was that there was not enough time allocated in the project to plan how the data would be migrated. The planning and completion of the data migration took six months longer than anticipated by stakeholders.)
- Make sure the data to be migrated are sampled correctly and are complete. (The comment most raised was that there was a big difference between the two organisation's data structures, which made the data mapping functionality extremely hard to comprehend and manage. The second most mentioned comment from participants was a lack of preparation before the data migration occurred, which impacted the time and effort that was needed to conduct the migration process. The next most common comments from participants were that the high level of cardinality of data imported between organisations made the data migration take significantly longer, and that it was difficult to map data between the organisations' DBMS structures.)
- Allocate enough dedicated resources, especially staff with the technical skills needed to perform tasks as required. This process of identifying the number of resources needed to complete the project should be identified by a DBA or project manager (External staff was acquired to help with the workload, map data structures and load the data into the organisation's DBMS, costing the organisation more than anticipated.)

With all the concerns identified and adhered to (refer section 5.5), an organisation in a similar position, may be confident when migrating large amounts of data from various sources, expecting limited complications.

5.6.3 Analysis design

After the data have been stored correctly into a DW environment, it is important to have a thoroughly designed analysis solution. The aim of the analysis environment is to have a

platform to build relationships between data, analyse the data, and add or remove business rules and logic, before sending the data through to the reporting environment.

From the analysis design (section 3.3.4 and section 4.3.3.3) of the case study and in association with tabular analysis modelling, by using data analysis expression (DAX) programming to be able to have a well-structured analysis structure, the data needs to go through several phases as presented in Fig. 5-8 below:

Phase 1 - Before information can be introduced into the analysis environment, relationships between the dimensional tables and fact tables need to be established. After the respective relationships between dimensional and factual tables have been established, data are imported into the analysis environment (section 4.3.3.3).

Phase 2 - A “date” dimension is developed and entails all information of the physical and financial date associated with the fact information, such as year, month, day and financial year.

Phase 3 - In this phase, the attributes that will be displayed to the users are formatted correctly and given the appropriate name structure. All key indication attributes should be hidden or excluded from user view in this phase (section 4.3.3.3).

Phase 4 - Appropriate hierarchies are designed in dimensional tables as prescribed by the organisation (section 4.3.3.3).

Phase 5 - Any additional measurements (as opposed to measurements already added to the DW environment) and KPIs should be created in this phase. Measures created in this phase include minor processing measures such as summing, minimum, maximum, average or conditional measurements (section 4.3.3.3).

Phase 6 - If applicable, any write-back functionality needed by the organisation to write information back to the analysis services from the reporting solution should be created in this phase. Most organisations write information back directly to the landing environment in the DW to be processed with the DW’s ETL process.

After the phases of data flow through an analysis environment have been completed, as depicted in Fig. 5-8, the analysis cube can be deployed to an analysis database. The reporting solution will then retrieve information form the analysis database when needed.

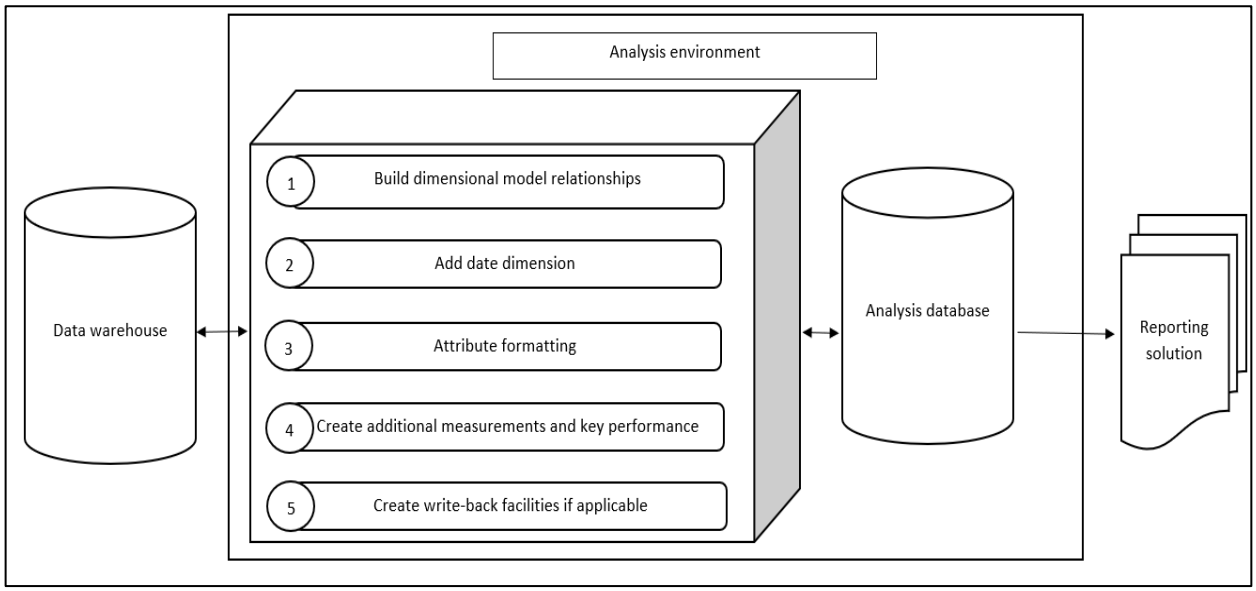


Figure 5-8: Analysis structure guideline

After a thoroughly developed analysis environment has been created, the organisation can be satisfied that the information retrieved from the analysis solution will abide by strict user-friendly naming conventions and additional business rules as needed.

5.6.4 Reporting design

The reporting environment plays a crucial role in any organisation. The reporting environment is a platform that displays the organisation's production information back to the organisation's users, and potentially to clients, to show operational, financial and business performance that is ultimately used to make critical business decisions.

After a thoroughly designed analysis environment has been structured and the data have been populated into an analysis database, the next logical step is to generate reports by querying directly from the analysis database, as opposed to from the DW database or, even worse, from the production database platform (section 4.3.3.4). The reason behind this is that the data positioned in the analysis database is already in a dimensional structure, relationships between dimensions and fact have been created, the data have been cleaned of any anomalies and formatted correctly, all business-oriented measurements and vital KPIs have already been incorporated, and business rules from most of the departments inside the organisation have already been developed. Furthermore, generating reports by using the analysis database removes the need for duplication of work as all needed work has already been done in the DW and analysis environments.

5.6.5 Business intelligence automation design

The BI automation process is important for any BI architecture to be able to store and keep data as current as possible with minimal human interaction needed.

From the BI automation design (section 3.2.6 and section 4.3.3.5) of the case study and, by association, concepts to consider when developing a BI automation solution as portrayed by Tomingas *et al.* (2015:8) (section 3.2.6), Fig. 5-9 presents steps to implement a thoroughly designed BI automation solution. These steps are:

1. Incremental automation process. The incremental automation process identifies data from the live operational databases that have been changed, deleted or inserted into the live operational database environment in the last few minutes. Only data that changes frequently is processed in this step to include changes made or new data added to the operational database. The incremental automation process then executes the BI architecture solution's ETL process according to the changed data only and inserts the data into an incremental section inside the live DW database structure. The incremental analysis environment then processes only data captured in this incremental process and stores the data in an incremental section of the analysis database structure. The incremental reporting solutions structure retrieves information from only the incremental analysis database structure and displays this information back to the business users. The information displayed back to the organisation in this reporting environment will be related to that day's operational procedures (section 4.3.3.5).
2. The historical automation process executes once daily and only after the close of the organisation's business hours when there is less performance strain on the operational server to which the BI architecture is deployed. The historical automation process runs the incremental ETL process as mentioned in step 1. First, all incremental information loaded throughout the day is retrieved. Thereafter, the historical automation process retrieves all other data that was not included with the incremental automation process throughout the day. The historical ETL process interjoins the data from the incremental ETL process and non-incremental ETL process into the historical DW section of the DW environment, while making sure that the data integrity and business rules are implemented as needed. Once the historical ETL process has been completed, the historical analytical structure re-processes the data to incorporate added or changed data into the historical section of the analytical database structure. Business users are then able to generate historical reporting solutions that include the latest data changes that occurred up until the previous working day (section 4.3.3.5).

By using the steps as provided in Fig. 5-8, the organisation can be sure that the BI architecture will automatically be updated with the latest production information from the ERP systems as soon as the business users have captured or changed data.

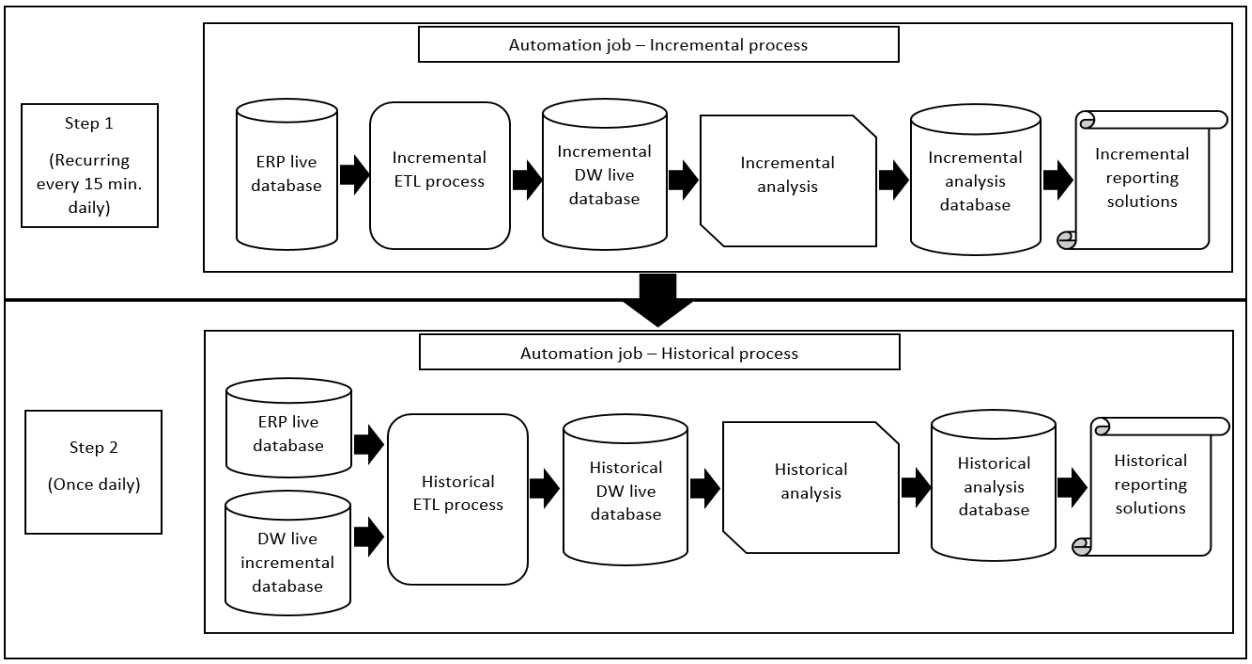


Figure 5-9: Business intelligence automation guideline

With the addition of the BI automation complete, the organisation has successfully implemented a durable BI architecture that will be able to cope with data growth throughout the organisation's expansion.

5.7 CONCLUSION

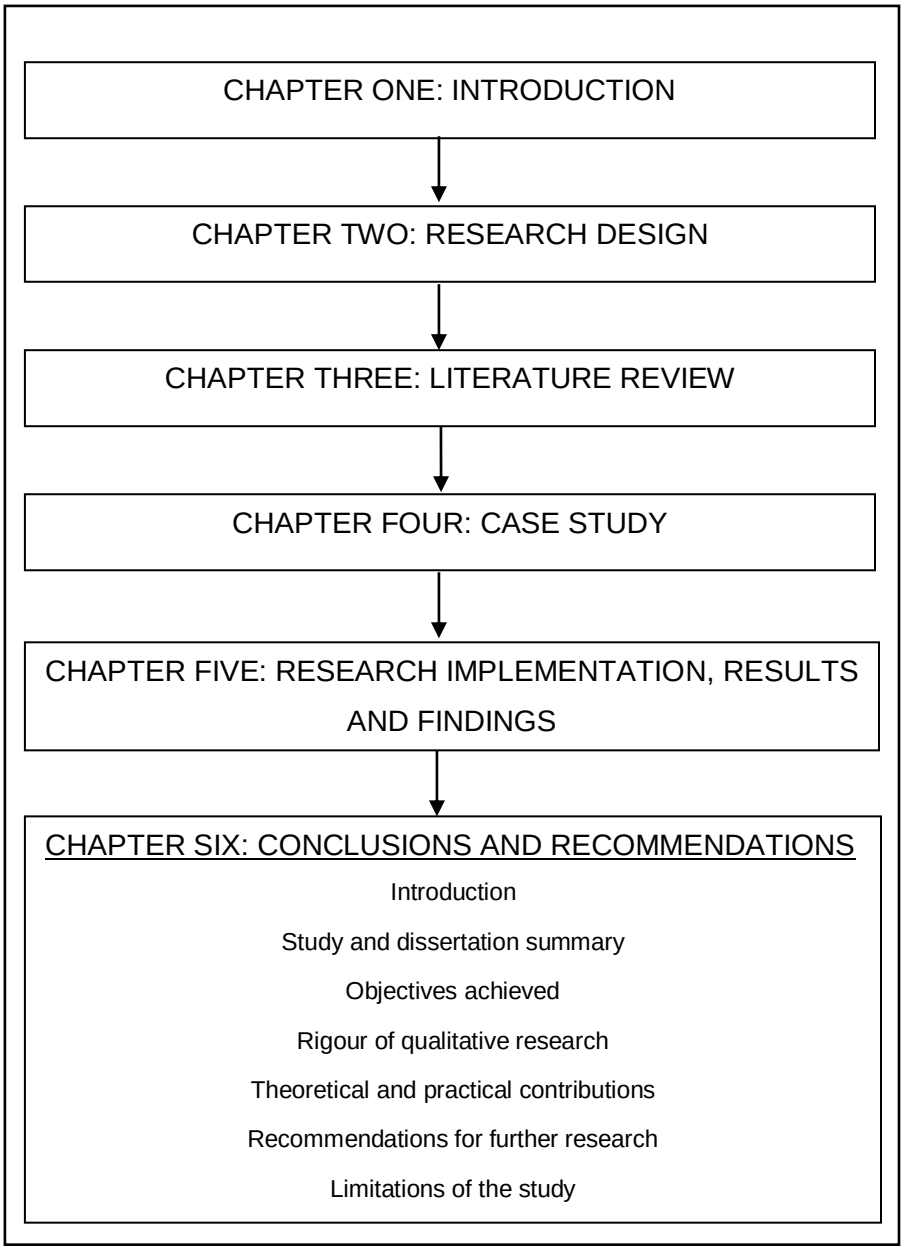
The chapter discussed the selection process for identifying and selecting participants, as well the interview questions and data collection process for three categories of participants. The process used to analyse data received from participants as well as the outcomes of the study in accordance to the empirical objectives (E01, E02, E03 and E04) were also discussed.

Guidelines for the migration of BI in a developing organisation were drafted according to the literature review, case study and outcomes to the study. Guidelines for the BI architecture, DW environment, analysis environment and BI automation environment were discussed.

Topics to be discussed in Chapter 6 include the conclusions of the research and recommendations for future studies. The chapter includes sections such as the study overview,

review of objectives to the study that have been met, a conclusion to the study, recommendations for future research, theoretical and practical contributions, and limitations of the study.

CHAPTER 6: CHAPTER FLOW



CHAPTER 6 CONCLUSION AND RECOMMENDATIONS

6.1 INTRODUCTION

The organisation used in this study grew from start-up to corporate size in just 10 years. With the growth came a large increase in production data captured by employees daily. The organisation quickly realised that the large amount of additional data resulted in the organisation's current business intelligence (BI) architecture not functioning as desired. The organisation decided to develop and implement a robust BI architecture, capable of processing the additional data and able to support any future data growth. After the upgraded BI architecture had been implemented, the organisation was tasked with a new challenge of integrating two database management systems (DBMS) into one and migrate the large amounts of data into their newly implemented BI architecture. The aim of this research was to study the upgrading of a BI architecture and data migration process in the organisation as a singular phenomenon and case, and to develop guidelines for migrating BI architectures in other developing organisations.

In the subsequent sections, several topics are addressed in conclusion to the study, namely the study and dissertation summary (section 6.2), objectives achieved (section 6.3), theoretical and practical contributions (section 6.4), recommendations for further research (section 6.5), limitations of the study (section 6.6) and finally a conclusion to the chapter (section 6.7).

6.2 STUDY AND DISSERTATION SUMMARY

The organisation used for the study was originally small but then experienced exponential data growth when it expanded. There was a need for the original BI architecture to be updated, and when the organisation bought out their most important business competitor, they not only had to deal with big data but also had to implement a data migration process from two different DBMSs into one.

The aim of this study was first to investigate, analyse and report on the way a BI architecture could be integrated into a growing organisation. The literature was used to define BI concepts used for the study. Organisation-specific BI-related and important concepts for the study were identified and discussed: the organisation's information and data flow process, initial and current BI architecture, project plan to develop a new BI architecture, the data warehouse (DW) implemented, and the data migration project plan. The second and main objective of the

research was to compile guidelines for possible use in other growing organisations experiencing exponential data growth or migration of BI.

The main objective of the research was to compile guidelines for migration of BI in developing (or growing) organisations. The literature was used to define BI concepts used for the study.

The research paradigm used for the study was interpretivism. A descripto-explanatory case study was used as the primary methodology. In the descriptive part, the old BI architecture, current BI architecture and data migration into the newly implemented BI architecture were discussed. Organisation-specific BI-related and important concepts for the study were identified and discussed using organisational-specific documentation and comments from employees: the organisation's information and data flow process, initial and current BI architecture, project plan to develop a new BI architecture, the data warehouse (DW) implemented, and the data migration project plan. The explanatory part of the case study was conducted through an empirical study. In the empirical study, semi-structured interviews were used as the method of data collection. Participants were identified and interviewed to capture three levels of knowledge and experience with regard to the BI development and data migration process of the organisation.

The questions asked to participants were formulated through the use of the literature on BI and data migration. Depending on the participants' levels of knowledge, all answers were mostly coherent and in agreement with the benefits of the newly developed BI architecture. Data collected from participants was then analysed using an appropriate qualitative analysis tool, ATLAS.ti where patterns and relationships from the qualitative data could be identified. An appropriate code book was drafted (section 5.4) and an associated code network was formulated for each category of participants (Appendix H, Appendix I and Appendix J). Seeing that there were only two managers interviewed, it was possible to show their direct statements. Microsoft Power BI was used to visually display the results for the technical staff and data capturers as the number of participants were more and the results could be made more visible and easier to read. The results and findings associated with the research objectives were presented, and the outcome of the research was discussed. Lastly, the guidelines for BI migration in a growing organisation were compiled and presented.

Chapter 1 provided an overview of the research. Chapter 2 discussed the research paradigms suitable for use in the information systems research environment, qualitative strategies of inquiry, research methods, data collection methods, data analysis and the study plan developed for this research. Chapter 3 presented the literature review on BI components such as the BI

lifecycle, evolution of BI and BI maturity, data migration phases, extract, transform and load (ETL) process, data scientists, and the difference between a NoSQL DBMS and a relational DBMS. Chapter 4 provided a discussion about the descriptive part of the case study such as the BI architecture implemented when the organisation was of start-up capacity, the BI development project plan, the current BI architecture and data migration from one organisation's DBMS to the current DBMS. Chapter 5 presented the explanatory part of the case study and the empirical research conducted, including the participant selection process, data collection techniques, the results of the data gathered and analysed from participants, findings and guidelines for BI migration in a developing organisation.

6.3 OBJECTIVES ACHIEVED

In this section, the objectives achieved by this study as set out in Chapter 1 are discussed in detail.

Fig. 6.1 presents the flow in which the events took place in the case study environment:

1. The BI architecture implemented in the organisation's first years comprised an initial server structure and an initial reporting solution.
2. A BI plan was developed to upgrade the organisation's BI architecture.
3. A BI architecture was implemented with key components such as the current BI server structure implemented, DW implemented, analysis services implemented, reporting solution implemented, and the automation process implemented.
4. The start-up and acquired organisations' data were migrated into the BI architecture of the start-up organisation, using a thorough data migration plan.

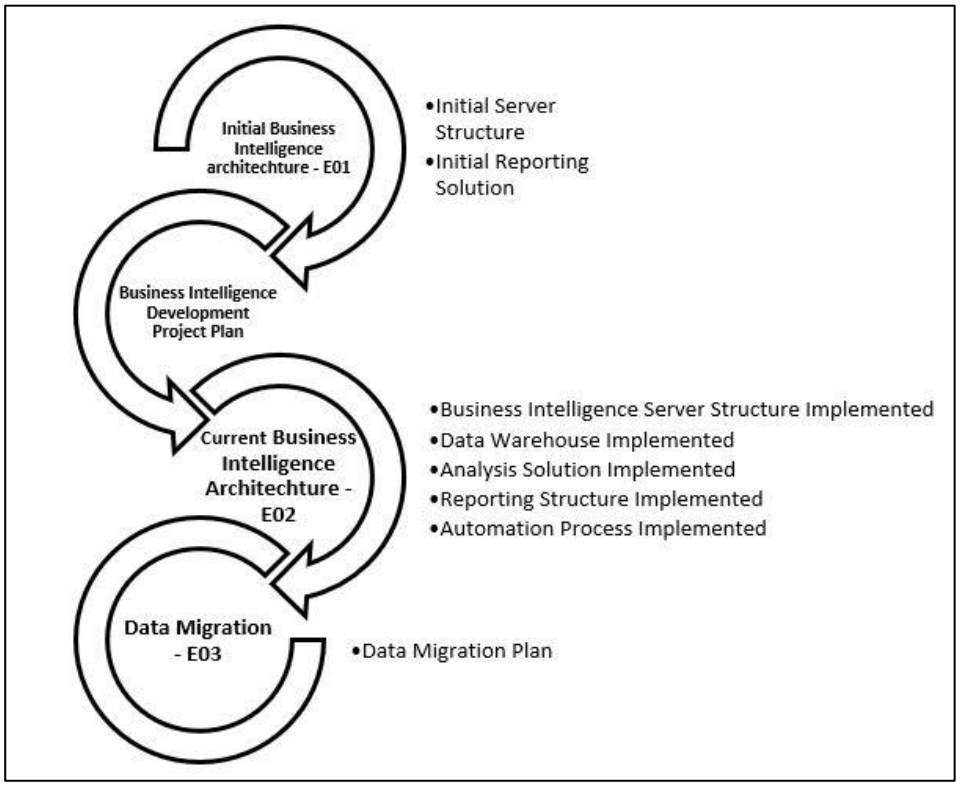


Figure 6-1: Case study event flow

In the following sub-sections, the objectives of the research and how they were accomplished are revisited: primary objectives (section 6.3.1), theoretical objectives (section 6.3.2) and empirical objectives (section 6.3.3).

6.3.1 Primary objectives

The primary objective of the study was:

> *To compile guidelines to assist developing organisations on how to overcome challenges of data growth while successfully implementing a BI architecture*.

Based on the literature review, descripto-explanatory case study and findings from the empirical study, guidelines associated with BI architecture implementation were compiled to assist developing organisations with data handling. The guidelines included BI architecture design guidelines as seen in Fig. 6-2 in the following phases:

1. Source data. The data in placed into the production server from various data sources.
2. Data warehouse. The data from the production server are placed through an ETL process where the data are transferred into the landing environment of the data warehouse. The data are then transformed into a multidimensional structure and placed into a staging

environment. Finally, the data are cleaned and performance-heavy additional measurements are added. The data are placed in a live environment after the ETL process is completed.

3. Analysis services. The data are transferred into an analysis services environment where links between the dimensional fact and dimension tables are made, Key performance indicators (KPI's) are created and any divisional light-process measurements are created. The analysed data are then placed into an analysis database.
4. Reporting services. The reporting services receive data from the analysis services and display the data to business users in multiple reporting solutions.

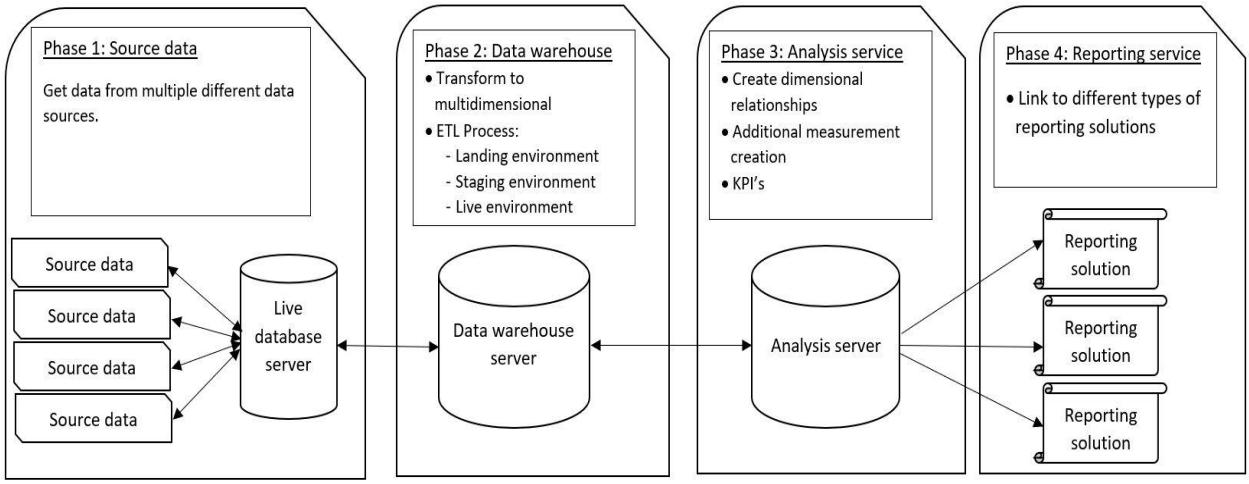


Figure 6-2: Business intelligence architecture design guidelines

Constraints to consider when integrating data from different DBMSs and migrating BI in developing or growing organisations were identified as:

- Allocate enough time for the planning and the scoring phase of the data migration project to avoid cost implications.
- Make sure the data to be migrated are sampled correctly and are complete.
- Allocate enough dedicated resources, especially staff with the technical skills needed to perform tasks as required. This process of identifying the number of resources needed to complete the project should be identified by a DBA or project manager.

6.3.2 Theoretical objectives

The theoretical objectives for the study were formulated and achieved as follows:

T01: To perform a literature review in support of all concepts key to the research, namely the evolution of a BI architecture, defining a DW and research conducted with regards to BI in growing organisations.

Chapter 3 provided the literature reviewed in support of BI architecture concepts such as the ETL process, DW design, analysis solution, reporting solution and the automation solution. Concepts related to data migration were also discussed in Chapter 3.

T02: To use the literature to describe the research paradigm, design and methods used for the research.

Chapter 2 provided an overview of the literature reviewed to describe the available and mostly used research paradigms available in the information systems (IS) field. Interpretivism was selected as the research paradigm for this study, with a descripto-explanatory case study as the qualitative strategy of inquiry. Interviews were used to collect qualitative data from participants. The study followed the principles as stated by Klein and Myers (section 1.6.4).

6.3.3 Empirical objectives

In accordance with the primary objective of the study, the following empirical objectives were formulated:

E01: To describe the BI architecture of the small organisation in its first year of existence.
E02: To describe the newly implemented BI architecture.
E03: To determine and report how data from two different database platforms (Microsoft and Oracle) could be integrated into one BI architecture.

Empirical objectives E01, E02 and E03 were addressed in Chapter 4 in this study by providing information regarding the organisation's BI architecture implemented in its first year of existence, the project plan of how the newly implemented BI architecture was built and the current implemented BI architecture. Chapter 4 also discussed the migration process of how data were migrated from one organisation's DBMS to another. Chapter 4 was the descriptive part of the case study.

E04: To describe the personal experiences and involvement of stakeholders over the course of developing a new BI architecture in the organisation.

For the explanatory part of the case study, an empirical study was performed, qualitative data were collected through semi-structured interviews, analysed and the results were presented. Empirical objective E04 was addressed in Chapter 5 by presenting the findings from data collected by participants in relation to their personal experiences and involvement as stakeholders over the course of migrating a new BI architecture in the organisation. From the data collected, four distinct themes could be identified: the organisation's old BI architecture, the old reporting solution, the organisation's new BI architecture and new reporting solution, and the data migration between two organisations' DBMSs.

In summary, this is how the empirical objectives of the research were achieved:

E01: The BI architecture of the small organisation in its first year of existence was discussed (Chapter 4).

E02: The researcher was able to portray how data from two different database platforms (Microsoft and Oracle) was integrated (Chapter 4).

E03: The impact of integrating data from two organisations on the start-up organisation's BI architecture and the way that the BI architecture had to be adjusted to be able to incorporate the additional size data were described (Chapter 4).

E04: Personal experiences and involvement of stakeholders in the course of developing a new BI architecture in the organisation were described (Chapter 5).

6.4 THEORETICAL AND PRACTICAL CONTRIBUTIONS

The theoretical contributions of the research to the field of study of BI are:

- The study provided guidelines to assist developing organisations on how to overcome challenges of data growth while successfully implementing a BI architecture (section 5.7).
- The study made use of a descripto-explanatory case study (section 2.3.4). The case was *described* in Chapter 4 and in addition, semi-structured interviews were utilised to collect data from participants to *explain* the case.

The practical contributions of the study include:

- Guidelines for organisations considering implementing a new or developing an existing BI architecture.

- Meaningful information for organisations planning to integrate large amounts of data from one DBMS structure to another, as well as advice on how to address complications that might arise.

Additionally, from the empirical study, the following practical advantages useful for other growing organisations were identified:

- An appropriate BI architecture implemented in an organisation can result in higher user productivity (section 5.5.3).
- An appropriate BI architecture implemented in an organisation can result in more user interaction with reporting solutions (section 5.5.2), data equality and consistency (section 5.5.2), and one source of true information incorporated from multiple user input structures across the organisation into one DBMS (section 5.5.1).
- It is possible to incorporate large amounts of data from one DBMS platform to another with minimal impact on an organisation's implemented BI architecture (section 5.5.1).

6.5 RIGOUR OF QUALITATIVE RESEARCH

Rigour of the study was assured as described in Section 1.6.4 and Table 6-1:

1. Interviews conducted with participants were recorded and then transcribed to ensure the authenticity and credibility of data collected from each participant.
2. Transcripts of participants from which data were gathered and used for the study, were checked using peer coding (as seen in Appendix K)
    a. Themes and codes were extracted by another person.
    b. Themes and codes were compared with the researcher's themes and codes.
    c. Differences were discussed and corrections were made.
    d. The most relevant codes and themes were used in the analysis of the results.
3. Integrity of data collected was therefore assured.

Table 6-1: Guidelines and principles of rigour adhered to by this study

Guideline or principle	Description	Reference in study
Authenticity guideline	The participants must be able to speak freely, their voices must be heard when spoken, and must be presented in an accurate way.	Section 5.3: An appointment was made with each participant and a consent form explaining the research was send to the participant prior to the interview. Interviews were recorded with consent from

Guideline or principle	Description	Reference in study
		participants. After each interview, the researcher compiled a summary of the interview, with added notes. See Appendix C, D, E, F, and G for interview questions. Section 5.5: The results of the interviews are presented.
Credibility guideline	A true researcher's perspective must be captured and portrayed.	Interviews were recorded and transcribed. The results were presented in Section 5.5.
Integrity guideline	Bias must be reflected, informants must be validated, and peer review of data collected and analysed must be ensured.	The description of the BI architecture development and upgrading process of the organisation was guided by an organisational compiled technical document (Anon, 2016). Participants from three different work-levels were interviewed (sections 5.5.1, 5.5.2 and 5.5.3). Peer coding was done to validate the codes extracted from the data.
The hermeneutic circle principle	The hermeneutic circle principle is primarily a representation of the human understanding of independent sections or parts of meaning from participants in the study. Human understanding is fundamental to all the other principles listed below. The researcher needs to consider and portray the participants' emotional state with regard to the study environment.	Section 5.4: The codes and themes as analysed from the transcripts are portrayed in Tables 5-1 to 5-5. ATLAS.ti was used to code the data.
The principle of contextualisation	The principle of contextualisation requires the researcher to have a critical reflection of the historical and social background of the study environment in order to be able to understand how the current situation emerged. This principle requires that the researcher understands the BI architecture migration in its context, from its origin to its current state.	Chapter 4: The contextualisation principle of Klein and Myers (1999:72) is adhered to in this chapter by describing the case study environment, the BI architecture that was used when the organisation was of start-up capacity, and how the BI architecture implementation currently functions.
The principle of interaction between the researcher and the subjects	The principle of interaction between the researcher and the participants is reflected in the structuring of the data received from the participants. The researcher needs to be able to relate to the participant while conducting the research to truly understand what is being described by the participant. In this study, the researcher was exposed to the study environment during the BI architecture evolution and data migration from one DBMS platform to another.	Section 5.3: The data collection approach is aligned with the principle of interaction between researcher and participant as stated by Klein and (Myers. 1999:72). The researcher was exposed to the study environment and could relate to the participants experiences and ask appropriate questions during the interview process.
The principle of abstraction and generalisation	The principle of abstraction and generalisation entails understanding the idiographic ideals relating to the	Sections 5.5 and 5.6: This section is aligned with the principle of abstraction and generalisation as stated by Klein and Myers

Guideline or principle	Description	Reference in study
	social and cultural actions of the data collected from the participants. This principle needs to be correlated with the first two principles. The researcher needs to analyse data collected from participants by abstracting meaning in the context of the study, leading to the extraction of guidelines applicable in other contexts.	(1999:72), where the researcher construct guidelines for the migration of BI in a developing organisation. These guidelines were abstracted from the descriptive part (Chapter 4) of the case study and through data analysed from participants by abstracting meaning in the context of the study (Chapter 5). The guidelines can be applied in other contexts.
The principle of dialogical reasoning	The principle of dialogical reasoning revises the actual findings and research design, guided by sensitively identifying differences between preconceptions ("stories told through the data"). The researcher should be aware that the conclusion of the research design and research data, analysed from participants, might be different. In the case of this study, implementations regarding BI architecture and data migration might differ from the opinions of the participants directly impacted.	Sections 4.3, 4.4 and 5.5.1 to 5.5.3: The actual findings of participants from three different work-levels were listed to identify differences between preconceptions about the processes of data migration and BI implementation (Chapter 4) and the feedback from participants (Chapter 5).
The principle of multiple interpretations	The principle of multiple interpretations involves different perspectives received from different participants in the same study environment or sequences of events expressed in multiple stories or narratives. Participants will describe the study environment as they experience it. The researcher should be aware that the opinions of participants impacted by the same study environment might differ.	Section 5.5: The researcher portrays different perspectives of participants in the same study in different forms such as senior management participants, technical staff participants and data capture participants.
The principle of suspicion	The principle of suspicion requires the researcher to be sensitive to possible distortions or biases in data collected from the participants. The researcher should not only rely on the words of the participants during the interview but should also look at behaviour such as facial expressions and body language when questions are answered.	Section 5.5: The researcher made notes during the interviews, but the study was of such a nature (description of a two-fold process and participants' perceptions of the process) that the notes were not used. However, the researcher reported on the data collected and feedback of participants and also used that to compile guidelines for the migration of BI in developing organisations. .

6.6 RECOMMENDATIONS FOR FURTHER RESEARCH

If future studies were to be conducted, a possible comparison study can be done between specific organisations using integrated BI architecture and organisations without BI architecture to report on the possible advantages and challenges. Further research can also be conducted

on whether an organisation should incorporate a BI architecture into an on-site or cloud-based environment, or a possible hybrid solution of the two environments. Lastly, further studies can be done on best practices and techniques related to the design, development and implementation of an appropriate BI architecture into an organisational environment.

6.7 LIMITATIONS OF THE STUDY

A first limitation to the study would be that other organisations' data policies, server structures and documented processes might differ from the study environment. A second limitation of the study concerns the fact that key people involved in the transformation process had left or been replaced by the organisation since its inception. A third limitation of the study would be that other organisations' reasons for implementing a BI architecture might differ from the study environment. Lastly, the case study was performed in one organisation; therefore, the guidelines developed are associated with a single organisation's experience of implementing a BI architecture and conducting data migration from one DBMS platform to another.

6.8 CONCLUSION

The study and dissertation summary were discussed in this chapter. Thereafter, the primary, theoretical and empirical objectives achieved through the research were discussed. Theoretical and practical contributions that the study made to the BI research community and the study environment were also discussed in this chapter. Furthermore, recommendations for future research were reviewed. Lastly, the limitations of the study were considered.

Through the perspective of the participants, the difference between the old reports and new reports measured personal capability growth, development in understanding and overcoming challenges, and growth within the BI architecture as well as within the organisation. The most important feedback was the improvement from the old solution to the new solution and how it increased user awareness of what was happening in the underlying data.

The proposed guidelines for the migration of business intelligence in a developing organisation may be useful for other start-up organisations in different context and environments.

BIBLIOGRAPHY

Agarwal, D., Aggarwal, C. & Elangovan, K. 2015. Microsoft Dynamics AX 2012 R3 Reporting Cookbook. Brimingham: Packt.

Agasisti, T. & Bowers, A.J. 2017. Data analytics and decision making in education: towards the educational data Scientist as a key actor in schools and higher education institutions. (*In* Johnes, G., Johnes, J., Agasisti, T. & López-Torres, L., *eds*. Handbook of contemporary education economics. Cheltenham: Elgar. p. 184-210).

Albano, A., Colazzo, D., Ghelli, G. & Orsini, R. 2015. Relational DBMS Internals. http://pages.di.unipi.it/ghelli/bd2/DBMS-Internals.pdf Date of access: 2019-04-12.

Alnoukari, M. & Hanano, A. 2017. Integration of business intelligence with corporate strategic management. *Journal of Intelligence Studies in Business*, 7(2).

Andersen, O., Thomsen, C. & Torp, K. 2018. Simple ETL: ETL Processing by Simple Specifications. (*In* Workshop proceedings of the EDBT/ICDT 2018 Joint Conference, Vienna, 26 March. S.l.: s.n.). http://ceur-ws.org/Vol-2062/paper10.pdf Date of access: 2019-04-12.

Anon. 2016. Technical Specification: Business Intelligence Upgrade

Anon. 2018a. Enterprise BI in Azure with SQL Data Warehouse. https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/data/enterprise-bi-sqldw Date of access: 2019-08-12.

Anon. 2018b. Extract, transform, and load (ETL). https://docs.microsoft.com/en-us/azure/architecture/data-guide/relational-data/etl#extract-load-and-transform-elt Date of access: 2019-08-11.

Aveyard, H. 2014. Doing a literature review in health and social care: A practical guide. London: McGraw-Hill.

Azam, M. & Tanweer, M.S. 2017. Impact of Business Intelligence and Predictive Analysis in Big Data. *International Journal of Advanced Research in Computer Science*, 8(5).

Azeroual, O. & Theel, H. 2019. The effects of using business intelligence systems on an excellence management and decision-making process by start-up companies: A case study. *arXiv preprint arXiv:1901.10555*.

Baskerville, R.L. & Wood-Harper, A.T. 1996. A critical perspective on action research as a method for information systems research. *Journal of information Technology*, 11(3):235-246.

Bernard, H.R., Wutich, A. & Ryan, G.W. 2016. Analyzing qualitative data: Systematic approaches. Los Angeles, CA: Sage.

Birks, M. & Mills, J. 2015. Grounded theory: a practical guide. Los Angeles, CA: Sage.

Bodislav, D.A. 2015. Transferring business intelligence and big data analysis from corporations to governments as a hybrid leading indicator. *Theoretical and Applied Economics*, 22(1):257-264.

Bowen, G.A. 2009. Document analysis as a qualitative research method. *Qualitative research journal*, 9(2):27-40.

Box, R.C. 2015. Critical social theory in public administration. London: Routledge.

Brannen, J. 2017. Mixing methods: Qualitative and quantitative research. London: Routledge.

Bryman, A., Bell, E. & Harley, B. 2018. Business research methods. Oxford: Oxford University Press.

Caudle, S.L. 2004. Qualitative data analysis. (*In* Wholey, J.S., Hatry, H.P. & Newcomer, K.E., 2015. Handbook of practical program evaluation. San Francisco, CA: Jossey-Bass. p. 417-438).

Coghlan, D. & Brannick, T. 2019. Doing action research in your own organization. Los Angeles, CA: Sage.

Collis, J. & Hussey, R. 2013. Business research: A practical guide for undergraduate and postgraduate students. London: Macmillan.

Côrte-Real, N., Ruivo, P. & Oliveira, T. 2014. The diffusion stages of business intelligence & analytics (BI&A): A systematic mapping study. *Procedia Technology*, 16:172-179.

Cuevas, A., Garcia, L. & Torres-Sánchez, M. 2017. Interactive BI Platform for Business Performance Analysis. Paper presented at the Data Management Conference 2017, Calgary, April. https://www.researchgate.net/publication/316609351_Interactive_BI_Platform_for_Business_Performance_Analysis Date of access: 2019-04-12.

Day, P. 2017. BI Solutions Using SSAS Tabular Model: For Starters.

Denscombe, M. 2014. The good research guide: for small-scale social research projects. London: McGraw-Hill.

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K. & Kyngäs, H. 2014. Qualitative content analysis: A focus on trustworthiness. https://journals.sagepub.com/doi/pdf/10.1177/2158244014522633 Date of access: 2019-04-12.

Erskine, S., Gasser, D., Van Hoecke, K. & Ismail, N. 2015. Microsoft System Center Reporting Cookbook. Birmingham: Packt.

Ferrari, A. & Russo, M. 2015. The Definitive Guide to DAX: Business Intelligence with Microsoft Excel, SQL Server Analysis Services, and Power BI. Redmond, WA: Microsoft Press.

Fields, E. & Sheppard, B. 2012. A New Approach to Business Intelligence: Rapid-fire BI. https://www.kennisportal.com/KP/TableauSoftware/Rapid-Fire_Business_Intelligence.pdf Date of access: 2019-04-12.

Fredstam, M. & Johansson, G. 2019. Comparing database management systems with SQLAlchemy: A quantitative study on database management systems. http://www.diva-portal.org/smash/get/diva2:1298578/FULLTEXT01.pdf Date of access: 2019-04-12.

Friese, S. 2019. Qualitative data analysis with ATLAS.ti. Los Angeles, CA: Sage.

Gadde, A. 2016. Optimization of Asset Allocation. https://repository.stcloudstate.edu/mme_etds/32/ Date of access: 2019-04-12.

Gallinucci, E., Golfarelli, M. & Rizzi, S. 2018. Variety-Aware OLAP of Document-Oriented Databases. (In Workshop proceedings of the EDBT/ICDT 2018 Joint Conference, Vienna, 26 March. S.l.: s.n.). http://ceur-ws.org/Vol-2062/paper02.pdf Date of access: 2019-04-12.

Gibbons, M.T. 1987. Interpreting politics. New York, New York: New York University Press.

Giménez-Figueroa, R., Martín-Rojas, R. & García-Morales, V.J. 2018. Business intelligence: an innovative technological way to influence corporate entrepreneurship. (*In* Mura, L., *ed*. Entrepreneurship: development tendencies and empirical approach. S.l.: InTech). InTech. https://www.intechopen.com/books/entrepreneurship-development-tendencies-and-empirical-approach/business-intelligence-an-innovative-technological-way-to-influence-corporate-entrepreneurship Date of access: 2019-04-12.

Glaser, B.G. & Strauss, A.L. 2017. Discovery of grounded theory: Strategies for qualitative research. London: Routledge.

Gosain, A. & Singh, J. 2015. Conceptual multidimensional modeling for data warehouses: a survey. (*In* Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014. Cham: Springer. p. 305-316).

Granville, V. 2014. Developing analytic talent: Becoming a data scientist. Hoboken, NJ: Wiley.

Gregor, S. & Hevner, A. 2013. Positioning and presenting design science research for maximum impact. *MIS quarterly*, 37(2):337-355.

Gupta, B., Goul, M. & Dinter, B. 2015. Business Intelligence and Big Data in Higher Education: Status of a Multi-Year Model Curriculum Development Effort for Business School Undergraduates, MS Graduates, and MBAs. *Communications of the Association for Information Systems*, 36:449-476.

Hammersley, M. 2016. Reading ethnographic research. London: Routledge.

Hammersley, M. & Atkinson, P. 2007. Ethnography: Principles in practice. London: Routledge.

Hancock, D.R. & Algozzine, B. 2016. Doing case study research: A practical guide for beginning researchers. New York, NY: Teachers College Press.

Hart, C. 2018. Doing a literature review: Releasing the research imagination. Los Angeles, CA: Sage.

Held, D. 1980. Introduction to critical theory: Horkheimer to Habermas. Vol. 261: Univ of California Press.

Herr, K. & Anderson, G.L. 2014. The action research dissertation: A guide for students and faculty. Los Angeles, CA: Sage.

Hevner, A. & Chatterjee, S. 2010. Design Research in Information Systems: theory and practice. New York: Springer.

Hofstee, E. 2006. Constructing a good dissertation. *Johannesburg: EPE*.

Huang, S.-C., McIntosh, S., Sobolevsky, S. & Hung, P.C. 2017. Big data analytics and business intelligence in industry. *Information Systems Frontiers*, 19(6):1229-1232.

Inmon, W.H. & Linstedt, D. 2014. Data Architecture: A Primer for the Data Scientist: Big Data, Data Warehouse and Data Vault. Amsterdam: Morgan Kaufmann.

Iyer, S.S. & Lakhtaria, K.I. 2017. Data Analytics Using Satellite Remote Sensing in Healthcare Applications.

Jabbar, A.M. 2015. Design and Develop an Information system for Court Data in the Republic of Iraq by using SSRS Reports with SSAS Cubes. *Iraqi Journal for Electrical & Electronic Engineering*, 11(1):105-109.

Jamshed, S. 2014. Qualitative research method: interviewing and observation. *Journal of basic and clinical pharmacy*, 5(4):87-88.

Kimball, R. & Ross, M. 2015. The kimball group reader: Relentlessly practical tools for data warehousing and business intelligence remastered collection. Hoboken, NJ: Wiley.

Kitchener, R.F. 2004. Logical positivism, naturalistic epistemology, and the foundations of psychology. *Behavior and philosophy*:37-54.

Klein, H.K. & Myers, M.D. 1999. A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, 23(1):67-93.

Kothari, C.R. 2004. Research methodology: Methods and techniques. New Delhi: New Age International.

Krueger, R.A. & Casey, M.A. 2014. Focus groups: A practical guide for applied research. Los Angeles, CA: Sage.

Kuderu, N. & Kumari, V. 2016. Relational Database to NoSQL Conversion by Schema Migration and Mapping. *International Journal of Computer Engineering In Research Trends*, 3(9):506-513.

Kumar, D., Dhyani, P. & Sharma, A.K. 2015. Cross Breed Data Migration using ACO Algorithm. *International Journal of Computer Applications*, 116(1):1-6.

Kumar, R. 2019. Research methodology: A step-by-step guide for beginners. London, UK: Sage.

Larson, D. & Chang, V. 2016. A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, 36(5):700-710.

Laurier, E. 2016. Participant and Non-participant Observation. (*In* Clifford, N., Cope, M., Gillespie, T. & French, S., *eds. Key Methods in Geography*. London: Sage.

Łęgowik-Świącik, S., Łęgowik-Małolepsza, M., Stępień, M. & Turek, T. 2016. Information Management in Dispersed Enterprises with the Use of Business Intelligence. Paper presented at MultiScience 30: MicroCAD International Multidisciplinary Scientific Conference, University of Miskolc, Hungary, 21-22 April. http://www.uni-miskolc.hu/~microcad/publikaciok/2016/C2_7_Legowik_Swiacik_Sylwia.pdf Date of access: 2019-04-12.

Linstedt, D. & Olschimke, M. 2015. Building a scalable data warehouse with data vault 2.0. Amsterdam: Morgan Kaufmann.

LoBiondo-Wood, G. & Haber, J. 2017. Nursing Research: Methods and Critical Appraisal for Evidence-Based Practice. 9th ed. St Louis, MO: Elsevier.

Lu, J. & Holubová, I. 2017. Multi-model Data Management: what's New and what's Next? (*In* 20th International Conference on Extending Database Technology (EDBT), Venice, 21-24 March. S.n.: s.l. p. 602-605). https://www.cs.helsinki.fi/u/jilu/paper/multi-model-data_EDBT.pdf Date of access: 2019-04-12.

Lunt, T. 2014. (*In* Page, S. & Connell, J., 2014. The Routledge handbook of events. p. 47-56).

Machi, L.A. & McEvoy, B.T. 2016. The literature review: Six steps to success. Thousand Oaks, CA: Corwin.

Mateen, A. & Ali, K. 2017. Optimization strategies through big-data migration in distributed cloud databases. (In Proceedings of the 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, 21-22 September. Piscataway, NJ: IEEE. p. 96-99).

McNiff, J. 2009. You and your action research project. London: Routledge.

Meehan, J., Aslantas, C., Zdonik, S., Tatbul, N. & Du, J. 2017. Data Ingestion for the Connected World. Paper presented at CIDR '17, Chaminade, CA, 8-11 January. http://cidrdb.org/cidr2017/papers/p124-meehan-cidr17.pdf Date of access: 2019-04-12.

Merriam, S.B. & Tisdell, E.J. 2015. Qualitative research: A guide to design and implementation. Hoboken, NJ: Wiley.

Mertens, D.M. 2014. Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods. Los Angeles, CA: Sage.

Mertler, C.A. 2019. Action research: Improving schools and empowering educators. Los Angeles, CA: Sage.

Mistry, R. & Misner, S. 2014. Introducing Microsoft SQL Server 2014. Redmond, WA: Microsoft Press.

Moser, C.A. & Kalton, G. 2017. Survey methods in social investigation. London: Routledge.

Müller, S. & Bundesbank, D. 2019. Creating comprehensive data worlds using standardisation. Paper presented at the 9th IFC Conference on "Are post-crisis statistical initiatives completed?" Basel, 30-31 August. https://www.bis.org/ifc/publ/ifcb49_46.pdf Date of access: 2019-04-12.

Munhall, P.L. 1988. Ethical considerations in qualitative research. *Western Journal of Nursing Research*, 10(2):150-162.

Musante, K. 2014. Participant observation. (*In* Bernard, H.R. & Gravlee, C.C., *eds.* Handbook of methods in cultural anthropology. Lanham, MD: Rowman & Littlefield. p. 251-292).

Myers, M.D. 1997. Critical ethnography in information systems. *MIS Quarterly.* https://www.researchgate.net/publication/220260372_Qualitative_Research_in_Information_Systems Date of access: 2019-04-12.

Myers, M.D. 2004. Hermeneutics in information systems research. (*In* Mingers, J. & Willcocks, L.P., *eds.* Social theory and philosophy for information systems. Hoboken, NJ: Wiley. p. 103-128).

Myers, M.D. & Klein, H.K. 2011. A set of principles for conducting critical research in information systems. *MIS Quarterly*, 35(1):17-36.

Neergaard, M.A., Olesen, F., Andersen, R.S. & Sondergaard, J. 2009. Qualitative description: the poor cousin of health research? *BMC medical research methodology*, 9(1):52.

Nikulchev, E., Pluzhnik, E., Biryukov, D., Lukyanchikov, O. & Payain, S. 2015. Experimental study of the cloud architecture selection for effective big data processing. *International Journal of Advanced Computer Science and Applications*, 6(6):22-26.

O'Brien, R. 1998. An overview of the methodological approach of action research.

Oates, B.J. 2006. Researching information systems and computing. London: SAGE.

Obeidat, M., North, M., Richardson, R., Rattanak, V. & North, S. 2015. Business intelligence technology, applications, and trends. *International Management Review*, 11(2):47-56.

Oestreich, T.W. 2016. Magic quadrant for business intelligence and analytics platforms. *Analyst(s)*, 501. http://www.aditi.com.br/upload/kcfinder/files/Gartner%20Magic%20Quadrant%20BI%202016.pdf Date of access: 2019-04-12.

Olszak, C.M. 2014. Business Intelligence in cloud. *Polish Journal of Management Studies*, 10(2):115-125.

Orlikowski, W.J. & Baroudi, J.J. 1991. Studying information technology in organizations: Research approaches and assumptions. *Information systems research*, 2(1):1-28.

Parahoo, K. 2014. Nursing research: principles, process and issues. Basingstoke: Palgrave Macmillan.

Parks, D.M.D. 2017. Defining Data Science and Data Scientist. https://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=8211&context=etd Date of access: 2019-04-12.

Rad, R. 2014. Microsoft SQL Server 2014 Business Intelligence Development Beginner's Guide. Birmingham: Packt.

Rahman, N. & Rutz, D. 2019. Building data warehouses using automation. (*In* Rapid Automation: Concepts, Methodologies, Tools, and Applications. Hershey, PA: IGI Global. p. 735-759).

Rea, L.M. & Parker, R.A. 2014. Designing and conducting survey research: A comprehensive guide. Hoboken, NJ: Wiley.

Remler, D.K. & Van Ryzin, G.G. 2014. Research methods in practice: Strategies for description and causation: Sage Publications.

Richards, G., Yeoh, W., Chong, A.Y.L. & Popovič, A. 2019. Business intelligence effectiveness and corporate performance management: an empirical analysis. *Journal of Computer Information Systems*59(2). https://www.researchgate.net/publication/318797663_Business_Intelligence_Effectiveness_and_Corporate_Performance_Management_An_Empirical_Analysis Date of access: 2019-04-12.

Richards, L. 2014. Handling qualitative data: A practical guide. Los Angeles, CA: Sage.

Roiger, R.J. 2017. Data mining: a tutorial-based primer. Boca Raton, FL: CRC.

Rokach, L. & Maimon, O.Z. 2008. Data mining with decision trees: theory and applications. 2nd ed. Hackensack, NJ: World scientific.

Rudestam, K.E. & Newton, R.R. 2014. Surviving your dissertation: A comprehensive guide to content and process. Los Angeles, CA: Sage.

Rule, P. & John, V. 2011. Your guide to case study research. Pretoria: van Schaik.

Sale, H.B., Patil, D., Thube, S. & Student, B. 2018. Crime Prevention with Data Warehouse using OLAP through Business Intelligence. *International Journal of Engineering Science*, 8(2):16017-16019.

Saunders, M., Lewis, P. & Thornhill, A. 2007. Research methods. *Business Students 4th edition Pearson Education Limited, England*.

Saunders, M., Lewis, P. & Thornhill, A. 2009. Research methods for business students. Essex. *Financial Times/Prentice Hall*.

Sauter, V.L. 2014. Decision support systems for business intelligence. Hoboken, NJ: Wiley.

Seers, K. 2012. Qualitative data analysis. *Evidence-based nursing*, 15(1):2.

Sekaran, U. & Bougie, R. 2016. Research methods for business: A skill building approach. Hoboken, NJ: Wiley.

Sharma, G. 2017. Pros and cons of different sampling techniques. *International Journal of Applied Research*, 3(7):749-752.

Sherman, R. 2014. Business intelligence guidebook: From data integration to analytics. Amsterdam: Elsevier.

Silverman, D. 2015. Interpreting qualitative data. Los Angeles, CA: Sage.

Silverman, D. 2016. Qualitative research. Los Angeles, CA: Sage.

Simon, H.A. 1996. The sciences of the artificial. Cambridge, MA: MIT press.

Simons, H. 2014. Case study research: In-depth understanding in context. (*In* Leavy, P. *ed*. The Oxford handbook of qualitative research. Oxford: Oxford University Press. p. 455-470).

Sirin, E. & Karacan, H. 2017. A Review on Business Intelligence and Big Data. *International Journal of Intelligent Systems and Applications in Engineering*, 5(4):206-215.

Smith, J.A. 2015. Qualitative psychology: A practical guide to research methods. Los Angeles, CA: Sage.

Soon, L. & Fraser, C. 2011. A Knowledge Management Approach: Business Intelligence in an Intranet Data Warehouse. *GSTF Journal on Computing*, 1(2):190-195.

Spradley, J.P. 1979. The ethnographic interview. Long Grove, IL: Waveland Press, Inc.

Spruit, M.R. & Sacu, C. 2015. DWCMM: The Data Warehouse Capability Maturity Model. *Journal of Universal Computer Science*, 21(11):1508-1534.

Stefanovic, N. 2015. Collaborative predictive business intelligence model for spare parts inventory replenishment. *Computer Science and Information Systems*, 12(3):911-930.

Stewart, D.W. & Shamdasani, P.N. 2014. Focus groups: Theory and practice. Los Angeles, CA: Sage.

Stockdale, R., Arefin, M.S., Hoque, M.R. & Bao, Y. 2015. The impact of business intelligence on organization's effectiveness: an empirical study. *Journal of Systems and Information Technology*, 17(3):263-285.

Straub, D., Boudreau, M.C. & Gefen, D. 2004. Validation guidelines for IS positivist research. *Communications of the Association for Information Systems*, 13(1):380-427.

Strauss, A. & Corbin, J. 1994. Grounded theory methodology. (*In* Denzin, N.K. & Lincoln, Y.S., *eds*. Handbook of qualitative research. Thousand Oaks, CA: Sage. p. 273-285).

Tavallaei, R., Shokohyar, S., Moosavi, S.M. & Sarfi, Z. 2015. Assessing the evaluation models of business intelligence maturity and presenting an optimized model. *International Journal of Management, Accounting and Economics*, 2(9):1005-1019.

Taylor, S.J., Bogdan, R. & DeVault, M. 2015. Introduction to qualitative research methods: A guidebook and resource. Hoboken, NJ: Wiley.

Thomas, G. 2015. How to do your case study. Los Angeles, CA: Sage.

Thorne, S. 2000. Data analysis in qualitative research. *Evidence-based nursing*, 3(3):68-70.

Tole, A.A. 2014. Cloud Computing and Business Intelligence. *Database Systems Journal*, 5(4):49-58.

Tomingas, K., Kliimask, M. & Tammet, T. 2015. Data integration patterns for data warehouse automation. (*In* Bassiliades, N., Ivanovic, M., Kon-Popovska, M., Manolopoulos, Y., Palpanas, T., Trajcevski, G. & Vakali, A., *eds*. New trends in database and information systems II. Cham: Springer). https://www.researchgate.net/publication/282379282_Data_Integration_Patterns_for_Data_Warehouse_Automation Date of access: 2019-04-12.

Vaishnavi, V.K. & Kuechler, W. 2015. Design science research methods and patterns: innovating information and communication technology. Boca Raton, FL: CRC.

Velimeneti, S. 2016. Data Migration from Legacy Systems to Modern Database. https://repository.stcloudstate.edu/cgi/viewcontent.cgi?article=1054&context=mme_etds Date of access: 2019-04-12.

Venable, J. 2006. A framework for design science research activities. (*In* Proceedings of the 2006 Information Resource Management Association Conference. Hershey, PA: Idea Group). http://desrist.org/desrist/content/design-science-research-in-information-systems.pdf Date of access: 2019-04-12.

Victor, C.P., Treschuk, J., Peoples, M. & Lewis, M. 2016. Interpretivist Paradigm: Essential for nursing research. Paper presented at the 7th Annual Qualitative Report Conference, Nova Southeastern University, 14-16 January.

Walsham, G. 1995. The emergence of interpretivism in IS research. *Information systems research*, 6(4):376-394.

Wang, Y., Kung, L. & Byrd, T.A. 2018. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, 126:3-13.

Watson, H.J. 2014. Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*, 34:1247-1268.

Wazurkar, P., Bhadoria, R.S. & Bajpai, D. 2017. Predictive analytics in data science for business intelligence solutions. (*In* Proceedings of the 2017 7th International Conference on

Communication Systems and Network Technologies (CSNT). Piscataway, NJ: IEEE. p. 367-370).

Wieder, B. & Ossimitz, M.L. 2015. The impact of Business Intelligence on the quality of decision making: a mediation model. *Procedia Computer Science*, 64:1163-1171.

Witten, I.H., Frank, E., Hall, M.A. & Pal, C.J. 2016. Data Mining: Practical machine learning tools and techniques. Amsterdam: Morgan Kaufmann.

Yasser, A.L.H. & Zota, R.D. 2016. Implementing Business Intelligence System-Case Study. *Database Systems Journal*, 7(1):35-44.

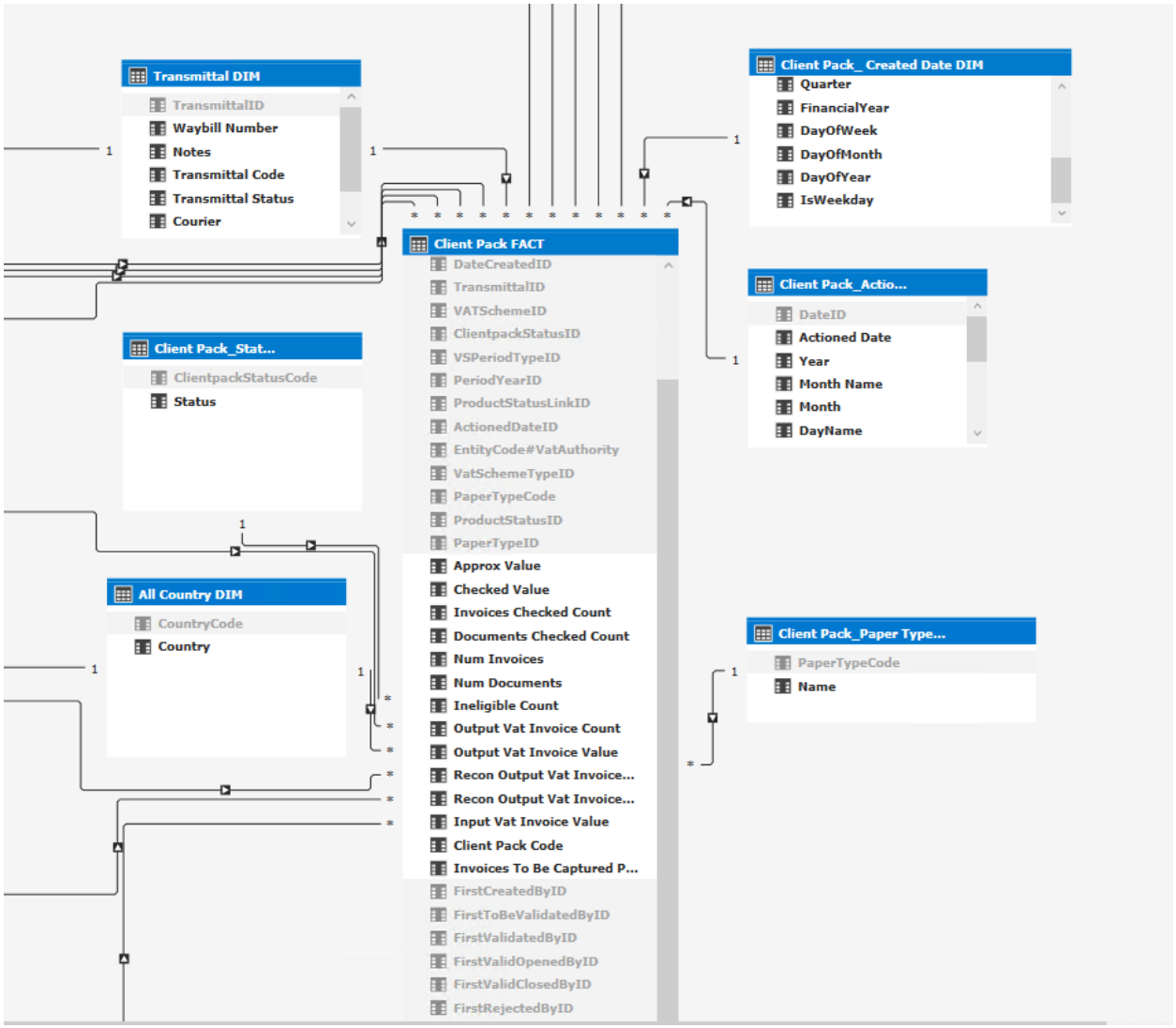
Yin, R.K. 2011. Applications of case study research. Los Angeles, CA: Sage.

Zainal, Z. 2007. Case study as a research method. *Jurnal Kemanusiaan*, 5(1):1-6.

Zikmund, W.G., Babin, B.J., Carr, J.C. & Griffin, M. 2013. Business research methods. 9th ed. Mason, OH : Cengage: Cengage Learning.

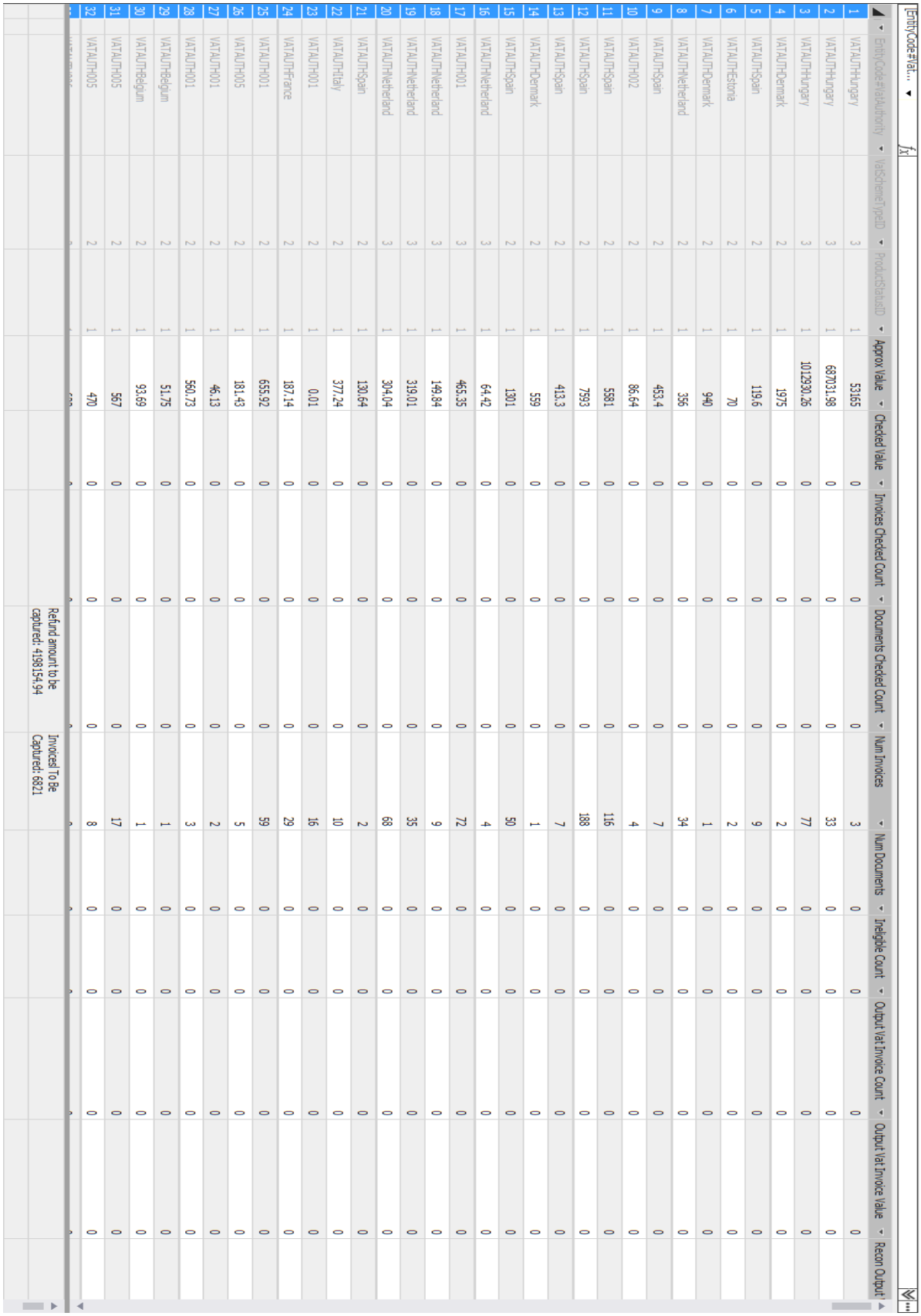
APPENDIX A: ANALYSIS DIMENSION STRUCTURE

Last Updated: 12 August 2019



APPENDIX B: ANALYSIS TABLE STRUCTURE

Last Updated: 12 August 2019



	EntityCode#VatAuthority	VatSchemeTypeID	ProductStatusID	Approx Value	Checked Value	Invoices Checked Count	Documents Checked Count	Num Invoices	Num Documents	Ineligible Count	Output Vat Invoice Count	Output Vat Invoice Value	Recon Output
1	VATAUTHHungary	3	1	53165	0	0	0	3	0	0	0	0	0
2	VATAUTHHungary	3	1	687031.98	0	0	0	33	0	0	0	0	0
3	VATAUTHHungary	3	1	1012930.26	0	0	0	77	0	0	0	0	0
4	VATAUTHDenmark	2	1	1975	0	0	0	2	0	0	0	0	0
5	VATAUTHSpain	2	1	119.6	0	0	0	9	0	0	0	0	0
6	VATAUTHEstonia	2	1	70	0	0	0	2	0	0	0	0	0
7	VATAUTHDenmark	2	1	940	0	0	0	1	0	0	0	0	0
8	VATAUTHNetherland	2	1	356	0	0	0	34	0	0	0	0	0
9	VATAUTHSpain	2	1	453.4	0	0	0	7	0	0	0	0	0
10	VATAUTH002	2	1	86.64	0	0	0	4	0	0	0	0	0
11	VATAUTHSpain	2	1	5581	0	0	0	116	0	0	0	0	0
12	VATAUTHSpain	2	1	7593	0	0	0	188	0	0	0	0	0
13	VATAUTHSpain	2	1	413.3	0	0	0	7	0	0	0	0	0
14	VATAUTHDenmark	2	1	559	0	0	0	1	0	0	0	0	0
15	VATAUTHSpain	2	1	1301	0	0	0	50	0	0	0	0	0
16	VATAUTHNetherland	3	1	64.42	0	0	0	4	0	0	0	0	0
17	VATAUTH001	3	1	465.35	0	0	0	72	0	0	0	0	0
18	VATAUTHNetherland	3	1	149.84	0	0	0	9	0	0	0	0	0
19	VATAUTHNetherland	3	1	319.01	0	0	0	35	0	0	0	0	0
20	VATAUTHNetherland	3	1	304.04	0	0	0	68	0	0	0	0	0
21	VATAUTHSpain	2	1	130.64	0	0	0	2	0	0	0	0	0
22	VATAUTHItaly	2	1	377.24	0	0	0	10	0	0	0	0	0
23	VATAUTH001	2	1	0.01	0	0	0	16	0	0	0	0	0
24	VATAUTHFrance	2	1	187.14	0	0	0	29	0	0	0	0	0
25	VATAUTH001	2	1	655.92	0	0	0	59	0	0	0	0	0
26	VATAUTH005	2	1	181.43	0	0	0	5	0	0	0	0	0
27	VATAUTH001	2	1	46.13	0	0	0	2	0	0	0	0	0
28	VATAUTH001	2	1	560.73	0	0	0	3	0	0	0	0	0
29	VATAUTHBelgium	2	1	51.75	0	0	0	1	0	0	0	0	0
30	VATAUTHBelgium	2	1	93.69	0	0	0	1	0	0	0	0	0
31	VATAUTH005	2	1	567	0	0	0	17	0	0	0	0	0
32	VATAUTH005	2	1	470	0	0	0	8	0	0	0	0	0

Refund amount to be captured: 4198154.94

Invoices To Be Captured: 6821

APPENDIX C: SENIOR MANAGEMENT BUSINESS INTELLIGENCE ARCHITECHTURE INTERVIEW QUESTION LIST

Last Updated: 12 August 2019

Question(s)	**Objective addressed**	**Literature**
What led to the consideration of implementing a business intelligence (BI) solution in 2?	Leading up to E01	Olszak (2014:116) state that introducing an integrated BI solution into an organisation can result in effective decision-making operations for distributed organisations.
How was the BI solution implemented in the organisation in 2?	E01: To describe the BI architecture of the small organisation in its first year of existence.	According to Watson (2014:1247) describes, Business intelligence allows for reliable and accurate analysis of day to day operations that results in more effective management.
Which tools were used in association with the reporting purposes throughout the organisation when the organisation was considered to be of start-up size?	E01: To describe the BI architecture of the small organisation in its first year of existence.	According to Agarwal *et al.* (2015:1), One of the most important aspects of an organisation is to present data through a reporting solution.
Was internal staff used to develop and manage the BI solution? If so, what were their initial roles in the organisation? If no, please specify the selection process of external staff selected for the project.	E03: To inspect and report on what impact the integration of the data from the two organisations had on the BI architecture and how the BI architecture had to be adjusted	Gupta *et al.* (2015:455) states BI specialists normally have skillsets associated with universities certifications such as BI courses in undergraduate; Master of Science and master's in business administration (MBA) courses.
What was the initial timeframe of completion for the BI solution project and was the project completed in the given timeframe?	E03: To inspect and report on what impact the integration of the data from the two organisations had on the BI architecture and how the BI architecture had to be adjusted	Kimball Life cycle (Kimball *et al.*, 2015:33) states that the timeframe given to develop a BI solution is critically important as allocating inaccurate time could mean complications to the entire BI solution development

Question(s)	**Objective addressed**	**Literature**
What, if any, complications arose during the BI solution project completion?	E03: To inspect and report on what impact the integration of the data from the two organisations had on the BI architecture and how the BI architecture had to be adjusted	Kimball Life cycle (Kimball *et al.*, 2015:33) states that the timeframe given to develop a BI solution is critically important as allocating inaccurate time could mean complications to the entire BI solution development
After the BI solution project was completed, which reporting solutions is being used throughout the company?	E03: To inspect and report on what impact the integration of the data from the two organisations had on the BI architecture and how the BI architecture had to be adjusted	According to Agarwal *et al.* (2015:1), One of the most important aspects of an organisation is to present data through a reporting solution.
Did the BI solution implemented in 2-18 provide more accurate information for decision making?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Erskine *et al.* (2015:1) states that reporting is a big aspect of decision making in organisations, without a reporting solution, decision making will be a difficult task to implement.
Did the newly implemented BI solution provide a more extensive variety of decision-making options to the business users?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Erskine *et al.* (2015:1) states that reporting is a big aspect of decision making in organisations, without a reporting solution, decision making will be a difficult task to implement
What would you say was the overall impression and impact that the BI solution had on the organisation and business users using the updated BI solution?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Gadde (2016:15) state that there are a few advantages to using SSRS and Power BI reporting solutions in an organisational environment such as data from multiple locations can be depicted in a graphical manner and is easy to read by management staff.

APPENDIX D: SENIOR MANAGEMENT DATA MIGRATION INTERVIEW QUESTION LIST

Last Updated: 12 August 2019

Question(s)	**Objective addressed**	**Literature**
What was the reason for the data migration from one organisation to another?	Lead up to E02	According to Lu and Holubová (2017:602), big data became a phenomenon in organisations through the years. Big data makes it possible for multiple approaches and technologies to interact and integrate with each other
What was the requirement criteria for selecting a project manager to manage the data migration project?	E02: To determine how data from two different database platforms (Microsoft and Oracle) may be integrated	Making human behavioural patterns and properties measurable, that in most cases could not have been measured before is the biggest reason for data migration (Franke *et al.*, 2016:372).
To the best of your knowledge, what was the initial timeframe of completion for the data migration project and was the project completed in the given timeframe?	E02: To determine how data from two different database platforms (Microsoft and Oracle) may be integrated	According to Kuderu and Kumari (2016:506), because of large quantities of data being integrated by organisations, data migration has become an important aspect in the organisational environment and an applicable timeframe need to be associated with the data migration.
If any, what complications was there with regards to the data migration project?	E02: To determine how data from two different database platforms (Microsoft and Oracle) may be integrated	Big data not only provides statistical inference opportunities, but even bigger analytical challenges as the dataset analysed is much bigger than normal (Franke *et al.*, 2016:371)
How were the problems as mentioned in question 2.4 addressed?	E02: To determine how data from two different database platforms (Microsoft and Oracle) may be integrated	Parks (2017:9) states, data migration specialists can be defined as individuals with a set of skill able to solve problems by associated with data migration from one DBMS to another.

Question(s)	**Objective addressed**	**Literature**
What would you do differently, if given the same possibility to migrate data in such scale from one organisations database management system platform to another again? Please explain your answer in full.	E02: To determine how data from two different database platforms (Microsoft and Oracle) may be integrated	After data migration has completed, organisations would analyse other ways of data migration by looking at the organisations data, analysing the data and building reports from the analytical data to make insightful organisational decisions (Agasisti & Bowers, 2017:184)

APPENDIX E: TECHNICAL STAFF BUSINESS INTELLIGENCE ARCHITECHTURE INTERVIEW QUESTION LIST

Last Updated: 12 August 2019

Question(s)	**Objective addressed**	**Literature**
Which of the BI solution objectives regarding design, development and implementation were clear and easy to grasp and which was not?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Gosain and Singh (2015:306) states that creating a bus matrix is most probably the most important step to perform using the Kimball DW design. This will help clear any questions regarding the star schema design and identify all fact tables and dimensional tables associated with the DW design.
What time was allocated for design, development and implementation of the BI solution?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Kimball Life cycle (Kimball *et al.*, 2015:33) states that the timeframe given to develop a BI solution is critically important, allocating inaccurate time could mean complications to the entire BI solution development
Explain the effectiveness of the extract, transform and load (ETL) tool used to clean, transform and load data from different sources into one central location.	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	According to Mistry and Misner (2014:1) states that Microsoft SSIS is a ETL tool available to retrieve new insights into big data with simplicity, by connecting to multiple data sources, retrieving the data, implementing an ETL process and providing cleaned, meaningful data back in one source of truth.
What impact did the design, development and implementation of analysis tool, used to analyse data from the DW, have on the organisation's decision making?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	According to (Ferrari & Russo, 2015), there are two different types of analysis categories in the Microsoft SSAS space to consider when creating a SSAS Solution namely, Multidimensional modelling and Tabular modelling

Question(s)	**Objective addressed**	**Literature**
Which organisational reporting needs, aligned with the policies and protocols of the organisation, is met by reporting tools connecting to the BI architecture?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Erskine *et al.* (2015:1) states that reporting is a big aspect of decision making in organisations, without a reporting solution, decision making will be a difficult task to implement
How does the current report development process differ from prE0BI architecture?	E01: To describe the BI architecture of the small organisation in its first year of existence.	According to Agarwal *et al.* (2015:1), one of the most important aspects of an organisation is to present data through a reporting solution.
How did you experience the transition from the old reporting tools to the new reporting tools?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Erskine *et al.* (2015:1) states that reporting is a big aspect of decision making in organisations, without a reporting solution, decision making will be a difficult task to implement
What was the reaction from staff utilising the newly implemented reporting tools?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Gadde (2016:15) states that there are a few advantages to using SSRS and Power BI reporting solutions in an organisational environment such as data from multiple locations can be depicted in a graphical manner and is easy to read by management staff.
What is the procedure to maintain the newly implemented BI solution?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	After the development of the DW is completed the BI developer will be in change to maintain the DW from time to time and the project manager will place provisions for excessive data growth during the DW existence Kimball Life cycle (Kimball *et al.*, 2015:33).

APPENDIX F: TECHNICAL STAFF DATA MIGRATION INTERVIEW QUESTION LIST

Last Updated: 12 August 2019

Question(s)	**Objective addressed**	**Literature**
What was the initial timeframe of completion regarding the data migration project and was the project completed in the given timeframe?	E03: To inspect and report on what impact the integration of the data from the two organisations had on the BI architecture and how the BI architecture had to be adjusted	According to Kuderu and Kumari (2016:506), because of large quantities of data being integrated by organisations, data migration has become an important aspect in the organisational environment and an applicable timeframe need to be associated with the data migration.
If any, what complications were there with regards to the data migration project?	E03: To inspect and report on what impact the integration of the data from the two organisations had on the BI architecture and how the BI architecture had to be adjusted	Big data not only provides statistical inference opportunities, but even bigger analytical challenges as the dataset analysed is much bigger than normal (Franke *et al.*, 2016:371)
How were the problems as mentioned in question 2.3 addressed?	E03: To inspect and report on what impact the integration of the data from the two organisations had on the BI architecture and how the BI architecture had to be adjusted	Parks (2017:9) states that data migration specialists can be defined as individuals with a set of skill able to solve problems associated with data migration from one DBMS to another.
What implications to business operations, if any, did the newly imported data have on the BI architecture?	E03: To inspect and report on what impact the integration of the data from the two organisations had on the BI architecture and how the BI architecture had to be adjusted	Nikulchev *et al.* (2015:22) states that the extensive growth in big data cause problems for developers and BI solutions, since the structural distribution of classical query optimisation theories do not accommodate query depth.

Question(s)	**Objective addressed**	**Literature**
In what manner was the data structure of the data imported different than the organisations data structure and database management system platform?	E02: To determine how data from two different database platforms (Microsoft and Oracle) may be integrated	Every organisations use unique manners to store data into DBMS platforms and so combining two data structures could be a complex task (Albano *et al.*, 2015:1)
What would you do differently, if given the same possibility to migrate data in such scale from one organisations data base to another again?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	After data migration has completed, organisations would analyse other ways of data migration by looking at the organisations data, analysing the data and building reports from the analytical data to make insightful organisational decisions (Agasisti & Bowers, 2017:184)

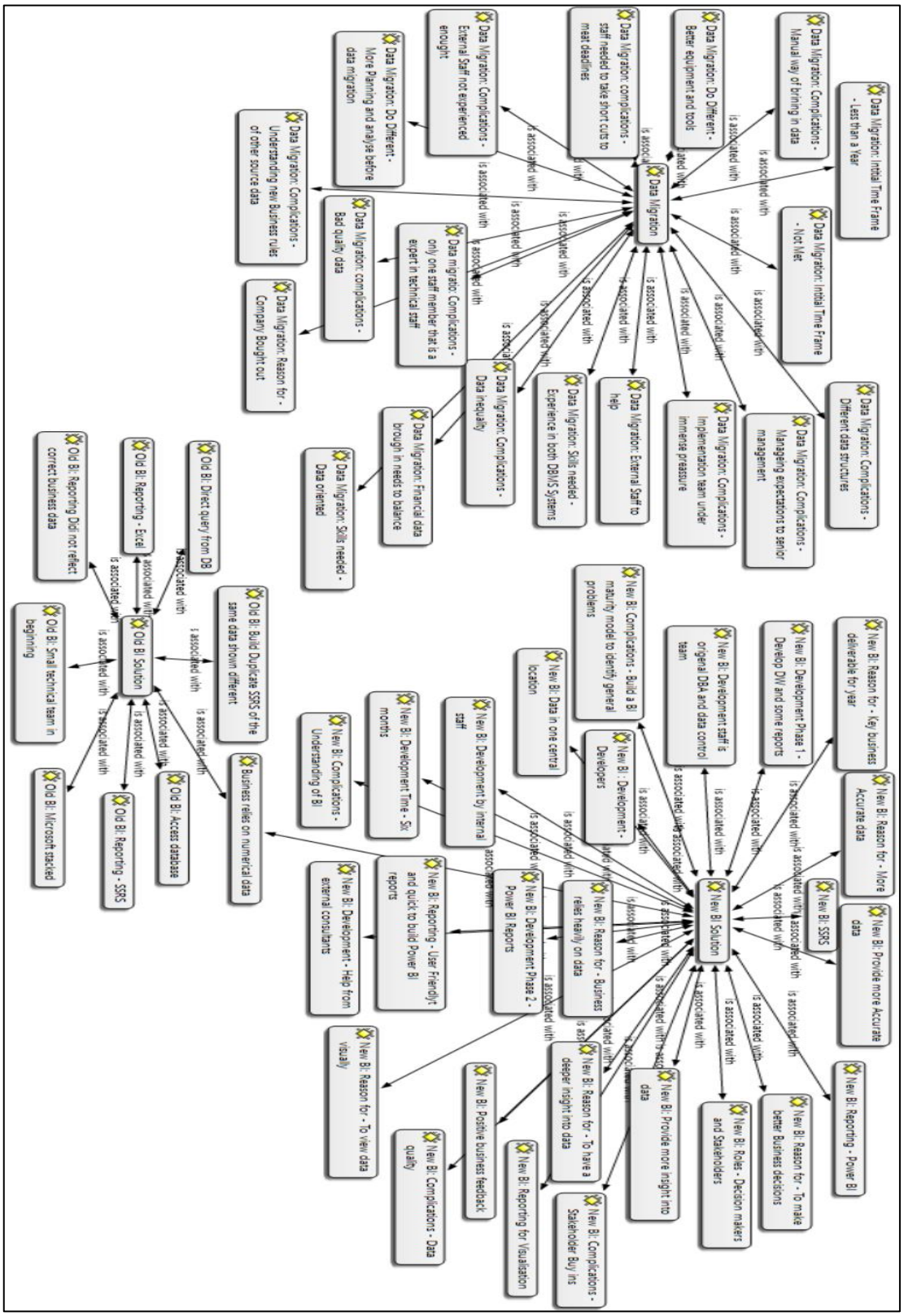
APPENDIX G: DATA CAPTURER INTERVIEW QUESTION LIST

Last Updated: 12 August 2019

Question	**Objective addressed**	**Literature**
What shortcomings did the old reporting solutions have?	E01: To describe the BI architecture of the small organisation in its first year of existence.	Erskine *et al.* (2015:1) states that reporting is a big aspect of decision making in organisations, without a reporting solution, decision making will be a difficult task to implement
Tell me about how you experienced the changeover and adjustment to the new reporting solution.	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	By the use of a SSAS cubes data sources, SSRS presents data in a visual perspective for organisations to make insightful decision making (Jabbar, 2015:105).
How do you feel about the new reporting solution? Please explain your answer.	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	By the use of a SSAS cubes data sources, SSRS presents data in a visual perspective for organisations to make insightful decision making (Jabbar, 2015:105).
How is the new reporting solution better or worse? Please explain your answer	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Gadde (2016:15) state that there are a few advantages to using SSRS reporting solutions in an organisational environment such as data from multiple locations can be depicted in a graphical manner and is easy to read by management staff.
What, if anything, would you change to the new reporting structure?	E04: To describe the personal experiences and involvement of stakeholders in the course of developing a new BI solution in the organisation.	Gadde (2016:15) state that there are a few advantages to using SSRS reporting solutions in an organisational environment such as data from multiple locations can be depicted in a graphical manner and is easy to read by management staff.

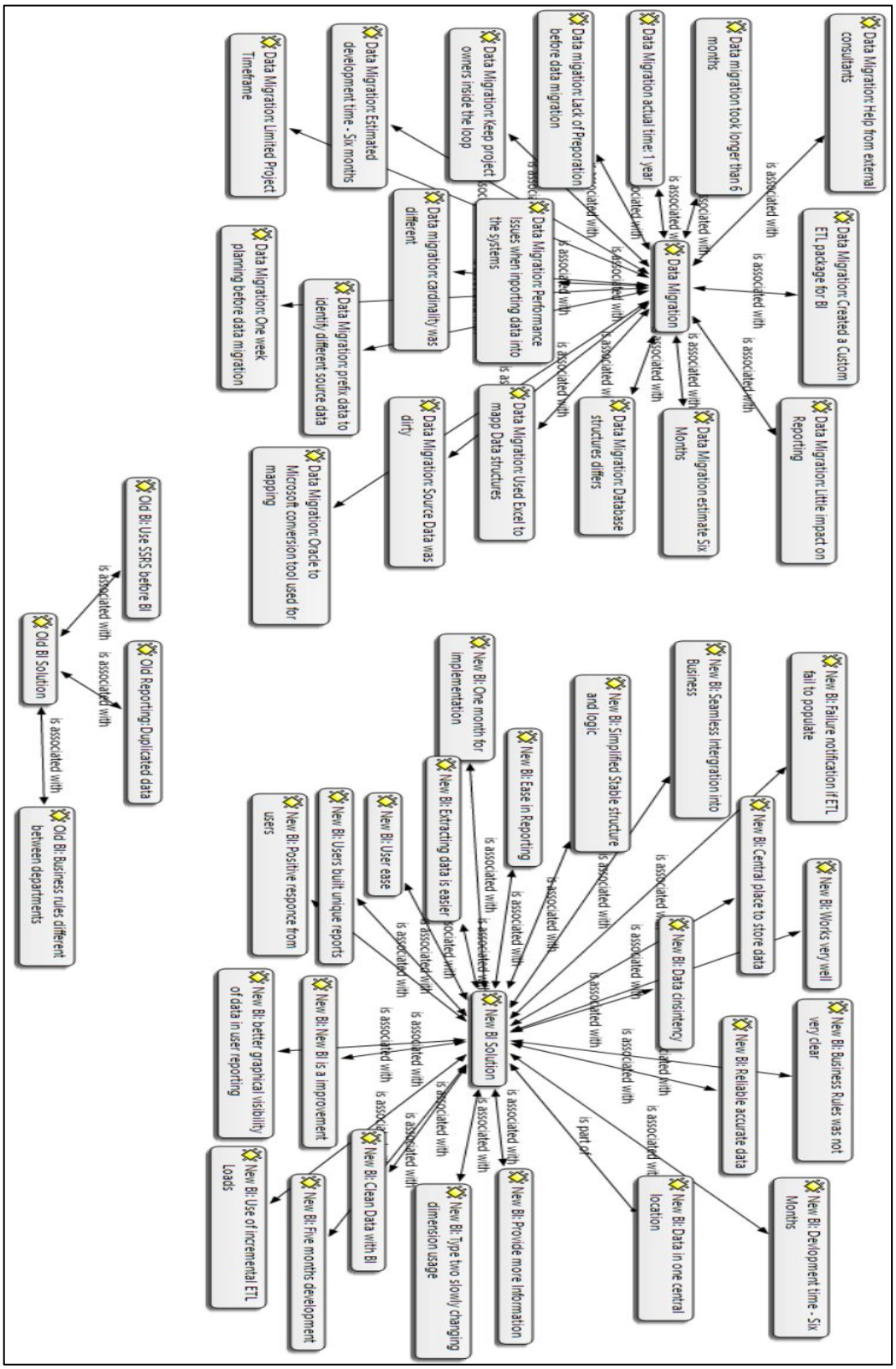
APPENDIX H: SENIOR MANAGEMENT PARTICIPANT FULL CODE NETWORK

Last Updated: 12 August 2019



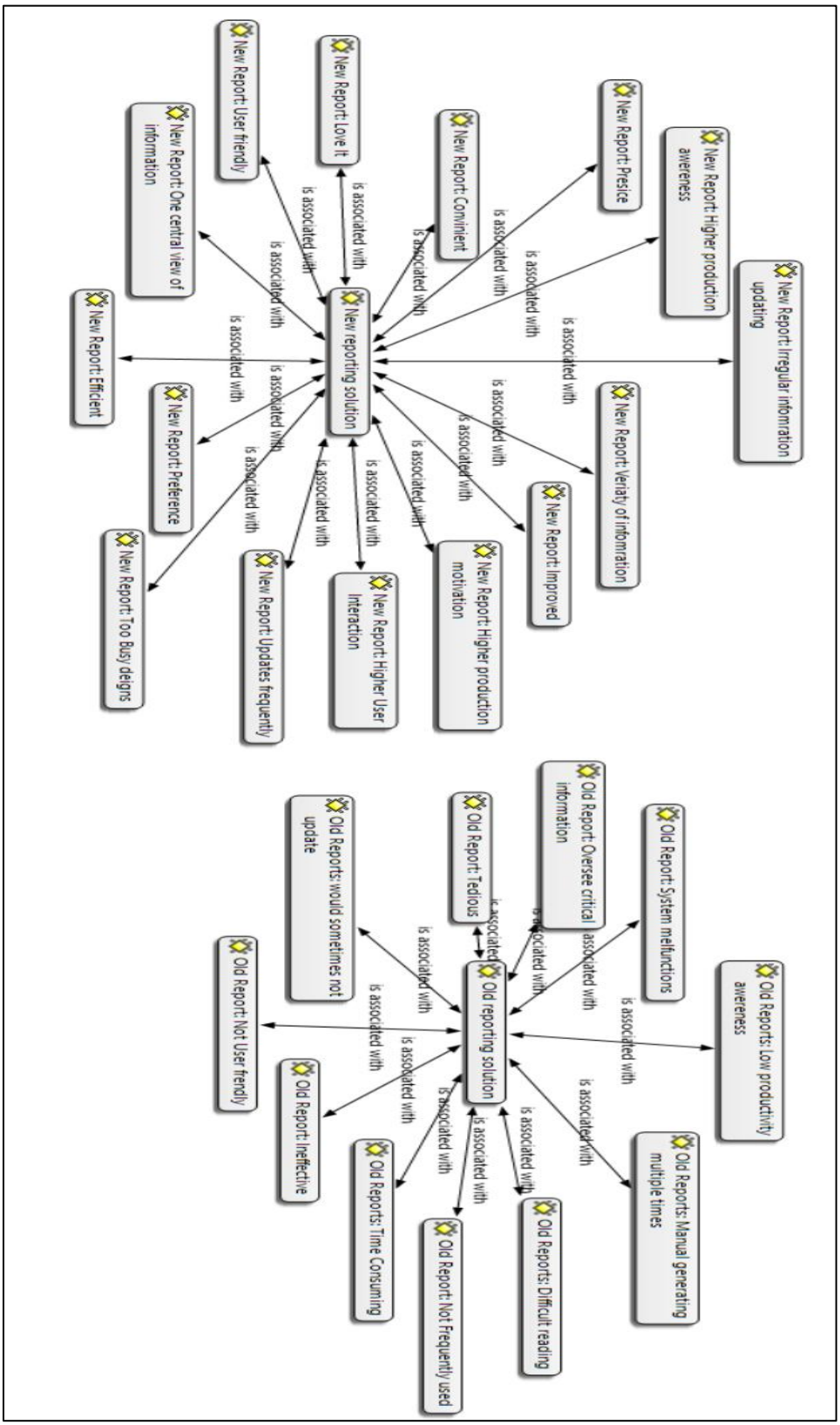
APPENDIX I: TECHNICAL STAFF PARTICIPANT FULL CODE NETWORK

Last Updated: 12 August 2019



APPENDIX J: DATA CAPTURER PARTICIPANT FULL CODE NETWORK

Last Updated: 12 August 2019



APPENDIX K: RIGOUR

Last Updated: 12 August 2019

Participant Type	Topic	Question Number	Researcher	Controller	Added /Altered codes
Senior Management	BI Solution	1.1	1	1	
Senior Management	BI Solution	1.2	0	1	Old BI: Reason For - Data Growth Problem (Added)
Senior Management	BI Solution	1.3	1	1	
Senior Management	BI Solution	1.4	0	1	Development staff skills originally DBA and Data control (Altered)
Senior Management	BI Solution	1.5	1	1	
Senior Management	BI Solution	1.5	1	1	
Senior Management	BI Solution	1,6	1	1	
Senior Management	BI Solution	1.7	1	1	
Senior Management	BI Solution	1.8	0	1	New BI: Complications - Understanding of benefits of BI (Altered)
Senior Management	BI Solution	1.9	1	1	
Senior Management	BI Solution	1.9	1	1	
Senior Management	BI Solution	1.10	0	1	New BI: Reason for - Business relies heavily on correct and consistent data (Altered)
Senior Management	Data Migration	2.1	0	1	Data Migration: All data including financial data brought through needs to balance

Participant Type	Topic	Question Number	Researcher	Controller	Added /Altered codes
Senior Management	Data Migration	2.2	0	1	Data Migration: Complications: Initial Time Frame - Not Met
Senior Management	Data Migration	2.3	1	1	
Senior Management	Data Migration	2.4	1	1	
Senior Management	Data Migration	2.4	1	1	
Senior Management	Data Migration	2.5	1	1	
Senior Management	Data Migration	2.6	1	1	
Technical staff	BI Solution	1.1	0	1	New BI: Complication: Business Rules was not very clear (Altered)
Technical staff	BI Solution	1.2	1	1	
Technical staff	BI Solution	1.3	1	1	
Technical staff	BI Solution	1.4	1	1	
Technical staff	BI Solution	1.5	0	1	New BI: User ease of use (Altered)
Technical staff	BI Solution	1.6	1	1	
Technical staff	BI Solution	1.7	0	1	New BI: Seamless transition and integration of reporting solution into Business (Altered)
Technical staff	BI Solution	1.8	0	1	New BI: Provide more Information faster (Altered)
Technical staff	BI Solution	1.9	1	1	
Technical staff	Data Migration	2.1	0	1	Data Migration: Complication - Actual Time: 1 year (Altered)

Participant Type	Topic	Question Number	Researcher	Controller	Added /Altered codes
Technical staff	Data Migration	2.2	0	1	Data migration: data anomalies and cardinality (Altered)
Technical staff	Data Migration	2.2	1	1	
Technical staff	Data Migration	2.3	1	1	
Technical staff	Data Migration	2.3	1	1	
Technical staff	Data Migration	2.4	1	1	
Technical staff	Data Migration	2.5	1	1	
Technical staff	Data Migration	2.6	1	1	
Data Capturer	BI Solution	1.1	0	1	Old BI: Not frequently updated (Added)
Data Capturer	BI Solution	1.1	1	1	
Data Capturer	BI Solution	1.2	1	1	
Data Capturer	BI Solution	1.3	1	1	
Data Capturer	BI Solution	1.4	1	1	
Data Capturer	BI Solution	1.4	1	1	