A FRAMEWORK TO MANAGE SENSITIVE INFORMATION DURING ITS MIGRATION BETWEEN SOFTWARE PLATFORMS

by

OLUSEGUN ADEMOLU AJIGINI

submitted in accordance with the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in the subject

INFORMATION SYSTEMS

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF J A VAN DER POLL

CO-SUPERVISOR: PROF J H KROEZE

JUNE 2016

ABSTRACT

Software migrations are mostly performed by organisations using migration teams. Such migration teams need to be aware of how sensitive information ought to be handled and protected during the implementation of the migration projects. There is a need to ensure that sensitive information is identified, classified and protected during the migration process. This thesis suggests how sensitive information in organisations can be handled and protected during migrations by using the migration from proprietary software to open source software to develop a management framework that can be used to manage such a migration process.

A rudimentary management framework on information sensitivity during software migrations and a model on the security challenges during open source migrations are utilised to propose a preliminary management framework using a sequential explanatory mixed methods case study. The preliminary management framework resulting from the quantitative data analysis is enhanced and validated to conceptualise the final management framework on information sensitivity during software migrations at the end of the qualitative data analysis. The final management framework is validated and found to be significant, valid and reliable by using statistical techniques like Exploratory Factor Analysis, reliability analysis and multivariate analysis as well as a qualitative coding process.

Keywords: Closed Source Software; Information classification; Information protection; Information security; Information sensitivity; IP Protection; IS Migration; IS theory development; IT control frameworks (CobiT, ITIL, ISO17799); Open source; Sensitive information.

DECLARATION

I declare that this research entitled *A framework to manage sensitive information during its migration between software platforms* is my own work and that all the sources that l have used or quoted have been indicated and acknowledged by means of complete references. I further declare that l have not previously submitted this work or part of it, for examination at UNISA or another qualification or at any other higher education institution.

Olusegun Ademolu Ajigini

PUBLICATIONS

The following publications emanated from this research:

Ajigini, O. A., Van der Poll, J. A. & Kroeze, J. H., 2012, 'Towards a Management Framework to protect Sensitive Information during Migrations', *The 2nd International Conference on Design and Modelling in Science, Education and Technology (DeMSet)*, Orlando, Florida, USA, (ISBN-13: 978-1-936338-76-4) ISBN-13: 978-1-936338-76-4 CD / ISBN-13, 6-13.

Ajigini, O. A., Van Der Poll, J. A. & Kroeze, J. H., 2014. 'Towards a Model on Security Challenges during Closed Source Software to OSS Migrations.' *The 9th International Conference for Internet Technology and Secured Transactions (ICITST Proceedings)*, London, UK, 8 – 10 Dec., 275-284. ISBN: 978-1-908320-31-5.

Ajigini, O. A., Van der Poll, J. A. & Kroeze, J. H., 2016. 'A Framework to Manage Sensitive Information during its Migration between Software Platforms.' *The African Journal of Information Systems*, Vol. 8, Issue 2, Article 2, 21 – 44. ISSN 1936-0282.

DEDICATION

This thesis is dedicated to my late father Pa Michael Oni Ajigini, my mother Mrs Beatrice Modupe Ajigini, my wife Annah and my children, Thabang, Omobusola and Khumo

ACKNOWLEDGMENTS

I wish to express my appreciation and profound gratitude to the following people for their assistance towards the completion of my thesis and degree:

1. My promoters, Prof J. A. van der Poll and Prof J. H. Kroeze both of University of South Africa (UNISA) for their excellent academic guidance, motivation, and support. I thank them greatly for imparting some of their knowledge and expertise of ICT research aspects to me.
2. I also acknowledge the guidance of Prof Sam Lubbe who was the initial supervisor of this research.
3. Mr Hennie Gerber, the UNISA Statistical Expert, who assisted me greatly in the quantitative data analysis section of the thesis.
4. All the employees of the following Government Departments/Agencies: State Information Technology Agency; Presidential National Commission; South African Department of Public Works; National Libraries of South Africa; South African Department of Arts and Culture; and South African Department of Social Development, who assisted me with the documents, information and study related interviews. I thank you all for your contribution, advice and support during the research work.
5. Special thanks to my beloved wife, Annah and my lovely children, Thabang, Omobusola and Khumo for their understanding, support, encouragement, mutual love and admiration throughout the journey of my life.
6. To my beloved parents, my late father Pa Michael Oni Ajigini and my mother Beatrice Modupe Ajigini for their advice, encouragement, guidance, love and financial support throughout my educational career. I will not forget my father's advice to us that 'we should

know the children thou art you'. I missed him a lot and May his gentle soul rest in perfect peace.

7. To other people who have contributed to my thesis one way or the other but whose names l have not mentioned earlier on, whom l thank for all their wonderful support and dedication towards the realisation of my thesis and degree.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Purpose of the Research

1.1.1 Introduction

This is a thesis on the development of a framework to manage sensitive information during its migration between software platforms. The thesis develops a management framework that can be used to protect and handle sensitive information during migration of software platforms. In this research, the author develops and validates a management framework for the migration of sensitive information during the migration of platforms by using a sequential explanatory mixed methods case study approach.

This chapter clarifies the statement of the problem investigated in this work by stating the needs in developing a management framework to manage sensitive information during its migration between software platforms. The research aim and objectives of the research which lead to the contribution of theory development in this area of Information Systems (IS) are further stated. The primary research and secondary research questions are enumerated with respect to the research work undertaken. The migration problems encountered by some South African Government Departments and Parastatals are further highlighted in this chapter. Specifically, the migration from Closed Source System (CSS) to Open Source System (OSS) is used to conceptualise resolving the research problem. A synopsis of all the chapters of the thesis is provided at the end of this chapter.

The study concentrates on South African government departments and parastatals that have performed software migrations. The main focus is the development and validation of a management framework that can be used to protect and handle sensitive information during its migration between software platforms. A good example of such platform migration is from Closed Source Software (CSS) to Open Source Software (OSS), also known as Free Open Source Software (FOSS) (Sarrab *et al.* 2013; Hudson 2015).

In South Africa, examples of such platform migrations include but are not limited to:

(a) migrations from proprietary systems to open source systems conducted during the eNaTIS migration by the Department of Transport (IT Web 2007),

(b) State Information Technology Agency (SITA) migration to FOSS (GITOC 2003),

(c) Presidential National Commission (PNC) migration to FOSS (PNC 2007),

(d) National Libraries of South Africa (NLSA) migration to FOSS (Novell Connection 2009),

(e) National Department of Arts and Culture migration to FOSS (S. Phala, *pers. comm.*)

(f) South Africa Department of Public Works migration to an open source asset management system (B. Zwane, *pers. comm.*).

SITA comprises about 3000 employees and provides ICT services to more than 30 national government departments, 80 provincial government departments and seeks to serve more than 230 city councils and local authorities (GITOC 2003). SITA has been tasked to lead and support the government institutions wishing to implement Free Open Source Software (FOSS) since 2001 (GITOC 2003).

However, based on a comprehensive review of articles in the literature, there is no management framework in place to guide SITA in protecting sensitive information during the software migrations. Therefore this research aims to improve this situation by developing and validating a management framework to manage sensitive information during software migrations.

The quantitative and qualitative research data are gathered from seven different organisations, namely State Information Technology Agency (SITA); South African Revenue Services (SARS); Presidential National Commission (PNC); National Libraries of South Africa, South African Department of Arts and Culture; South African Department of Public Works and South African Department of Social Development. The reasons why some of these Government organisations are chosen for conducting the research are explained in the paragraphs that follow.

The researcher performed this research in SITA because the South African government, through the State Information Technology Agency (SITA) migrated desktop applications from proprietary to open source platform in 2008 (SITA FOSS Focus 2009). SITA launched an initiative to explore possibilities of using FOSS in its environment and in the broader Government in October 2004. Consequently, SITA formed partnerships with certain South African government departments, national and provincial.

SITA migrated all SITA users from the current environment to the FOSS platform with the same functionality and capability in 2008 (Department of Public Service and Administration 2006; Weilbach & Byrne 2011). UbuntuLinux was implemented at SITA as the preferred operating system and this project was implemented in two phases, namely, front and back office. Phase one entailed the application of front office while phase two entailed the implementation of a centralised FOSS email solution, centralisation of user data storage and the migration of current directory services to FOSS directory services. Finally the model implemented was replicated to other Government Departments based on their FOSS requirements (Department of Public Service and Administration 2006; Weilbach & Byrne 2011). During recent visits to SITA and some South African government departments, the researcher determined that UbuntuLinux is still being used in SITA and these government departments as at 2014.

The Presidential National Commission on Information Society and Development (PNC) carried out a feasibility survey on migration to Open Source Software in 2006 (PNC 2007). PNC used the services of a vendor namely, Impi Linux (Pty) Ltd. to perform the migration of their IT systems to FOSS. The project was undertaken by following the standard SDLC phases namely, Initiate; Analyse; Design and Build; Implement and Close (Brown *et al.* 2014; Shelly & Rosenblatt 2013; Rob 2015). The project was a pilot one aimed to deploy open office on Windows and to evaluate the compatibility with Windows users in the Department of Communications. This led to the total migration to Linux following the results of the pilot project (PNC 2007).

The National Libraries of South Africa (NLSA) migrated 725 desktop computers and five blade servers to a Novell SUSE Linux Enterprise Desktop (SLED) environment in 2007. They saved R5.5 million in software licensing as well as improved security and user productivity. A vendor, MESO ICT Solutions, performed the migration for NLSA in two different phases, Planning and Deployment (Novell Connection 2009).

The National Department of Arts and Culture performed an in-house migration from a proprietary system to FOSS in 2007 (S. Phala, *pers. comm.*). Microsoft Exchange was migrated to Kolab, an OSS groupware mail server. The following FOSS products were implemented: Open Suze for desktop; Open Office for productivity; Fire Fox for web browsing; Drupal for intranet site and GLPi for the help desk.

The National Traffic Information System is a register supporting the National Road Traffic Act, 1996 (Act No. 93 of 1996) and the NaTIS/eNaTIS regulations. A vendor, TASIMA, performed the migration from a dispersed database architecture to an Integrated traffic systems using free and Open Source Software in 2009 (IT Web 2007).

1.1.2 Statement of Problem

Some of the hardest challenges that security researchers and professionals are faced with today include the prevention, detection, and responding to data leakage by authorised users, or insider threats (Huth *et al.* 2013). Thus, information in organisations has to be protected in accordance with how sensitive, critical and valuable it is. However, it should not depend on the storage media, the processing manual or automated systems, or the methods of the information distribution.

The protection of information in accordance with its sensitivity is substantiated by section 5 of the ISO17799 standard which stipulates that information should be classified according to its actual value and level of sensitivity so that the appropriate level of security can be deployed. ISO17799 Newsletter Issue 9 (2007) maintains that a system of classification should ideally be easy to comprehend and to manage; can be used to define the level of protection the information is given; and utilised uniformly throughout the whole organisation. Organisations need to protect the confidentiality and privacy of their sensitive information using document security technologies (Deshmukh & Pande 2014).

The view of Thompson and Kaarst-Brown (2005) is that there is a key task for intelligence and security informatics when discovering how sensitive information is categorised and classified by humans. Because of this, they point out that there is a need for research that discovers ways that sensitive information is different from other organisational information and also to understand how people conceptualise sensitive information with a view to discover which factors inspire their assessment of the sensitivity level.

A definition of sensitive information will assist organisations to understand how sensitive information differs from other organisational information and also how people should conceptualise sensitive information. Therefore, a preliminary definition of sensitive information is given below.

A preliminary definition of sensitive information is:

Sensitive information is information that can cause harm to an organisation or a person when it is revealed (Nawafleh *et al.* 2013).

Based on a comprehensive review of articles in the literature, there is currently no management framework to manage information sensitivity during migration of platforms in academic research. Thompson and Kaarst-Brown (2005) also contend that there is no research conducted on sensitive information formal classification schemes, despite the call from the US federal government on the need for research on sensitive information classification. Furthermore, they highlight that the need for research in information sensitivity is due to some security solutions that need organisations to classify their information based on the levels of their sensitive information. This need also requires organisations to categorise sensitive information in their communications network (Liddy 2001; Rakers 2010). There is the need to classify information assets and organisational data in terms of the risk of unauthorised disclosure (Rodgers 2012). Identity information such as user profiles and financial data should be protected when migrating between systems (Park *et al.* 2011; Pearson *et al.* 2007). They also highlight the importance of resolving the problem of migrating sensitive information between systems in dynamic environments (e.g. data centres) and they propose a policy-based approach to control and secure transfer of sensitive data across platforms.

As part of reported unauthorised access, the media reported that hackers have gained access into the Electronic National Traffic Information System (eNaTIS) (ITWeb 2007). According to this source, two individuals were arrested for allegedly trying to bypass the eNaTIS system to provide illegal roadworthy certificates for a number of vehicles. This raised fears that the eNaTIS is not highly secured and the security of the system needs to be investigated and enhanced.

The South African government policy on free and Open Source Software states that current proprietary software should be migrated to Free Open Source Software (FOSS) whenever comparable software exists (Department of Public Service and Administration 2006). The SITA FOSS Programme Office conducted a survey of all South African National Departments to assess and review the implementation of FOSS in their departments. The results of the survey show that 51% of all national departments have FOSS implementation strategies, 15% are using Z-linux based mainframe systems, 22% use FOSS web servers, while 41% use Linux and/or other FOSS at their back end, 7% use FOSS web standards and 12% are found to implement FOSS on their desktops (SITA FOSSFocus 2009).

SITA used some FOSS components like the Java platform, Eclipse, Trinidad, Glassfish, Liferay, Maven, EJB3, Subversion and Hibernate to build the Integrated Financial Management System (IFMS) which provides financial management, supply chain management, human resources management and business intelligence functions to South African national and provincial departments (SITA FOSSFocus

2009). Data was migrated from the old system to the new IFMS. The problem is, however, that SITA does not have a management framework that can be used to ensure that sensitive information is protected during such migration processes. This was revealed during the author’s discussion with some ICT staff working at SITA in 2009 (A. Webb, *pers. comm.*). Based on a comprehensive review of articles in the literature, SITA does not have a management framework that can be used to ensure the protection of sensitive information during software migrations as at 2014.

Government information consist mostly of sensitive information and there is a risk when SITA migrates the government desktop platform from a proprietary platform to open source. However, the protection of such sensitive information was not taken into consideration during the migration (SITA FOSSFocus 2009). The researcher argues that such a risk can be mitigated by applying a management framework that can assist in the protection and handling of sensitive information during the migration.

There are three main interrelated concepts in this research: firstly, it is the development of the management framework, secondly, it is the transfer of sensitive information between platforms e.g. from a proprietary platform to an open source platform and thirdly, it is the migration of platforms, for example from proprietary to open source. These three related concepts are illustrated in Figure 1.1 below.

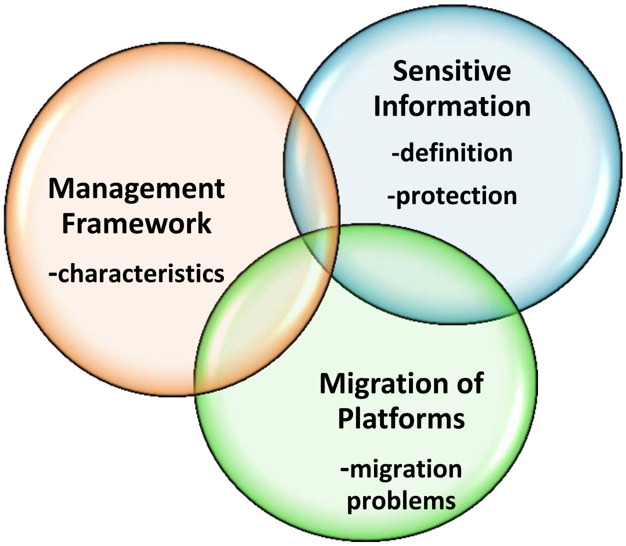


Figure 1-1: Relationship between a Management Framework, Sensitive information and Migration of platforms

The following problems are envisaged during the migration of sensitive information across platforms:

(a) There is the possibility of intruders trying to gain unathourised access into the system during such migration process (Crossler *et al.* 2013; Juneja 2013),

(b) Viruses and intruders can also invade the system during the migration process (Huth *et al.* 2013),

(c) Data integrity needs to be maintained during the migration and data corruption has to be prevented (Huth *et al.* 2013; Chavhan *et al.* 2013),

(e) Information leakage (Ahmad *et al.* 2014; Garfinkel 2014),

(f) Information theft (Von Solms & Van Niekerk 2013; Deshmukh & Pande 2014),

(g) Identity theft (Kirda & Kruegel 2005; Park *et al.* 2011),

(i) Phishing is a fraudulent online identity theft that is used to steal sensitive information e.g. passwords of banking clients and clients' credit card information (Kirda & Kruegel 2005; Park *et al.* 2011),

(j) Stealing sensitive information, e.g. account details and cookies, and getting hacked during the process (Gupta 2010; Juneja 2013).

Sensitive information residing in the databases to be migrated needs to be identified, classified and protected during the migration process. Such classification schema can then be used to facilitate sensitive information protection during similar migrations. This is done in order to ensure that the information is classified in accordance with their degree of sensitivity and that adequate protection measures are then applied to them as classified. There is a need for the protection of sensitive information during software migrations.

1.1.3 Research Aim and Objectives

The aim of this research is to develop a management framework for the migration of sensitive information during software migrations. This is a framework to manage the transformation as a process similar to the Software Development Life Cycle (SDLC).

The objectives of this research are to:

- conceptualise information sensitivity during platform migrations, e.g. from proprietary software to open source software.
- define the protection measures that should be undertaken during the migration from one platform to another one, e.g. from a proprietary platform to an open source platform.
- enrich the theory of information systems with respect to information sensitivity conceptualisation.
- develop a management framework that can be used to protect and handle sensitive information during migration of software platforms.

1.2 Research Questions

The primary research question is:

RQ: How should organisations manage sensitive information during its migration between software platforms?

Secondary research questions are:

SQ1: What are the differences between sensitive information and other information capital in an organisation?

SQ2: What are the protection mechanisms during the migration of information from one platform to another, e.g. from a proprietary platform to an open source platform?

SQ3: What would be the properties of a management framework for the migration of sensitive information during platform migrations?

SQ4: Why is the management framework necessary to protect sensitive data during software migrations?

1.3 Delineations of the Research

The delineations of this research are:

- The focus of this thesis is the development of a management framework to manage information sensitivity during software migrations. Data is collected from the following organisations namely State Information Technology Agency (SITA); South African Revenue Services (SARS); Presidential National Commission (PNC); National Libraries of South Africa, South African Department of Arts and Culture; South African Department of Public Works and South African Department of Social

Development. These organisations have performed platform migrations such as migration from a proprietary platform to an OSS platform. The data is then subjected to quantitative and qualitative analysis to obtain the final management framework.

- This research is not the traditional information security research, rather it focuses on the management area of information security.

1.4 Value of the Research (Rationale)

This research contributes to defining sensitive information, understanding how information can be classified, identifying specific migration problems, and developing a framework to manage sensitive information during software migrations.

The resulting management framework can be used to protect sensitive information between software migrations. Additionally, the research work contributes to the ICT theory by developing and validating the management framework on migration of platforms.

1.5 Research Design and Methodology

1.5.1 Underlying Philosophical Paradigm

Research strategies in Information Systems (IS) differ in their underlying philosophical paradigms and IS researchers are expected to understand different paradigms underlying their research strategies (Oates 2006). IS philosophical paradigms include positivism, interpretivism, critical research and pragmatism (Oates 2006).

The underlying philosophical assumptions in this research work utilises the pragmatism philosophical paradigm as explained in section 4.2.4. Pragmatism offers a general approach to the philosophical challenges facing the mixed methods research (Morgan 2014; Teddlie & Tashakkori 2009; Yardley & Bishop 2008). The case study methodology is used to carry this research by using multiple cases (data triangulation). Methodology is a strategy of enquiry guiding a set of procedures while methods are techniques used in analysing data to create knowledge (Denzin & Lincoln 2000; Cresswell 2009; Petty *et al.* 2012).

1.5.2 Case Study Research

Case study research is one of the ways of conducting social science research while experiments, surveys, histories and the analysis of archival information are the others (Yin 2009). Case study research is conducted in an actual life situation by the researcher and there is no distinction between the research phenomenon and the real life context, especially when there is no difference between phenomenon and context (Yin 2009).

The case study research is used as the methodology in this research work, and it is carried out by using the mixed method approach. Multiple sources of evidence (data triangulation) as explained by Yin (2003) is followed to conduct this research. The results from these cases are analysed using both quantitative and qualitative data analysis to develop the management framework on information sensitivity during the migration of platforms. The underlying philosophical paradigm used by the researcher is pragmatism, which substantiates the trustworthiness and dependability of the case study research. The case study research is conducted in some South African government departments and parastatals that have performed platform migrations as explained in section 1.1.1.

1.5.3 Mixed Methods Research

Mixed methods research has been defined by Johnson and Onwuegbuzie (2004) as an approach requiring the researcher to combine the two paradigms (quantitative and qualitative), methods, concepts or language. They argue that a mixed methods approach gains from the views and strengths of each method by acknowledging the importance and presence of truth and impact of human experience. Mixed methods research is defined by Tashakkori and Creswell (2007) as the collection and analysis of data and then integrating the findings by drawing inferences from quantitative and qualitative approaches.

Mixed methods research is used in this work to enhance and validate the management framework on information sensitivity as illustrated in section 4.3.4 and is elaborated upon in Chapter 4.

1.5.4 Data Gathering

Data was gathered in the government organisations and agencies that are mentioned in section 1.1.1. Data triangulation was used to collect the data, that is, data was collected from many different sources following Yin's (2003) data triangulation methodology. A questionnaire was developed and forwarded to 250 respondents in various government organisations and agencies. The author of this thesis received 90 completed questionnaires. The responses were then collated using a spreadsheet and the data was imported into the JMP SAS software (SAS 2014) for data analysis and more discussion on quantitative analysis follows in Chapter 5.

The quantitative research questions in Chapter 5 were enhanced by the qualitative analysis in Chapter 6 by using open-ended and in-depth interviews to validate the preliminary management framework that resulted from the quantitative analysis. The qualitative interviews were recorded on tapes and were later transcribed. Recording requires consent, hence ethical clearance was therefore obtained from the University of South Africa's ethics committee. A letter obtained from the UNISA ethics committee is shown in Appendix D(C). The transcripts were subsequently imported into the NVIVO version 10 software (Buchanan & Jones 2010; Edhlund & McDougall 2013) for further qualitative analysis. A more comprehensive description of the data gathering is explored in section 4.5.

1.5.5 Data Analysis

Two types of data analysis were performed, namely quantitative data analysis and qualitative data analysis in order to validate the management framework. A pilot (item analysis) was undertaken to test the reliability of the questions asked in the questionnaire and explained in Chapter 4. During this pilot quantitative data analysis, the questionnaire was validated by testing the reliability of the constructs in the questionnaire using item analysis (Cronbach Alpha) (Cronbach 1951; Cronbach & Meehl 1955).

Twenty-five respondents completed the first version of the questionnaire and then the data was analysed using statstical techniques to validate the constructs and obtain the final questionnaire. The final questionnaire was analysed using statistical analyses, namely exploratory factor analysis, item analysis, reliability analysis and Spearman's correlation analysis which are elaborated on in Chapter 5. Exploratory factor analysis is used to discover the constructs in the measuring instrument while item analysis was used to check the reliability of the constructs in a measuring instrument (Tate 2003; Wiid & Diggines 2013). After the pilot quantitative data analysis, discussed in Chapter 4, the descriptive and Spearman's correlation analyses were performed and elaborated on in Chapter 5.

During the qualitative data analysis, the audio tapes containing the interviews were transcribed and analysed using the NVIVO version 10 software. A bottom-up approach (content analysis) grounded in data was used to develop the management framework on information sensitivity abductively. The framework was validated using open-ended and in-depth interviews in the government organisations that have performed platform migrations. The detailed data analysis of this study is presented in Chapters 5 and 6.

1.6 Contributions of this research

The contributions of this research are:

- Development of a management framework that can be used to protect and handle sensitive information during migration of software platforms.
- Development of two frameworks (Rudimentary Management Framework and Preliminary Management Framework) and a model (Model on Security Challenges).
- Development of a formal (mathematical) description of sensitive information.
- Two international conference papers from the literature review section of the thesis.
- One international journal paper from the main work of the research.
- Contributions to the reduction in the research gap between academia and the industry.
- Conceptualisation of information sensitivity during platform migrations.
- Defining the protection measures that should be undertaken during platform migrations.
- Enriching the theory of information systems with respect to information sensitivity conceptualisation.

These contributions are further elaborated upon in the paragraphs that follow below.

The key contribution of this research to knowledge is the development of a management framework that can be used to protect and handle sensitive information during migration of software platforms. Based on a comprehensive review of articles in the literature, the researcher has not encountered any management framework that can be used to protect and handle sensitive information during migration of software platforms. The researcher reviewed the literature and was unable to find anything pertaining to a management framework that can be used to protect and handle sensitive information during migration of software platforms.

Two frameworks and a model were developed in this research before the final management framework was developed. The first one was the development of a rudimentary management framework. This framework was developed from the review of the literature by surveying for what should be the properties of a management framework that can be used to protect and handle sensitive information during migration of software platforms. The second one was the development of a model on the security challenges during OSS migrations. This model was also developed from the review of literature. The last one is the development of a preliminary management framework from quantitative analysis. This preliminary management framework was developed as a result of performing quantitative analysis while ensuring that the validity and the reliability of the constructs were maintained. A formal description of sensitive information was developed in this thesis.

Some of the parts of the research had been published in two international conference proceedings namely: The 2nd International Conference on Design and

Modelling in Science, Education and Technology (DeMset), Orlando, Florida, USA (2012) (Ajigini *et al.* 2012); and The 9th International Conference for Internet Technology and Secured Transactions (ICITST), London, UK (2014) (Ajigini *et al.* 2014). The third journal article had been published by the African Journal of Information Systems (AJIS) (Ajigini *et al.* 2016). This third journal article is the main work of the thesis that had been published by the AJIS. These articles are listed in the Publication's section of the thesis.

The research also contributes to the reduction in the research gap between academia and the industry. There have been calls for increased collaboration between the industry and academia as far back as the 1980s (Tartari & Breschi 2012). The outcome of the research may well assist the industry in information security management. Therefore, it may be concluded that the study can be regarded as a valuable and original one within the context of IS research (cf. Hassan 2007).

According to Hassan (2007), a valuable research in IS is one that contributes to the IS field itself and not to its parent disciplines or other fields e.g. management or computer science. He concludes that original research is a surbodinate to its discursive formation. Original IS research should focus on searching for concepts and theories within the IS discipline itself and not from other disciplines. He stresses that valuable research must be sensitive to the demands of changes in the academic and discursive environment. IS research must also uncover hidden insights and expose relationships that are silent in order to be valuable, relevant and original (Hassan 2007).

Other contributions of this research include the conceptualisation of information sensitivity during platform migrations; defining the protection measures that should be undertaken during platform migrations and enriching the theory of information systems with respect to information sensitivity conceptualisation.

1.7 From Rudimentary Management Framework to Final Management Framework

The movement from the rudimentary management framework to the final management framework is depicted in Figure 1-2. The rudimentary management framework is the first framework to be developed as illustrated in Chapter 2 (Figure 2-2). This is followed by the model on security challenges during OSS migration which is depicted in Chapter 3 (Figure 3-1). Thirdly, the preliminary management framework is developed and shown in Chapter 5 (Figure 5-27). Lastly, the final management framework is developed in Chapter 6 (Figure 6-5).

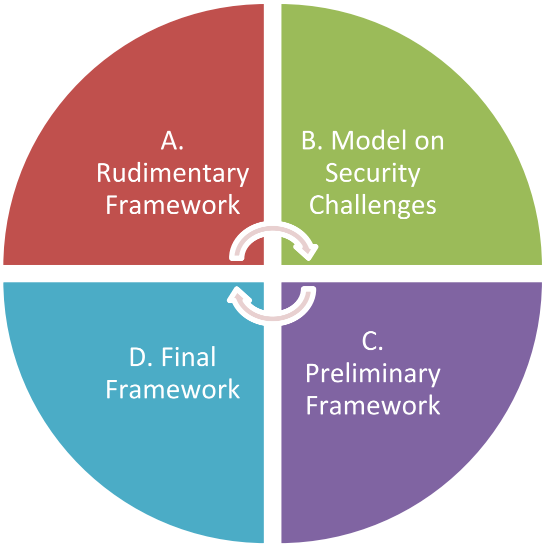


Figure1-2: The Movement from the Rudimentary Framework to the Final Framework

1.8 Chapter Layout of the Thesis

Chapter One: Introduction

This chapter highlights the statement of the problem investigated and the purpose of the study, research goals and objectives, research questions, delineations and limitations. It also briefly outlines the research design and methodology of the research.

Chapter Two: The Rudimentary Management Framework

The literature review on the rudimentary framework is explored (and a publication[1] resulted from this literature review). The rudimentary framework on information sensitivity developed from the literature review is also presented.

Chapter Three: Security Challenges During OSS Migrations

The literature review on the security challenges during OSS migrations is outlined. The model on the security challenges during OSS migrations developed from the literature review is presented (and a publication[2] emanated from this literature review).

[1] Ajigini, O. A., Van der Poll, J. A. & Kroeze, J. H., 2012, 'Towards a Management Framework to protect Sensitive Information during Migrations', *The 2nd International Conference on Design and Modelling in Science, Education and Technology (DeMSet),* Orlando, Florida, USA, (ISBN-13: 978-1-936338-76-4) ISBN-13: 978-1-936338-76-4 CD / ISBN-13, 6-13.

[2] Ajigini, O. A., Van Der Poll, J. A. & Kroeze, J. H., 2014. 'Towards a Model on Security Challenges during Closed Source Software to OSS Migrations.' *The 9th International Conference for Internet Technology and Secured Transactions (ICITST Proceedings)*, London, UK, 8 – 10 Dec., 275-284. ISBN: 978-1-908320-31-5.

Chapter Four: Research Design and Methodology

The first part of this chapter highlights an overview of research paradigms, research methods, and research methodologies and further explains why pragmatism, case study and mixed methods are used in the research. The activities conducted during the data gathering process and data analysis are elaborated upon.

Chapter Five: Preliminary Management Framework: Quantitative Data Analysis

The quantitative data analysis using statistical methods is explained and a preliminary version of the management framework on information sensitivity during platform migrations is developed from the analysis. The JMP version 11 software was used to perform the statistical analysis.

Chapter Six: Final Management Framework: Qualitative Data Analysis

The qualitative data analysis using NVIVO version 10 to perform content analysis in order to refine and validate the preliminary management framework obtained in Chapter Five is explored. The analysis also includes the findings from the qualitative interviews. This process results in developing the final management framework which is further illustrated in this chapter.

Chapter Seven: Conclusions and Recommendations for Future Work

In this chapter, the conclusions and recommendations for future work are explained. This includes the synopsis of the research questions and the individual chapters. A journal article[3] resulted from the main work of this thesis.

[3] Ajigini, O. A., Van der Poll, J. A. & Kroeze, J. H., 2016. 'A framework to Manage Sensitive Information during its Migration between Software Platforms.' *The African Journal of Information Systems*, Vol. 8, Issue 2, Article 2, pp. 21 – 44. ISSN 1936-0282.

The thesis concludes with the References and the Appendixes (Appendix A to Appendix E).

1.9 Conclusion

This chapter has explored the purpose of the research and the problem investigated, including the goals and objectives of the research. The research aims to develop a management framework to manage sensitive information between software migrations. The primary and secondary research questions are provided with a view to resolve statement of the problem. Moreover, the chapter includes a brief description of the research design and methodology by acknowledging the use of the case study methodology and the mixed methods approach.

The next chapter is about the literature review, including the development of a rudimentary management framework that resulted from the contextualisation of the literature.

Chapter 2

The Rudimentary Management Framework

2.1 Introduction

In the previous chapter, the researcher outlined the purpose of the work and the problem statement including the aim and objectives of the research. The researcher's focus in this chapter is the critical reviewing of the literature on the main aspects of the research. Sensitive information is defined from a synthesis of definitions found in the literature. The protection and management of sensitive information are explored in sections 2.3 and 2.4.

Open Source Software (OSS) and Closed Source Software (CSS) are defined. The various OSS initiatives undertaken by the South African Government and other Foreign Government initiatives are explored. Section 2.8 explores the overview of platform migrations in two different sub-sections – general IS migrations and OSS migrations. Some of the popular OSS projects are addressed in section 2.9, and section 2.10 highlights the benefits of OSS versus CSS security. The security of OSS features are addressed in section 2.12 while section 2.15 concludes the chapter.

The literature review of this chapter as well as the next chapter is tailored towards addressing the five research questions stated in section 1.2 and also includes the foundational concepts of the management framework on information sensitivity. This research is being embarked upon in order to maintain confidentiality, integrity and availability of sensitive information between software migrations. This is why the researcher develops a management framework to manage such sensitive information between software migrations in this thesis. The content of this chapter was synthesised into a research publication (Ajigini *et al.* 2012).

2.2 What is Sensitive Information?

Some authors have defined sensitive information in the literature (Gennotte & Trueman 1996; McCullagh 2007; Thompson & Kaarst-Brown 2005; TJNAF 2007) to name but a few. Table 2-1 depicts the definitions of sensitive information from different authors.

TABLE 2-1

Definitions of Sensitive Information

Authors	Definitions of sensitive information by each author
Gennotte and Trueman (1996)	Protected information used to increase the prospect of the result for the organisation, group, or person handling the information.
ALRC (2000)	Information pertaining to a person's race or ethnicity, political orientation, religious relationship, philosophical inclination, profession, trade union or association, sexual orientation or criminal practices.
Thompson and Kaarst-Brown (2005)	Information about the owner that is concealed by the owner. It is also information known to a person about an organisation that the person does not want to reveal outside the organisation.
TJNAF(2007)	Information that can cause damage to the government, laboratory or persons if such information is made known to people that do not require knowledge of such information in the discharge of their duties.
McCullagh (2007)	The European Union defines sensitive data as information that exposes the ethnicity, political orientation, religious or philosophical affliation, health, sexual beliefs, and membership of trade union.

NIST (2008)	Sensitive information is defined as any information that, if lost, misused, modified by unauthorised persons, will result in undermining the national interest, federal programmes performance, individual privacy entitlement as enshrined in the Privacy Act, that is not approved by an Executive Order or Congress Act and which is expected to be hidden in line with national defence interest or foreign policy. According to the US Computer Security Act of 1987, agencies are required to categorise and distinguish their sensitive systems, train their employees in computer security and create computer security plans.
NIH (2008)	Sensitive information is when the loss of confidentiality, availability, or integrity of such information could have a disastrous unfavourable effect on individuals as well as organisational belongings.
Nawafleh *et al*. (2013)	Sensitive data is any information which, if leaked, can lead to the destruction of the person or the organisation and may include personal information as well as the organisation's information.

In Table 2-1, some authors (Nawafleh *et al.* 2013; NIH 2008; NIST 2008) define sensitive data from the point of losing information that can lead to destruction of a person or an organisation which can be disastrous to them. Other authors (Gennotte & Trueman 1996; Thompson & Kaarst-Brown 2005; TJNAF (2007) define sensitive data as protected information that should be concealed and not revealed to other people. Therefore it can be inferred from Table 2-1 that sensitive information is protected information that should not be lost, ortherwise such loss is detrimental to the person or the organisation and efforts should be made to conceal such information and prevent it from being revealed to other people.

The concepts in Table 2-1 are synthesised into the following definition of sensitive information specified as definition 2.1:-

Definition 2.1:

Sensitive information is protected information that the owner does not want to reveal to others and which is not to be divulged outside the organisation, as well as information concerning an individual's ethnic origin or race, criminal record, sexual preferences or practices, and other information that include political beliefs, political association membership, trade union membership, religious associations or philosophical opinions; efforts should be made to conceal such information not being revealed to other people.

Examples of sensitive information include (Jericho Forum 2009; Nawafleh *et al.* 2013):

- Identity number
- Patient's personal information
- Student's personal information
- Organisational financial data
- Students' records (e.g. marks, study plans)
- Employees personal information
- University research data
- Credit card number
- Bank pin number
- University special legal data

This view of sensitive information is incorporated into the management framework on information sensitivity during software migrations. This can assist in differentiating between sensitive information and non-sensitive information during the software migration process so that application protection measures can be applied to the sensitive information. The researcher will work on all these aspects of sensitive information in this study.

The above descriptions of sensitive information lean towards instantiations, i.e. stating examples of sensitive information, yet the researcher embarks on a formal approach to defining sensitive information in section 2.5.1.

2.3 Management of Sensitive Information

Sensitive information from US government networks is being gathered by well-funded Chinese groups (Graham 2005). This has led to national security concerns in the USA and the extent of these intrusions and the nature of data exposed is not fully known (Casey 2006). There is the pressure for organisations to enforce good corporate governance, secure sensitive information and comply with standards (Fakhri *et al.* 2015).

A new challenge for management is to keep the vast amounts of information contained within computer-based information systems secure in organisations (Taylor 2006). Information systems store private or confidential information, e.g. identity numbers, employee salaries and patient's records, that need to be secured properly (Bhatt & Dongre 2014). This problem is exacerbated by the advent of 'big data'.

There is also a special sensitivity surrounding any personal health data and medical records. The role of sensitive information as well as trust in decreasing the concerns about privacy of medical information have been investigated by Rohm and Milne (2004) and they conclude that consumers are concerned about their medical history and records. Iroju and Ikono (2013) point out that some security violations taking place within the healthcare systems include: information theft; unauthorised view of patients' information; eavesdropping of patient information over a network; and unauthorised destruction of patients' data. Also, storing data in the 'cloud' has not yet countered all user fears of security, especially with regard to public cloud. In South Africa, legally, no financial instituition is allowed to store financial information outside the geographical borders of the country, and with the public cloud, it could be located anywhere, unknown to the owner of the data (FATF/OECD 2009).

The management of sensitive information related to their business ought to be very important to all organisations (Rakers 2010). Some authors (Nurse *et al.* 2015; Kale *et al.* 2015; Ahmad *et al.* 2014 ; Chavhan *et al.* 2013; Arai and Tanaka 2009) have highlighted the importance of avoiding information leakage for a computer system's handling of a company's sensitive information. Sensitive information should be encrypted and technology should make it possible to assign the decryption key between the users dealing with the sensitive information (Arai & Tanaka 2009). This process of encrypting and decrypting data forms part of the management framework on information sensitivity during software migrations.

The complexity of the information systems required for its safe-keeping increases as the amount of personal data stored and processed by companies increases (Naik & Ghule 2013; Acquisto *et al*. 2006). Internal employees of organisations might try to gain unauthorised access to the information while others might unintentionally put organisational information at risk (Sampemane 2015; Choi *et al.* 2014; Taylor 2006). This is why information security problems due to the integration of organisations into the World Wide Web draw considerable attention from investigators and experts (Ma & Pearson 2005b).

Sensitive information is kept by the social constraints defined by social connections that has to be protected since absent protection affects the organisational image. This is why the security of IT software and the network controls must be taken into consideration when designing and implementing new software systems (Scholz 1990). The information in a smartphone (a wireless network device) is more at risk than in a computer because of lack of security controls (Vargas *et al.* 2012). Network control is embedded as part of the components used to build up the management framework on information sensitivity.

2.4 Protection of Sensitive Information

Organisations need to to know who and when their information is accessed since information has become one of the most important resources for them (Vargas *et al.* 2012). Information is a resource that has strategic value to an organisation and exists in many forms like written or printed documents, electronic files, microfilms and videotapes (Fung & Jordan 2002). Information has been regarded by Duri *et al.* (2004) as the new currency of the global economy. Correct information is expected to support decision-making or to provide service at the appropriate time. Therefore the integrity of the information cannot be compromised and data protection is vital in order for the users to be assured of their privacy and that the data meets the service provider's integrity requirements (Duri *et al*. 2004; Chavhan *et al.* 2013). Business activities should be the first to be protected in any security program (Fakhri *et al.* 2015).

Corporations have been motivated to invest in information security by safeguarding their confidential data and their customers' personal information (Kalyvas *et al.* 2013; Acquisto *et al.* 2006). The non-protection of sensitive information can damage the reputation of an organisation (Kalyvas *et al.* 2013; Rasmussen 2008). Organisations must protect their sensitive information throughout its lifecycle. This has led Taylor (2006), in carrying out research using case studies and intergroup bias theory, to investigate the current strategies to protect organisational information. Taylor's study shows that organisational information is at risk due to employee behaviour which can be intentional or unintensional and management should educate employees on the different behaviours that can cause information security risks. Organisations need to take into cognisance the human aspect of information security and include both information that is within and outside their computer-based information systems as part of their information security definition (Taylor 2006).

There has been wide media coverage of many incidents involving the disclosure of sensitive information due to leakage in recent years (Ahmad *et al.* 2014). They stress that sensitive information leakage through unknown avenues is a serious problem to management, mostly caused by mobile devices, cloud computing, network technologies and social media. They maintain that due to the leakage, organisations may suffer from reputational damage, revenue loss, and loss of productivity. According to them, the leakage can be prevented by using technical measures to control information access, e.g. passwords, encryption, logging mechanisms, firewalls, and intrusion detection systems. Organisations ought to have a data protection strategy as part of a data leak prevention solution (Gupta 2010). From the recommendations above, the inclusion of strategy is in fact a part of the information management framework components developed by the researcher.

Data leakage includes different types of crimes perpetrated by insiders, theft of personally identifiable information, theft of intellectual property, an insider passing sensitive or classified information to an unauthorised third party (Huth *et al.* 2013). Data leakage and theft have been classified by McCormick (2008) into three stages: (a) obtaining access, (b) downloading data and (c) sharing data. Therefore, it is crucial to protect sensitive information during migration of software platforms in order to avoid leakage and theft of sensitive information. The people’s roles and responsibilities are part of the management framework developed by the researcher as well as their training and awareness of information security in their environment.

Database privacy has to be maintained (Olivier 2002). Olivier expresses the importance of database privacy and maintains that the challenge of database privacy is how personal information is enabled in databases in a way that balances society's needs with those of the individual. Server security should also be implemented because servers store or process sensitive information belonging to their organisations (Martinez *et al.* 2013). Database level encryption is considered when protecting data by using keys. Organisations use database-level access to control types of information that can be shared among users (Bayuk 2009). Organisations store sensitive information in databases, but database security has not been given the much attention it requires as other areas of information security (Sodhi 2015). Database technology is incorporated as part of the management framework components developed by the researcher on information sensitivity because of the above recommendations.

Proceeds from information theft were estimated at $105 billion US worldwide in 2004 (Swartz 2005). Similarly, the cost of electronic crimes was estimated by the FBI to be approximately US$10 billion a year (Ma & Pearson 2005b). Due to the high level of cyber crimes, the United States Congress passed a series of bills in November 2002 to allocate one billion dollars for research on cybersecurity with the aim of combating terrorist attacks on private and government computer systems (Ma & Pearson 2005a). Moreover, Diffie (2008) indicates that information security is a vast field that involves vasts amounts of money, publications and practitioners when compared to all computer science areas a half-century ago.

Information theft can also have non-financial implications. This is also the view of Bruce (2003) who points out that breaches of information systems can have non-financial implications like a negative impact on a company's status, trust, goodwill, deficit in potential sales and competitive advantage. Also, losing sensitive information by organisations may cause confidential information leakage that can cause financial loss (Sarrab & Bourdoucen 2013; Chavhan *et al.* 2013).

The cost component is included by the researcher as part of the management framework due to the above recommendations.

A majority of security incidents are caused by organisational employees that violate IS policies. Therefore, a proper working environment needs to be created to enhance employee compliance to organisations’ IS policies (PWC 2008; Whitman & Mattord 2008; Kolkowska 2011). This has called for the establishment of environments that ensure good security behaviour by transforming the culture of the organisation and setting up an information security culture (Knap *et al.* 2007; Thomson *et al.* 2006). As part of protecting sensitive information, Rasmussen (2008) maintains that the detail of how sensitive information is labelled, stored, distributed and destroyed must be contained in their data security policies. Security scholars have identified the lack of awareness of security policies among users to be a major cause of failure (Abraham 2011). Culture, policies and procedures form part of the components used to build the management framework.

2.5 Information Sensitivity and Information Classification

Data classification is the grouping of data into homogenenous groups (Kaushal *et al.* 2015). Data classification allows organisations to apply protective markings on documents and messages in both visual and metadata forms (Tankard & Pathways 2015). Such tools lead to the improvement of data security and also make employees to be aware of what constitutes sensitive data so that they can protect it. Data classification tools are used with data leakage prevention tools in order to prevent security breaches. Nawafleh *et al.* (2013) state that there is a gap in the way that organisations control, monitor and protect their business environment including their electronic data assets. The concept of information classification is mostly not implemented or totally absent in many organisations (Gupta 2010).

The need for research on sensitive information has been stressed by Thompson and Kaarst-Brown (2005) as well as Nawafleh *et al.* (2013). More research should be done to comprehend how sensitive information should be conceptualised and also to understand the difference between sensitive information and other organisational information. Thompson and Kaarst-Brown (2005) argue that because these research gaps hinge directly on the information, research to accomodate them should be done at the same time with research efforts that are linked to IS security architecture, such as systems that have multiple layers of security. The developments in the IT field have created the need to understand sensitivity cues (Thompson & Kaarst-Brown 2005).

Many authors explain the reasons for information sensitivity classification (The Open Group 2009; Bradley 2007; Chang *et al.* 2009; Nawafleh *et al.* 2013; Rodgers 2012). Organisations classify their information so that they can have control over who accesses their sensitive information or confidential information; protect their sensitive information or confidential information and make it easy to find their sensitive information (The Open Group 2009). A well-planned data classification system enables data to be easily retrieved and located (Kaushal *et al.* 2015). Information classification is used to build up the management framework on information sensitivity developed by the researcher.

Information classification is important during the information protection process and there are many different classification schemes available. The Open group (2009) came up with a four-level classification scheme called the ‘G8 Traffic Light Proposal’ and these are:

- Red: Highly sensitive
- Amber: Sensitive
- Green: Normal Business
- White: Public

The BS17799 Classification scheme has five levels namely (Jericho Forum 2009):

- Top secret
- Highly confidential
- Proprietary
- Internal use only
- Public documents

The above BS17799 Classification scheme is of particular significance in this thesis, since these levels are adopted by the researcher as the basis for a formal approach to defining sensitive information in section 2.5.1.

Nawafleh *et al.* (2013) propose a three-level classification scheme namely:

- Sensitive data classification
- Private classification
- Public classification

The Nawafleh *et al.* (2013) data classification scheme is similar to the one proposed by Rodgers (2012) in which the data can be classified into a three-level classification system: Public; Internal and Confidential.

The suggestion of the researcher is that organisations should use the BS 17799 classification scheme to classify their sensitive information during migrations of software platforms since it is an international standard that has been validated and widely approved. The formal description of sensitive information presented in section 2.5.1 builds on the five-level BS17799 Classification scheme.

The United States federal government has elaborated on the importance of research on how sensitive information should be classified and understood (Thompson &

Kaarst-Brown 2005). The US government also stresses the importance of conducting research to comprehend the classification of sensitive information and this is because the US classification of national security information is outdated and new research initiatives are needed in this area (Thompson & Kaarst-Brown 2005). This view is also supported by McCullagh (2007), who also adds that the current classification of sensitive data is outdated and ineffective for determining the conditions of data processing.

The report compiled by the US General Accounting Office (GAO 2000) has highlighted the need for categorising data used by all federal agencies. The terrorist attack on the World Trade Center in New York and the Pentagon in Washington DC on September 11, 2001, moved the intelligence committees of both houses of the US Congress to propose the review of the statuses, policies and procedures governing the classification of national security information (US Congress 2003). The Open Group (2009) states that the recent information classification systems are used by specialists and only a small portion of the information is labelled. Organisations should be able to classify information based on its sensitivity, taxonomy, probability and use the classification to be able to protect sensitive information in their organisations (The Open Group 2009). This will enable them to understand which information should be the most protected and which one could be the least protected.

Some authors (The Open Group 2009; Bataller 2012; Richardson & Michalski 2007; Fowler 2003) have also stressed the need to have a classification system for information in order to realise the goal of performing a sensitivity assessment. The view of Farrell (2002) is that organisations must perform sensitivity assessment even if they know the various protection requirements for information located in both their electronic and manual systems. However, such ideas on sensitive information protection can be guided by using a management framework on information sensitivity which is the intent of this study. Sensitivity assessment is incorporated in the management framework on information sensitivity because of the recommendations above.

Next a formal approach to defining sensitive information is presented using the above BS17799 Classification scheme.

2.5.1 Towards a Formal Description of Sensitive Information

The BS17799 Classification scheme defines five (5) levels of sensitive information, namely, Top secret, Highly confidential, Proprietary, Internal use only, and Public documents. Following this classification, the researcher views information in an organisation as potentially covering the whole spectrum of these classifications. It is also plausible that sensitive information in a financial institution is different from likewise information in the medical sphere or the defence force which often work with mission- or safety critical software. Subsequently, the space of all sensitive information is divided into spheres (Financial, Medical, etc.) as indicated in Figure 2-1. As per BS17799, such information is spread over the five (5) levels as shown.

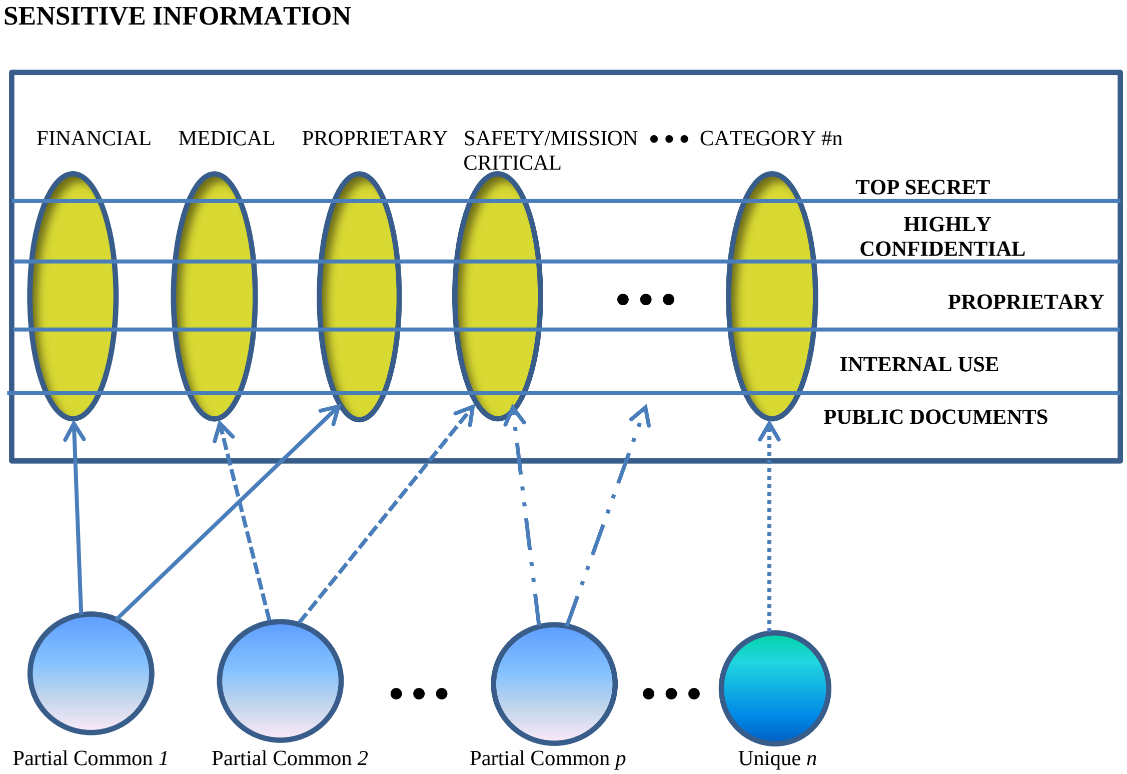


Figure 2-1: Formal Description of Sensitive Information

Some sub-spheres of sensitive information in a sector (or organisation) may overlap with similar information in one or more other sectors, hence the blue circles at the bottom of the figure. For example, in Figure 2-1 sensitive information in the Financial world may overlap with such information in a Proprietary sector. Sensitive information in safety- or mission critical organisations may overlap with a sector or grouping not shown in the diagram. Following standard mathematical set-theoretic notation, the universal set of all sensitive information is indicated by the outer rectangle.

Following Figure 2-1 and the above description, sensitive information *common* (*C*) to all sectors is defined as the arbitrary intersection (∩) of the sensitive information in each, i.e.:

<u>Definition 2.2:</u> **Common features (*C*) among all sectors (organisations)**

$$C = \bigcap_{i = 1}^{n} S_i$$

where,

Si ∈ {FINANCIAL, MEDICAL, PROPRIETARY, SAFETY/MISSION CRITICAL, …, CATEGORY #n}

<u>Definition 2.3:</u> **Partial Common features**

In a similar vein sensitive information that may be common to only certain sectors (e.g. sector k, C_k) but not all of them is defined as:

$$Partial\ C_k = \bigcap_{i = 1}^{2 \leq k < n} S_i$$

where as before,

Si ∈ {FINANCIAL, MEDICAL, PROPRIETARY, SAFETY/MISSION CRITICAL, …, CATEGORY #n}

Naturally, the above mathematical description of sensitive information may be developed further through appropriate qualitative investigations in the various sectors indicated in Figure 2-1.

Next the important topic of open-source software versus closed-source, or Proprietary software is discussed.

2.6 What is Open Source Software (OSS)?

Open source software is software that is distributed with its source code and open-source license which can be free or not free depending on the license under which the software is distributed (Stoyanov & Kordov 2014). Hansen *et al.* (2002) describe open source as software of which the source code is disseminated with the executable program. The software comes with a licence that allows users and developers to change and redeploy the software. OSS is software that is issued under OSS licensing formats (Gwebu & Jang 2011; Vintila 2010; Stoyanov & Kordov 2014). Pearson (2000) writes that open source is a term which is the opposite of closed source in the sense of having its source code freely available for anyone to make enhancements or correct errors.

A comprehensive definition of open source software is outlined by the Open Source Definition (OSD) and they list eight requirements of OSS as (OSI 2014):

(a) Integrity of the author's source code,

(b) Derived works,

(c) Source code availability,

(d) Free redistribution,

(e) No license restriction on other software,

(f) No discrimination against fields of endeavour,

(g) No discrimination against persons or groups,

(h) License must be neutral regarding the technology.

FOSS is described by Rafiq and Ameen (2009) as computer software that has its source code available under a license allowing users to use, enhance and modify the software and which can be redistributed in unmodified or modified form. The implementation of Free Open Source Software (FOSS) has been adopted by many governments globally (Mtsweni & Biermann 2010). The South African (SA) government has been at the forefront of advocating the use of FOSS (Johnston & Seymour 2005). Mtsweni and Biermann (2010) indicate that a number of governments implemented FOSS on their servers and workstations. Additionally, migrations from Proprietary Software to FOSS were performed worldwide. Therefore, it is important to protect sensitive information during such migrations in organisations.

The perpetual rise of OSS has been a feature of the software industry during the past ten years (Kemp 2009). Subsequently, the open source movement as a new paradigm for software development has gained increasing interest in recent years (Raghunathan *et al.* 2005). One such example is the South African government, which has been in the vanguard of using OSS since 2001 by adopting policy recommendations in 2003 (Miscione & Johnston 2010).

The gradual proliferation of articles and reports in the mainstream media presented evidence for the awareness to, and increased prominence of open source (The Guardian 2004). This is also the view of Hoepman and Jacobs (2007) that there are many publications on OSS advantages and disadvantages. The view of Gartner (2008) is that at least 80% of all commercially available software solutions would have had functional open source components by 2012.

A study conducted by the International Data Corporation (IDC) indicates that the OSS market will grow at an annual rate of 22.4% and might reach US$8.1 billion in 2013 (Little & Stergiades 2009). OSS software market is projected to cost US$46 billion (Statista 2015) which is far more than the forecast in 2013. Gwebu and Wang (2010) have pointed out that low acceptance rates of OSS continue to reduce its share of the market. In later work, Gwebu and Wang (2010) emphasise that OSS benefits would not be fully realized until it is accepted and used by the mainstream software users. However, in 2014, the OSS acceptance rates has shot up because 485 of the 500 supercomputers are running some form of Linux, 75%

of large companies use Linux in the Cloud as against 23% that use Windows, while Android is used in 83.6% smartphone shipments (Moody 2015).

There are many classifications of open source and closed source in the literature e.g. (Gwebu & Wang 2011; Hansen *et al*. 2002; Vintila 2010). In this section, open source software is viewed as software of which the source code is supplied alongside with the executable program, having the right of redistribution and open standards. It is free to use, but could be paid for (e.g. paying for the medium on which such software is distributed) as illustrated in Table 2-2. The researcher uses the migration from a proprietary software to an open source software to develop the management framework on information sensitivity, therefore the distinction between open source and closed source needs to be clarified.

In Table 2-2, our definition of open source covers column 1 (both quadrants A and C). Examples of closed source and open source programs are listed in each of the four quadrants. For example, Ubuntu Linux is a free, open source program in column 1, row 1 (quadrant A), while MS Office is a closed source software that is paid for and resides in column 2, row 2 (quadrant D). It should be noted that free software may be distributed on a medium such as a CD which is not free.

TABLE 2-2
Classification of Open and Closed Source Software

	Open source software	**Closed source software**
Free	**Quadrant A** e.g. Ubuntu Linux, Apache	**Quadrant B** e.g. Internet Explorer, Adobe Acrobat Reader
Paid For	**Quadrant C** e.g. Red Hat, MySQL	**Quadrant D** e.g. MS Office, MS Windows operating systems

2.6.1 History of OSS

Kemp (2009) writes that the US academia started the foundation of the OSS movement in the 1960s, when there was a cultural attitude of opposition to the restrictive nature of exclusive rights under intellectual property laws. Richard Stallman, an ex-MIT academic, launched the Free Software Foundation (FSF) in 1985, which is dedicated to the development of free software as a non-profit body (Kemp 2009).

Richard Stallman is considered to be the founder of the open source movement because he holds very strong philosophical beliefs that all users of computers should have the freedom to enhance any software in order to share software and also to support their needs (Pearson 2000). He started to write the GNU software (an acronym for ‘GNU’s Not Unix). He also developed a licencing system for GNU software called ‘copyleft’. The FSF was set up to further the development of the GNU software (Pearson 2000).

The FSF freedom philosophy is openly anti-business and this concept was promoted to a wider business community in 1997 by a group of free software community leaders (Pearson 2000). This group came up with the name ‘open source’ and they came up with a definition to provide the requirements for open source software. This is the group that created the Open Source Institute (OSI) that manages the Open Source Definition (OSD).

According to Kemp (2009), the FSF oversaw the GNU project, which was a mass collaboration, to create a free full operating system that will replace the UNIX system under the GPL – the GNU General Public Licence (GPL). In 1992, the operating system kernel (known as GNU Hurd) had not been completed, though all the other necessary components had been completed. The GNU software was combined with Linux in 1992 (a new kernel) to have a complete operating system, a combination known as GNU/Linux and licenced under the GPL.

Linus Torvalds, a 21-year-old Finn and a computer scientist at Helsinki University, developed Linux, which was firstly publicly released in September 1991 (Pearson 2000; Kemp 2009; Nazeer *et al.* 2015). The name ‘Linux’ is derived from his first name and ‘UNIX’. Linux is freely available over the Internet and suggestions for enhancing the system are requested from the public. Linux is being used and adopted commercially by many computer manufacturers and it is a competitor for the Microsoft Windows operating system – a closed system (Nazeer *et al.* 2015; Pearson 2000).

According to Kemp (2009), the Open Source Initiative (OSI) was established by Eric Raymond and Bruce Perens in 1998 to promote OSS on pragmatic grounds. Part of the function of the OSI is to review and approve licences conforming to the Open Source Definition (OSD) which was carved out from the OSI (Open Source Initiative). The OSD has eight requirements that must be adhered to before software can be allowed to be classified as open source as explained early in this section.

2.7 What is Closed Source Software (CSS)?

Closed source software (CSS) is a term invented as an antonym for OSS and is used to refer to any program whose licensing terms do not qualify as OSS. This implies that a user will have the binary version of the software they are licensed to without any copy of the program’s source code. A user of closed source software might not be able to render modifications to the software, although in certain situations, it can be de-compiled or reverse engineered.

CSS is based on the assumption that software development is managed by a team of specialised developers and best practices project management since it is a highly specialised process, all of which result in new releases and enhancements from time to time (Raghunathan *et al.* 2005).

In this section, closed source software (CSS) is defined as software to which users might not be able to render modifications and can be free or paid for as illustrated in Table 2-2. An example of closed source software that is free as shown in Table 2-2 is Internet Explorer which can be downloaded from the Internet.

2.8 Overview of Platform Migrations

Many organisations migrate their legacy systems to modern systems as a result of mergers and reorganisations (Razavian & Lago 2014) and they have to address demands for high quality, fast delivery and decreasing costs during such migrations. Pearson *et al.* (2007) have highlighted that the problem of migrating sensitive information between systems in dynamic environments is crucial as distributed computing expands and this information has to be secured and protected. Users have been considered to be the most problematic aspect of migration because user skill and discomfort are not easily quantified (Drozdik & Kovacs 2005). This section investigates the following kinds of migration in the Literature: (a) General IS Migrations and (b) OSS Migrations.

2.8.1 General IS Migrations

IT systems migration involves moving from the source system state to the target system state (Pieta 2010) and it has potentially undesirable effects from a security point of view. Kazimir (2012) indicates that change is the outcome of migration resulting from organisational mergers, acquisitions, or business optimisation, for instance. This type of change includes:

(a) Developmental change – enhancement of business applications and IT infrastructure;
(b) Transitional change – from an initial state to the target state;
(c) Transformational change – demise of the old state due to natural disasters, collapse of IT;
(d) Infrastructure and then moving to the new state.

IT migration has been defined as a type of interim change that occurs repeatedly (Kazimir 2012) which is aimed to improve the organisational function and enhance business by transitioning from an old state to a new state. Therefore, the migration process has to be planned and implemented using best practices in terms of management, maintenance, support, and IT models, e.g. ITIL, CobiT or CMM/CMMi (Pieta 2010).

Different types of IT migration include:

(a) Application migration – the process of redeploying applications to newer platforms and infrastructure (Kazimir 2012; Torchiano *et al.* 2011), therefore, the SLAs (Service Level Agreements), data portability, long-term costs, user management and security should be considered before the migration;
(b) Business process migration – migration from the organisation's 'as is' state to the 'to be' state and it can involve moving data from storage devices, databases or applications to others depending on the changes required;
(c) Data migration – this is regarded as the basic IT migration because information in IT systems is stored in form of data. It involves transferring data using ETL (extract, transfer & load) processes between storage types, formats, database applications of IT systems (Kazimir 2012);

(d) Data centre migration – a relocation of servers to another location in order to meet the needs of the organisation in terms of its software and databases by taking into consideration energy efficiency and higher computing density units (EMC Corporation 2011);

(e) Database migration – the transferring of data from a different operating system platform or from one database to new hardware;

(f) End-user equipment migration – this involves transferring the user environment between two computer systems and it involves the movement of workstations to end-user workplaces, migration of their data and profiles, among others;

(g) Server migration – the migration to a different physical server, the server instance and its services as well as data and it involves the migration of data, features, server roles, and operating system settings;

(h) Storage migration – this is the migration of data between storage systems;

(i) System migration – the transfer of data, programs and settings from the old IT system to the new one and it involves migrating system, user, email, network settings and data between the systems;

(j) Website migration – transferring website files from one web hosting organisation to the other. This involves the transfer of the web site HTML files and images, scripts, or applications, the web site media files, MySQL databases and e-mail configuration replication.

Generally, IT migrations involve the movement of application, system, data, process servers, and storage from the old system to the new platform and it should be executed by following proper actions and processes (Kazimir 2012). This research focuses mainly on the application and data migration as explained in this section. The information management framework developed by the researcher is based on the application and data migration.

2.8.2 OSS Migrations

Oram (2011) highlights the various successful OSS migrations and these include: (a) Migration from Microsoft Office to OpenOfice.org and also from the Windows Operating System (OS) to Debian GNU/Linux by Munich; (b) Migration to OSS by the Brazillian stratum of educated professionals; (c) Migration to OSS by OSS advocates and civil society organisations and (d) Migration from Microsoft Office to OpenOffice.org in mid-2000 by Massachusetts State Government, USA.

The factors governing the sustainability of OSS Migrations have been investigated by James and Van Belle (2008), using a qualitative and thematic analysis approach. The work of the Shuttleworth Foundation has increased the knowledge and importance of OSS in South Africa in recent years (James & Ven Belle 2008). Hislop (2004) also pointed out that several South African municipalities have migrated to OSS with various levels of success. SITA migrated all SITA users from the proprietary environment to the FOSS platform with the same functionality and capability in 2008 (Department of Public Service and Administration 2006; Weilbach & Byrne 2011). The Presidential National Commission on Information Society and Development (PNC) also migrated users from proprietary environment to Linux in 2006 (PNC 2007).

A survey about OSS usage at universities and research centres indicates that 60% of servers, 42% of database systems, 67% of email systems and 87% of tools for managing contents of universities are based on OSS (CENATIC 2009). Research on characterising OSS migration initiatives has been performed by Heredero *et al.* (2010). They found that software migrations from proprietary to open source depend on organisational and contextual factors such as the IT resources accessibility, organisational climate, organisational complexity, political support, why the change is needed and the project leadership style. Organisational factors are part of the management framework developed by the researcher.

An overview of OSS migration and criteria for migration challenges has been presented by Geetha (2012). He points out that organisations migrate to OSS from legacy systems because the legacy systems are difficult to integrate with the newer technologies. The OSS migrations can include:

- Language or code migrations;
- Operating systems migrations;
- Data migrations;
- User interface migrations;
- Architecture migrations.

All the above-mentioned migration types are part of the management framework developed by the researcher. A description of these migration types follows: Language or code migration is the process of moving software between languages (Nguyen *et al.* 2014). Operating systems migration is the movement of the entire operating system including the processes, file systems and network connections (Hansen & Jul 2004). Data migration is the process of data transfer between databases, storage types and computers in an automated way with less human intervention (Geetha 2012).

User interface migration is the movement of user interfaces and it can be between devices where the users are allowed to control several services from one device (Svensson & Magnusson 2004). Ghiani *et al.* (2015) propose that migrating a user interface between devices can be accomplished in two ways: (i) pull migration – where the migration commences at the target device and (ii) push migration – where the user interface is pushed from the source device to the target device. Architecture migration is the movement from a legacy software arhitecture to a new software architecture with enhancements carried out within the constraints of the legacy software architecture (Cala *et al.* 2004).

2.9 Most Popular OSS Projects

The emergence of open source software (OSS) indicates that it is a recent major development in Information Technology (Chengalur-Smith *et al*. 2010). They point out that the commonly used OSS products are Apache web server, MySQL database, the OpenOffice office suite, the Firefox web browser, the Linux operating system and the DRUPAL content management system. They maintain that the business value that OSS brings to organisations includes tangibles such as cost savings and reliability as well as intangibles such as innovation and flexibility. Allen (2012) also stresses that software innovation has been democratised by OSS; however, he points out that there are doubts whether this innovation can be used in business applications where the end users are not the individual developers.

Table 2-3 provides short descriptions of some important OSS projects cited by different authors.

TABLE 2-3
Examples of Important OSS Projects

OSS Projects	Project Description	Author
Linux	Linux is the foremost enterprise server operating system. It is the open source equivalent of UNIX. It comes with many distributions and these include Ubuntu, RedHat, SUSE, Open SUSE and Debian.	Raghunathan *et al.* (2005) Kemp (2009) Stol *et al.* (2009) Chengalur-Smith *et al.* (2010) Frej *et al.* (2015) O'Neill (2012) Ratajczak (2015) Markus *et al.* (2014) Lu *et al.* (2014)

Perl	It was designed by Larry Wall. It is the software supporting the most ‘live content’ on the Internet. It has been around for 22 years and runs on more than 100 different platforms. It integrates easily with polpular databases and it is an ideal web programming language.	Raghunathan *et al.* (2005) Stevenson *et al.* (2015) Harvey (2015)
Python	It was designed by Guido van Rossum. It is used to integrate systems more effectively as a scripting or glue language that connects existing components together. It is used for rapid application development. Its strengths are its speed, flexibility and readable syntax.	Raghunathan *et al.* (2005) Goyal *et al.* (2015) Alomari *et al.* (2015) Singh *et al.* (2015) Harvey (2015)
PHP	It is a general purpose scripting language used in web development by enabling developers to write dynamically generated pages quickly.	Harvey (2015)
GNU Project	It is a set of high quality programming tools from the Free Software Foundation’s GNU project.	Raghunathan *et al.* (2005) Singh *et al.* (2015) Stoyanov and Kordov (2014)
Apache	It was originally created by Rob McCool in 1995. Currently, it is used in over 63% of today’s web servers which is more than Microsoft's IIS at 23%. It is the most popular web server for more than a decade. It is secure, efficient and extensible.	Raghunathan *et al.* (2005) Kemp (2009) Stol *et al.* (2009) Chengalur-Smith *et al.* (2010) Harvey (2015) O’Neill (2012) Stoyanov and Kordov (2014)

Mozilla	It is the open source version of its well-liked browser product Communicator and it is Netscape's next-generation web browser. Nescape produced the first commercially successful web browser called Navigator in 1994 as closed source software. Netscape decided to make the Navigator source code freely available under a project known as Mozilla in 1998 because the Microsoft internet explorer was made free.	Raghunathan *et al.* (2005) Kemp (2009) Stol *et al.* (2009) Pearson (2000) Friedman and Friedman (2015) Hudson (2015)
Ghost-script and Ghostview	Ghost-script is a postscript editor and printer. Ghostview enables the viewing of postscript files.	Raghunathan *et al.* (2005)
Open-Office.Org	It is the open source spreadsheet, word processor and other general-purpose office application software from Sun Microsystems.	Raghunathan *et al.* (2005) Chengalur-Smith *et al.* (2010) Singh *et al.* (2015)
Drupal	It is a content management system. It is a website tool which is very easy to set up a site for small organisations. It comprises of a community of more than a million developers, supporters and users. It has thousands of distributions, modules and extensions that can be used to set up a site easily.	Chengalur-Smith *et al.* (2010) Stoyanov and Kordov (2014) Harvey (2015)
FireFox	Mozilla FireFox is a secure and efficient web browser. It has a high speed and better privacy protection when compared with other browsers. It is made by Mozilla.	Singh *et al.* (2015) Stoyanov and Kordov (2014) Harvey (2015)
Alfresco	It combines simple web content management with document management and it is suitable for very small organisations.	Harvey (2015)

Android OS	It is mostly used for smartphones and tablets. It is the most popular mobile operating system.	Stoyanov and Kordov (2014) Harvey (2015)
WordPress	It runs on more than 60 million websites and blogs.	Stoyanov and Kordov (2014 Harvey (2015)
DevC++	It is a programming language.	Stoyanov and Kordov (2014)
Joomla	It is a very popular web content management solution like Drupal. It has more than 50 million downloads and its users include eBay, the United Nations, General Electric.	Harvey (2015) Stoyanov and Kordov (2014)
Moodle	It is an online education/eLearning tool. It has about 65 million users and it is easy to use. It is flexible, customisable, scalable and secure.	Harvey (2015)
MySQL	It is a web technology tool. It is the most prevalent open source database in the world. It is available in free community edition, paid standard, enterprise and cluster carrier grade editions. Users include Facebook, MTV Networks, Verizon wireless, Wikipedia.	Stoyanov and Kordov (2014) Harvey (2015)
Java	It was originally developed by Sun, but it is now owned by Oracle. It allows developers to write code that can run on multiple operating systems. It is the most popular programming language in the world.	Harvey (2015)
GNS3	It is a networking and communications tool	Stoyanov and Kordov (2014)

Table 2-3 highlights some examples of the important OSS projects undertaken. Linux is the leading enterprise server operating system. It is the open source equivalent of UNIX. It comes with many distributions and these include Ubuntu, RedHat, SUSE, Open SUSE and Debian. (Raghunathan *et al.* 2005 ; Kemp 2009;

Stol *et al.* 2009; Chengalur-Smith *et al.* 2010; Frej *et al.* 2015; O’Neill 2012; Ratajczak 2015; Markus *et al.* 2014; Lu *et al.* 2014). Perl is the software supporting the most ‘live content’ on the Internet. It has been around for 22 years and runs on more than 100 different platforms. It integrates easily with popular databases and it is an ideal web programming language (Raghunathan *et al.* 2005; Stevenson *et al.* 2015; Harvey 2015).

Mozilla is the open source version of its popular browser product Communicator and it is Netscape's next-generation web browser. Nescape produced the first commercially successful web browser called Navigator in 1994 as closed source software. Netscape decided to make the Navigator source code freely available under a project known as Mozilla in 1998 because the Microsoft internet explorer was made free (Raghunathan *et al.* 2005; Kemp 2009; Stol *et al.* 2009; Pearson 2000; Friedman & Friedman 2015; Hudson 2015). The table indicates that there are many popular OSS products in the market and they are widely used in the industry.

OSS is being developed and used by companies like Google, eBay and presently Facebook (Manfield-Devine 2008). This implies that the distinction between OSS and closed software might not be crucial because the OSS vendors are now the commercial enterprises. The services, support, and guarantees of continued development given by major OSS distributors such as Red Hat are the same as the closed source software vendors (Manfield-Devine 2008).

Manfield-Devine (2008) highlights that the same development tools, practices and at times, the same developers are being used by both OSS and closed source vendors. Tools and processes used by OSS vendors include: public bug tracking, regression tests, security architecture review, code-scanning (simple pattern matching, or static analysis), systems tests, and penetration testing. Many closed source software vendors use these tools because they are well known and trusted. Security tools are part of the components of the management framework on information sensitivity.

2.10 Benefits of OSS vs. CSS – a comparison

Table 2-4 is a comparison of the benefits of OSS and closed source software by different authors. The comparison reveals that there are more profound benefits of OSS than for closed source software.

TABLE 2-4
Comparing the Benefits of OSS and CSS

Benefit / Characteristic	Open Source Software (OSS)	Closed Source Software (CSS)	Author
Reliability	OSS has increased reliability over closed source software. The reason is that OSS is usually critically examined by many independent and enthusiastic developers during all its developmental stages.	The reliability of some closed source software is lower than that of OSS. The reason is that CSS is produced by a smaller number of developers who work against tight deadlines under much pressure.	Pearson (2000) Fitzgerald (2006) Gallegoa *et al.* (2008) Singh *et al.* (2015)
Sense of Urgency	There is little sense of urgency in OSS projects; there are few or no strict deadlines, and no hierarchical team structure in OSS developments.	Due to stringent deadlines to be met, there is a sense of urgency of CSS projects. There is a hierarchical team structure in closed source projects – the corporate world.	Kamthan (2007)

Quality	The quality of OSS is perceived to be higher than that of CSS. This is because many developers examine the software, facilitating the detection of errors. The quality of OSS products should be higher than for CSS if there is competition between them in the market	CSS is perceived to have a lower quality than OSS. Developers outside the closed group cannot detect errors because the source code is generally not publicly available.	Fitzgerald (2006) Khanjani & Sulaiman (2011) Gallegoa *et al.* (2008)
	Generally there are no formal inspections in the quality of OSS programs and no broad testing. There is little evidence to support rigorous measurements in OSS.	Quality of CSS could be higher than quality of OSS if there is no competition in the market.	Raghunathan *et al.* (2005)
		Formal inspections are conducted in CSS projects as well as broad testing. Rigorous measurement is performed in CSS implementations.	Kamthan (2007)

Innovation and Flexibility	OSS has more flexibility than CSS – source code is publicly available.	CSS has less flexibility than OSS due to its code being closed.	Fitzgerald (2006) Singh *et al.* (2015) Stoyanov and Kordov (2014)
	OSS enables innovation to modify the software without any restriction by providing users with the autonomy and being flexible.	Users are not allowed to see the source code and this restricts innovation. But it facilitates the security and reliability of the software. They have targeted innovation that is business focused rather than technology focused.	Gallegoa *et al.* (2008) Daniel (2009) O’Neill (2012)
Software Requirements	Requirements are mostly absent in OSS projects. There is little systematic effort in addressing Capability Maturity Models (CMMs). There is also little evidence of using the Unified Modelling Language (UML) or any form of systematic modelling in OSS.	Requirements are used in CSS projects. The Capability Maturity Model (CMM) is well addressed in CSS projects. Closed source projects make use of UML or other modelling techniques.	Kamthan (2007)
Vendor lock-ins	There is no vendor lock-in associated with OSS. The user is independent of the vendor.	CSS is dependent on the vendor. Therefore, there is vendor-lock in.	Fitzgerald (2006) Gallegoa *et al.* (2008) Singh *et al.* (2015)

Cost	OSS tends to be free; and have low acquisition cost, except for having to pay for the media on which the software may be distributed (e.g. on a CD).	Most CSS are not free and have a higher acquisition cost than OSS. However, in some situations closed source software's total cost of ownership (TCO) is lower than that of open source.	Fitzgerald (2006) Gallegoa *et al.* (2008) Vintila (2010) Raghunathan *et al.* (2005) O'Neill (2012) Singh *et al.* (2015) Stoyanov and Kordov (2014)
	The total cost of ownership may roughly be the same as for some closed source programs.	TCO for closed source and open source software could roughly be the same.	Daniel (2009)
Adherence to standards	The use of standards is limited to data formats like the Hypertext Markup Language (HTML), or the Extensible Markup Language (XML).	Closed source projects normally adhere to most IT standards during implementation.	Kamthan (2007)

Usability / Ease of code errors identification and problem solving	Most OSS products offer code error reporting tools. These tools assist in faster detection of errors and rapid finding of solutions. OSS usually lacks usability because it is developer-centric. Ability to correct errors is limited to users with technical expertise.	Generally, it requires a much longer period to resolve errors in CSS, due to non-availability of code error reporting tools. Closed source programs do not lack usability. They employ expert usability testing techniques and usability is ranked higher than in OSS.	Vintila (2010) Singh *et al.* (2015) O'Neill (2012) Circoria *et al.* (2012) Daniel (2009) Khanjani & Sulaiman (2011) Singh *et al.* (2015)
Operating Systems	OSS products have operating systems that surpass the CSS operating systems because their source code can be altered. Users can adapt the OSS to their operating systems. The cost of such a diversity of operating systems tends to be higher in closed source systems due to their high development costs.	It is more expensive to change the operating system source code of a CSS. Development costs are generally high. Users usually have to wait for a next release of the software.	Vintila (2010)

Documentation	Most OSS projects are weak on documentation. OSS products are not legally bound to produce documentation such as manuals or guides.	Most CSS projects produce manual and quality documentation. Closed source programs are legally required to supply documentation such as user manuals and guides.	Kamthan (2007) Daniel (2009)
Personalisation	This is the degree to which developers are able to write applications in the way they want the application to look and be used. OSS developers use personalisation a lot in their work in order to change the look and feel of a product, so that it can integrate seamlessly with their working environment. This enhances their efficiency and mood.	CSS developers are generally not allowed to attach personalisation to their work. Company standards and policies have to be adhered to and CSS is designed to accommodate the generic software market.	Vintila (2010)

Service and Product Support	OSS products come with many learning materials obtainable from the developer’s site or other locations supporting the OSS product. A large community of users and developers support OSS products by designing tutorials and short articles on how the product should be used.	Closed source systems are supported by a support team and they usually make use of printed material or books which come at a cost.	Vintila (2010) O’Neill (2012) Singh *et al.* (2015)
	User groups are available and support is delivered via forums and blogs. Issues may, or may not be resolved soon.	Closed source programs have a high response service. Ongoing support is provided to the customer. Support to the users of CSS is arguably the greatest advantage of using CSS.	Daniel (2009) Stoyanov and Kordov (2014)
Plug-in functionality	Plug-ins are readily available for OSS products. OSS developers and users can extend the functionality of their product by using plug-ins to write their own modules which can be integrated with the OSS product.	It is more difficult to write plug-ins for Closed Source Systems than OSS because documentation is not as rich as the OSS. The source code is also not readily available.	Vintila (2010)

Highly specialised Applications	OSS programs are less likely to be used to develop highly specialised applications.	CSS can be used effectively to develop highly specialised applications.	Raghunathan *et al.* (2005)
	There is little evidence that formal specifications are used in OSS projects and this limits the use of OSS in safety-critical software.	Formal specifications are used in closed source projects and this enhances their use in safety-critical software.	Kamthan (2007)
Best-practices Project Management	PM practices are usually lacking in most OSS projects and this could undermine the product's quality.	Most closed source projects use best-practices project management techniques, all of which enhance a product's quality.	Raghunathan *et al.* (2005)
	Release management guidelines are informal in OSS and there are often version proliferation and implementation issues.	Most closed source projects follow release management guidelines.	Kamthan (2007)
Reshaping the IT Industry	OSS has reshaped the IT industry in terms of server technology, cloud/virtualisation, content management and mobile.	CSS is slower in reshaping the IT industry when compared to the OSS.	Silic and Back (2015)

<u>Discussion of Table 2-4</u>

The reliability of some CSS may be lower than that of OSS owing to fewer programmers who develop closed source software, working against tight deadlines and under a fair amount of pressure (Pearson 2000; Fitzgerald 2006; Gallegoa *et al*. 2008). OSS is now considered as a reliable software (Singh *et al.* 2015).

Closed source software is perceived to have a lower quality and lower flexibility than OSS due to the non-availability of the source code (Fitzgerald 2006; Gallegoa *et al*. 2008; Khanjani & Sulaiman 2011). However, Raghunathan *et al*. (2005) and Kamthan (2007) argue that CSS is of a higher quality than OSS, provided that there is little or no competition in the market.

Most CSS implementations make use of a modelling Language like UML, as well as incorporating the Capability Maturity Model (CMM). In contrast, OSS implementations usually do not make use of any modelling techniques like UML, neither do they use the CMM (Kamthan 2007).

The Total Cost of Ownership (TCO) of OSS and closed source software are roughly the same (Daniel 2009). OSS suppliers are charging for additional items, extra administrations and these combinations reduce the gap in the TCO between the two types of software (Singh *et al.* 2015).

Closed source programs do not lack usability, documentation or service/product support, whereas OSS programs usually lack usability and documentation (Daniel 2009; Kamthan 2007; Singh *et al.* 2015). No vendor lock-in is associated with OSS but closed source software is characterized by vendor lock-ins (O’Neill 2012; Fitzgerald 2006; Gallegoa *et al*. 2008).

The differences between open source and closed source are not conclusive, but in a finer analysis are slightly in favour of open source (Raghunathan *et al.* 2005). Khanjani *et al*. (2011) concur, stating that OSS yields more benefits than CSS. More enthusiastic developers are involved in developing, testing and evaluating the code of OSS programs. OSS has reshaped the IT industry in terms of server technology, cloud/virtualisation, content management and mobile (Silic & Back 2015).

OSS software maintenance costs might be substantially lower than that of CSS, and the reliability of OSS is higher than that of CSS (O'Neill 2012). Thus the quality of OSS is higher than that of CSS, therefore a management framework on information sensitivity might well improve on the quality of OSS when migrated from proprietary to OSS by ensuring that the sensitive information are protected during such migrations. The benefits of OSS outweigh that of CSS as indicated in Table 2-4.

2.11 OSS Initiatives

This section describes various OSS initiatives undergone by different national governments (South African and foreign governments). The use of OSS gained momentum in the last decade in both public and private organisations (Weber 2004b; Marsan *et al.* 2012; Di Bella *et al.* 2013). Internationally, governments see OSS as a tool that can assist them to enhance affordable service delivery due to its low cost of implementation and maintenance (Mutula & Kalaote 2010). However, Oram (2011) reiterates that procuring OSS has proven difficult in governments and it is hard to get information on government usage of OSS.

Some studies have shown that OSS has a tendency to be cheaper to procure but more expensive on subsequent consultation and maintenance (Hauge *et al.* 2010;

Drozdik & Kovacs 2005; Gallopino 2009; Poulter 2010; Mutula & Kalaote 2010; Singh *et al.* 2015). The close to zero licence costs of OSS does not necessarily translate to lower costs (Shaikh & Cornford 2012; Singh *et al.* 2015; Stoyanov & Kordov 2014). Cost is a factor that the researcher included as part of the information management framework on information sensitivity because of the reasons above.

2.11.1 South African Government Initiatives

The South African Government has strongly stated the yearning to use FOSS since 2001 (Miscione & Johnston 2010). The South African cabinet accepted two FOSS policy submissions, one by the National Advisory Council on Innovation (NACI) in 2002 and the other by the Department of Arts and Culture, Science and Technology in 2003 (Weilbach and Byrne 2010). The Government IT Officers' Council's (GITOC) FOSS Working Group compiled the 2003 FOSS policy for government (Cabinet Memorandum No. 29 of 2003) and this promoted the use of FOSS in the SA Government (Weilbach & Byrne 2010).

A FOSS policy was approved by the South African Cabinet in 2007 (Weilbach & Byrne 2010), stipulating that all future software should be based upon open standards and encouraged the migration of current government software to FOSS (GCIS 2007). The State Information Technology Agency (SITA) with the Council for Scientific and industrial Research (CSIR) established a project office (FOSS Programme Office) that will oversee the implementation of this policy (Weilbach & Byrne 2010). Despite these decisions of government, FOSS adoption has not met its targets (Weilbach & Byrne 2010).

The South African government started implementing FOSS within its departments since 2006 and has a target of 60% for back-end servers running FOSS (Vital 2006). However, the results of a survey conducted by Weilbach and Byrne (2010) from November 2007 to March 2008 suggest that FOSS is not yet widely used within the South African government. They conclude that FOSS implementations in the SA government are rather few.

The South African government is still using FOSS in the development of their software systems. For example, FOSS components are used to develop the Integrated Financial Management System (IFMS) as explained in section 1.1.2. This was done to lower the cost of supporting the software (e.g. licence costs) and also to improve on the quality and productivity of the IFMS software. Protecting sensitive information during migrations is a way of improving the quality of a software system migration and this research should assist the South African government to improve the quality and productivity of their software system migrations. FOSS is being used by South African government departments and many FOSS migrations have been performed by government departments, and this is why the researcher used some of the employees of these government departments to participate as respondents to the questionnaires in this research.

2.11.2 Foreign Government Initiatives

This section outlines some of the various foreign government initiatives on the adoption and implementation of FOSS in their organisations.

(a) Indian Government: According to Miscione and Johnston (Miscione & Johnston 2010), the Indian Government supports the use of FOSS and has clear policies in this regard. Sharma and Adkins (Sharma & Adkins 2006) claim that India has implemented many projects in support of FOSS adoption. FOSS implementations have been carried out in many countries, e.g. China (Yeo *et al*. 2006), Pakistan (Rafiq & Ameen 2009), and the South Americas (Hedgebeth 2007). The Indian government has come up with policies on Linux and other OSS for several years now (Manthena 2011).

(b) Malaysian Government: The Malaysian government provided comprehensive implementation guidelines for FOSS adoption (Thomas 2007) and approximately 128 Malaysian state agencies migrated desktop users to FOSS by March 2008 as detailed in the Malaysian Public Sector Open Source Software Master Plan (TMPSOSSMP 2008).

(c) Brazilian Government: The Brazilian government also implemented and adopted FOSS (Lewis 2007) and has a large number of FOSS developers and contributors (Mtsweni & Bierman 2010). According to SERPRO (2005), FOSS was used by almost 60% of state departments in Brazil in 2005. Shaw (2011) has pointed out that a group of Brazilian proponents of social change joined the FOSS communities and accelerated FOSS adoption by many Brazilian Government Agencies during the earlier part of the Lula Administration. The competence of IT professionals' impacts on the Brazilian FOSS adoption and the use of FOSS in Brazil has sky-rocketed because many Brazilian educated professionals are committed to FOSS (Oram 2011).

(d) German Government: The German government also implemented many FOSS projects: migration from MS Exchange 5.5 to KOLAB (Nagler 2005), migration of 14000 Windows desktop and laptop computers by the Munich Municipality in 2004 to Linux and OpenOffice.org (Kovacs *et al.* 2004), migration of 10,000 desktop machines by the German Foreign Office to FOSS across 300 sites in 2007 (Otter 2007). The central administration of Germany signed an agreement with IBM to supply FOSS products based on Linux at a reduced price (Mutula & Kalaote 2010).

(e) US Government: The US Government launched its recovery .gov Website known as Drupal and it was based on an Open Source Content Management System (Scola 2009).

(f) British Government: The British government adopted a policy on FOSS in 2002 (Mutula & Kalaote 2010). The objectives of this policy include the use of products based on open standards, and avoiding problems of over-dependency on a specific supplier. The policy enhances the use of FOSS in all publicly funded British organisations (Central Government Departments and their Agencies), local governments, non-departmental public institutions, the National Health Service (NHS) and the educational sector.

(g) French Government: France set up the Agency for Information and Communication Technology (AICTA) in 2001 and it facilitates the use of FOSS by public agencies (Nagler 2005).

(h) Spanish Government: The Spanish Ministry of Industry, Tourism and Trade gave financial support for FOSS implementation to various government institutions and autonomous administrations (CENATIC 2008). Some FOSS implementations include GNU/Linux, Guaclalinux, Guadainfo, Linkat, Council of Zaragoza and MAX.

Many foreign governments migrated to FOSS in order to lower the cost of supporting the software (e.g. license costs) and also to enhance the quality and productivity of their systems. Protecting sensitive information during migrations is a way of improving the quality of a software system and the research reported on in this thesis may well assist foreign governments to improve on the quality and productivity of their software migrations with respect to information sensitivity protection.

2.12 Addressing OSS Security

Many IS security researchers have concentrated on the development of algorithms and protocols for the encryption, authentication and integrity of data (Hussain *et al*. 2005; Lafuente 2015; Choi *et al.* 2014; Kaushal *et al.* 2015). They maintain that since operating systems (e.g. Windows, UNIX, Linux) do not protect sensitive information by default, three security levels (low/medium/high) can be introduced to protect sensitive information.

Tools are used to protect sensitive information during software migrations. For example, Brin *et al*. (1995) state that copying sensitive files to removable media can be blocked by some tools which also disallow sensitive files to be included in email attachments by using copy detection techniques. Tools like digital rights management can be used to protect sensitive information by using encryption (Kale *et al.* 2015; Ku & Chi 2004). Tools and encryption are part of the components of the management framework due to the reasons above.

Adherence to policies is important during software migrations and Kurita *et al*. (2007) propose a technique to trace and regulate how programs read sensitive information by establishing security policies that grant or deny permissions to output devices, as well as saving and protecting sensitive data in adherence to such policy. An information flow control model that is used for protecting and sharing sensitive information was proposed by Arai and Tanaka (2009). Program execution environments were built and separated based on the type of information and privileges were given based on the execution environment. Adherence to policy is included by the researcher as part of the management framework for the reasons above.

This section covers the standard ways of resolving security problems during OSS migration; and also motivates the need to have a management framework for information sensitivity during software migrations. Some of these components e.g. tools, policies, monitoring and controlling of sensitive data, have been used to build and develop the management framework. Section 2.14 motivates the use of a management framework in conjunction with existing solutions.

2.13 Properties of a Management Framework

There is the requirement for research to comprehend human conceptualisations of sensitive information and also to find the difference between sensitive information and other organisational information for security purposes (Thompson & Kaarst-Brown 2005). They maintain that most sensitive information can be regarded as not being steered by technology. This implies that security solutions might not be based on technology alone, but also with the management of the processes involved in the protection of the system. This is why PoliVec (2002) points out that organisations need to isolate information based on its sensitivity as suggested by some security solutions. Jones (2002) supports this view and states that more technology cannot resolve security problems; rather the basic models of security being employed by organisations need to be managed.

Organisations ought to perform sensitivity assessment as Farrell (2002) suggests that organisations must perform sensitivity assessments to elicit the different security requirements for information in both manual and electronic systems even if the organisations may already have an understanding of the different protection needs for information in both manual and electronic systems.

Information classification is important when protecting sensitive information and the British Standards Institute (2000) indicates that organisations need to know which information necessitates the most security and which may require less protection using their levels of information sensitivity. They emphasise the need of a classification system to realise this goal. Organisations should be able to classify information based on its sensitivity and use such classification to protect sensitive information in their organisations (Kaushal *et al.* 2015; Thompson & Karst-Brown 2005). Based on a comprehensive review of articles in the literature, organisations should use the BS17799 classification scheme to classify their sensitive information during migrations of software platforms since it is an international standard that has been tested and widely approved.

The examination of business rules can assist in protecting sensitive information. Liddy (2001) indicates that business rules should be examined to provide a foundation for information categorisation with respect to information sensitivity. For this reason business rules are included as part of the management framework developed by the researcher.

During new software systems design and implementation, the security and network controls are important aspects that ought to be considered. Scholz (1990) indicates that when new software systems are being designed and implemented, the security of the system and the network controls ought to be taken into consideration. Network control is one of the factors used to build up the management framework. More specifically, ‘wireless’ networks are highly vulnerable to hacking and should be protected by using network controls (Vargas *et al.* 2012; Scholz 1990).

The important aspects that affects the core of a security program that protects sensitive information include integrity, confidentiality, monitoring access, identifying authorised uses and the flow of information, and having knowledge of where information is at any point in time (Biot-Paquerot & Hasnaoui 2009). Integrity, confidentiality, monitoring access, identifying authorised uses and the flow of information are part of the components of the management framework.

A five-step suggestion is made by Cate (2006) to universities for the management of their sensitive information: stopping collecting data for the sake of data collection; implementing protection tools; commitment to privacy and security; getting involved in the legal debate on privacy rights; training and creation of executive leadership with resources to manage sensitive information. Following these suggestions, tools, training and data classification are all part of the management framework.

In order to safeguard sensitive information, organisations should: educate employees, validate the people and systems and update the program with changes as needed; mitigate risk by adopting insurance coverage; develop and implement policies and procedures to protect sensitive information; assess organisational data with a dedicated data security team; enforce hardware and software standards to eliminate unknown factors that access sensitive information without being authorised to do so (Augustinos 2009).

Rakers (2010) highlights that managing sensitive information involves people, technology and information, but the people are the most critical component, yet it is the most neglected part when managing sensitive information. There should be attention on processes, policies and technology when managing sensitive information (Lacey 2010; Oyelami & Ithnin 2015; Tankard & Pathways 2015). People, policies, processes, technology and information are all part of the management framework . Policies, organisational data, enforcing standards, training and risk assesments are all part of the management framework.

Managing sensitive information involves the following guiding principles: the development of a clear objective; the alignment of the objective with the organisational strategy; using multiple methods to accomplish the objective and understanding and planning for change (Ma *et al*. 2009). Organisational strategy is part of the management framework.

Changes in employee awareness, attitude and behaviour should be facilitated when protecting sensitive information. The view of Da Veiga and Eloff (2010) is that the employee behaviour should be focused on when managing sensitive information. Employee awareness, attitude and behaviour are part of the management framework.

Data accountability is also important when data is being protected. Pearson (2009) highlights that organisations have to value accountability when handling data and build mechanisms for accountable and responsible decision-making. He maintains that obligations to protect data must be observed by all who process data, independent of where such processing occurs. Data accountability is part of the management framework.

The overall goal is to decrease privacy risk and as with security, however, it is necessary to take this into consideration from the outset of the migration process and not just add privacy mechanisms at a later stage. Table 2-5 illustrates the building blocks for a management framework. The table synthesises the abbove observations on the components of a management framework as per suggestions in the literature.

TABLE 2-5
Building Blocks for a Management Framework

Component in framework	Author(s)	Suggestion or Challenge Noted
Classify and categorise sensitive data / Develop a data classification system	Thompson and Kaarst-Brown (2005) Kaushal *et al.* (2015)	Suggest that organisations should classify and categorise sensitive information based on the behaviours of people in organisations
	PoliVec (2002)	Suggests that organisations should segregate information based on their sensitivity
	British Standards Institute (2000) Tankard and Pathways (2015)	A classification system is needed to address security issues
Address the basic models of security within an organisation	Jones (2002) Nazareth and Choi (2015)	Suggests more technology cannot resolve security problems but basic models of security employed by organisations ought to be addressed.
Commit to privacy and security by the organisation / Deploy protection tools to protect sensitive data / Assign executive leadership to manage sensitive information.	Cate (2006) Tankard and Pathways (2015) Performanta (2015)	Points out the five steps to manage sensitive information: commitment to privacy and security; protection tools; no unnecessary data collection; executive leadership to manage sensitive information and participation in legal debates
Assess the organisational data / Enforce hardware and software standards	Augustinos (2009) Oyelami and Ithnin (2015)	Suggests ways to protect sensitive information: Policies and Procedures; organisational data assessment; hardware and software standards enforcement
Train users on how to handle sensitive information	Da Veiga and Eloff (2010) Augustinos (2009)	Focus on employee behaviour, employee training; systems/people validation and risk mitigation

Perform a sensitivity assessment	Farrell (2002)	Suggests organisations ought to perform sensitivity assessment to identify different protection needs for information
Understand the business rules	Liddy (2001)	Indicates business rules should be examined to provide a foundation for information classification with respect to sensitivity
Consider confidentiality, integrity, identifying authorized uses, monitoring access and the flow of information	Biot-Paquerot Hasnaoui (2009)	Indicate that confidentiality, integrity, identifying authorised uses, monitoring access and the flow of information and knowing where information is at any point in time
Guiding principles	Ma *et al.* (2009)	Indicate four guiding principles to manage sensitive information: develop a clear objective; align the objective with organisational strategy; use multiple methods to accomplish the objective and understand and plan for change
Focus on policies, processes, technology, a change in employee awareness, attitude and behaviour	Augustinos (2009)	Suggests five ways to protect sensitive information and one of them is policies and procedures
	Lacey (2010) Oyelami and Ithnin (2015) Tankard and Pathways (2015)	Argue that there should be a focus on policies, processes, technology, a change in employee awareness, attitude and behaviour
	Rakers (2010)	Points out that there are three primary aspects when managing sensitive information and these are people, technology and information

Value accountability and build mechanisms for accountable and responsible decision-making	Pearson (2009)	Advises organisations to value accountability when handling data. Build mechanisms for accountable and responsible decision-making

Table 2-5 reveals that there is the need to classify and categorise sensitive data as well as to develop a data classification system (Thompson & Kaarst-Brown 2005; PoliVec 2002; Kaushal et al. 2015; British Standards Institute 2000). The basic models of security should be addressed within an organisation (Jones 2002; Nazareth & Choi 2015) and organisations should train their users on how to handle sensitive information (Da Veiga & Eloff 2010; Augustinos 2009). Organisations should focus on policies, processes, technology, as well as a change in employee awareness, attitude and behaviour (Augustinos 2009; Lacey 2010; Rakers 2010; Oyelami & Ithnin 2015; Tankard & Pathways 2015).

From Table 2-5, the building blocks for a management framework include data classification, security models, protection tools, assessing organisation data, user training on sensitive information handling, business rules, policies and procedures, changing employee awareness, attitude and behaviour. To protect sensitive information during the migration from a proprietary to a FOSS platform, the development of a management framework with the steps/actions as indicated in Table 2-5 is suggested:

(a) Develop sensitive information policies and procedures (Augustinos 2009; Oyelami & Ithnin 2015; Tankard & Pathways 2015);

(b) Know what sensitive information you have to migrate (Federal Trade Commission 2009);

(c) Classify the information to be migrated (Kaushal *et al.* 2015; Thompson & Karst-Brown 2005; Tankard & Pathways 2015);

(d) Encrypt sensitive information stored or transmitted electronically (Kaushal *et al.* 2015; Lafuente 2015; Ku & Chi 2004);

(e) Keep only the sensitive information you need and comprehensively destroy sensitive information that are no longer needed (Federal Trade Commission 2009);

(f) Train users (managers/developers/analysts and others) who will migrate the sensitive information (Da Veiga & Eloff 2010; Augustinos 2009);

(g) Use privacy-enhanced technologies (Cate 2006; Tankard & Pathways 2015; Performanta 2015);

(h) Develop a response plan to a security breach of sensitive information (Federal Trade Commission 2009).

2.14 Towards a Rudimentary Management Framework

The Rudimentary Management Framework is synthesised from the building blocks in Table 2-5 and is illustrated in Figure 2-2.

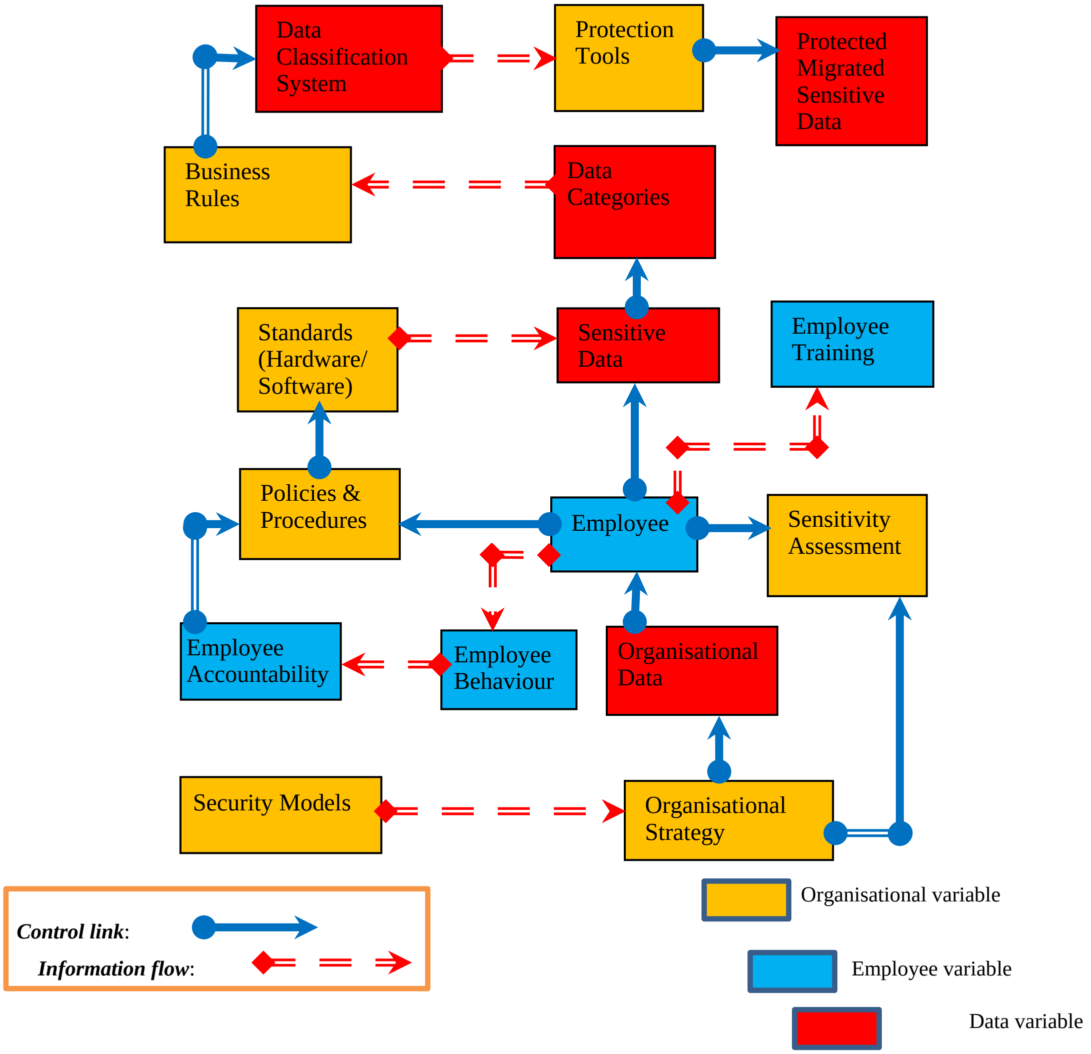


Figure 2-2: The Rudimentary Management Framework of Information Sensitivity during Software Migrations

This is a preliminary framework obtained from the literature and a more comprehensive one follows in the thesis.

Discussion on the Rudimentary Management Framework

Organisations migrating sensitive information have to develop security models to support their organisational strategy. The organisational strategy can incorporate how organisational data will be protected and handled. Organisations ought to develop clear objectives to manage sensitive information through a dedicated data security team.

Employees handling organisational data have to be trained on how to protect sensitive information and the changes in employee awareness, attitude and behaviour ought to be facilitated. Employees need to perform sensitivity assessment as part of the organisational strategy on the protection of their organisational data.

Policies and procedures on sensitive information need to be developed and enforced by management. Employees have to be made accountable to ensure that sensitive information protection is in line with the policy and procedures governing sensitive information. Such policies and procedures have to be used to enforce hardware and software standards in order to eliminate unknown factors that access sensitive information.

Data has to be categorised into categories using business rules and data classification systems. Data has to be classified into categories of critical importance and in accordance to the cost involved in collecting, organising and maintaining the data. Organisations need to examine business rules in order to provide a foundation for information categorisation with respect to sensitivity.

The information to be migrated needs to be classified using the data classification system. Sensitive information needs to be encrypted using the data protection tools and privacy-enhanced technologies. Organisations need to develop a response plan to be implemented when the security of sensitive information is breached.

2.15 Conclusion

In this chapter, sensitive information is defined, based on definitions from different authors in the literature. Moreover, the understanding and being able to identify sensitive information has facilitated the development of the comprehensive management framework on information sensitivity during software migrations which follows later in the thesis.

The desirable properties and the building blocks of such a management framework are noted (see section 2.13), and on the strength of these, a preliminary and high-level framework for sensitive information protection during software migrations is defined. Some of the concepts/components in this chapter are used to develop the management framework, the one which follows later.

Although many researchers claim that OSS platforms have much security, due to their openness (Hoepman & Jacobs 2007; Walker 2004; Wheeler 2005; Witten *et al.* 2001) more still has to be done. This chapter argues in favour of a management framework to address the protection of sensitive information in migrating from a proprietary platform to a FOSS platform.

The next chapter further explores the literature on the security challenges during OSS migrations leading to a proposed model on these issues.

Chapter 3

Security Challenges During OSS Migrations

3.1 Introduction

The rudimentary management framework is proposed in the previous chapter as a result of reviewing the literature and conceptualising the categories obtained from the literature in order to develop such framework. In this chapter, the main focus is on establishing a need for developing a model that can be used to address the security challenges during closed source software to OSS migrations. The rudimentary framework will be augumented with the security aspects in this chapter.

The layout of the chapter is as follows: section 3.2 compares OSS and CSS security. The security challenges in OSS are explored in section 3.3 while section 3.4 explains the challenges during migrating from closed source to open source. A model is proposed to address a number of security challenges during the migration from CSS to OSS (see section 3.5). The model is synthesised from the framework proposed by Aner & Cid (2010) and will augument the rudimentary framework to protect sensitive information during system migrations, suggested in Chapter 2. How the model can be implemented is explored in section 3.6 while section 3.7 concludes the chapter. The content of this chapter was synthesised into a research publication (Ajigini *et al.* 2014).

3.2 Comparing OSS and CSS Security

Hansen *et al*. (2002) emphasise the importance of analysing a whole OSS system when performing an extensive security investigation. Such analyses include the application software, its source code, and the tools used for developing the object code. Examples are compilers, operating systems, hardware and the whole development environment.

Different authors have different perceptions when they compare OSS security with that of CSS as shown in Table 3-1. The table reveals that the security of OSS is far beter than that of a CSS system.

TABLE 3-1
Comparing OSS and CSS Security

Characteristic	**OSS security**	**CSS security**	**Author**
Publishing of Designs and Protocols	OSS designs and protocols are published and these contribute to the security of the systems. This may reveal logical errors in the security of the system.	Closed source designs and protocols are not published.	Hoepman and Jacobs (2007)
Finding and correcting security vulnerability	It is easier to find and correct code errors in OSS than in CSS owing to the openness factor.	Open and closed approaches to security are rather similar. Correcting errors in CSS is dependent on the programming team that developed the program – the source code is not publicly available.	Dwan (2004) Manthena (2011) Schryen (2011)

Reliability	OSS is considered to be reliable as CSS because of the extensive work on reliability that has been performed on them.	CSS is developed by organisations and professional teams. This leads to unvalidated alteration and consistent trustworthy.	Singh *et al.* (2015)
Checking and Testing of Code	OSS users have the freedom to validate and test the code of the OSS product that they want to use in order to ascertain its quality and security.	Because users do not have the choice to validate and test the code in closed systems, the author stresses that OSS initial coding tends to be of a higher quality than that of CSS.	Manfield-Devine (2008)
Time to fix security vulnerabilities	OSS communities fix security vulnerabilities twice as quickly than CSS	It takes more time (twice more) to fix the security vulnerabilities of CSS than OSS.	Singh *et al.* (2015) Pokarna *et al.* (2015)
Controlled Environment Development	OSS is often viewed as having security issues because OSS development is not in a controlled environment.	CSS is perceived to be developed in a controlled environment by a dedicated team of developers with a common goal and thus it might be seen as being more secure. The source code is only seen and modified by this team. The software is comprehensively audited, and this reduces the risk of back door Trojans and further limits the risk of code errors or other software issues.	Daniel (2009) Manthena (2011)

Availability	OSS is freely available over the internet and they have 24 by 7 support from the online community and discussion forums.	Only the trial version of CSS is accessible free of charge for downloading and testing	Singh *et al.* (2015)
Closedness or openness of software code – security through obscurity	OSS improves software transparency, security and trustworthiness because users and developers can validate an OSS program's functionality and security, due to the availability of its source code.	The authors stress that the security of software is dependent on the user and not necessarily its closedness or openness. CSS can also be as secure as OSS.	Hansen *et al.* (2002) Circoria *et al.* 2012
	They highlight that it is easier to correct bugs in OSS systems, thereby enhancing the quality of code. This could also lead to the use of better project management and quality control. Open source users can independently evaluate the security for themselves. The real exposure of the system can be assessed and the gap between perceived and actual exposure is diminished.	CSS does not allow users of such software to evaluate its security for themselves. This does not allow users to easily discover weaknesses and 'patching' is not possible by users.	Hoepman and Jacobs (2007)

Analysis of published vulnerabilities	No substantial differences in terms of the severity of vulnerability were found between open source and closed source.	The vulnerability severity found between open source and closed source are perceived to be the same.	Schryen (2009) Schryen (2011)
	More and faster patches can be found in open source systems. Patches for open source systems are released faster than for closed source systems.	Patches for vulnerabilities of closed systems are released weeks or months after the discovery of the vulnerabilities and this increases the risk of using the system.	Hoepman and Jacobs (2007)
	Patch management is harder to co-ordinate in open source systems because OSS comes in many different versions. Patches will not be available for some distributions and they may be vulnerable to attacks while others are being patched.	The authors claim that it is easier to manage patches in a closed source system than in an open source system.	Clake *et al.* (2009)
	OSS products are more secure than CSS products. However, their general pattern of vulnerability detection is similar.	CSS products are less secure than OSS products.	Pokarna *et al.* (2015) Walia *et al.* (2006)

Discussion of Table 3-1

Closed source designs and protocols are not published, whereas the OSS designs and protocols are published enhancing the security of OSS programs since logical errors may be revealed (Hoepman and Jacobs 2007). This is also the view of some authors (Dwan 2004; Manthena 2011; Schryen 2011) that it is easier to find and correct errors in OSS than in CSS because of the openness of OSS code. OSS users have the freedom to validate and test the code in order to ascertain its quality and security, therefore OSS initial coding tends to have higher quality and security than CSS (Manfield-Devine 2008).

CSS is perceived to be more secure than OSS because it is developed in a controlled environment by a dedicated team of developers with a common direction (Walia *et al.* 2006 ; Daniel 2009; Pokarna *et al.* 2015). Moreover OSS is considered to be reliable equally as CSS and the time to fix security vulnerabilities is twice as quickly than CSS (Singh *et al.* 2015), although Hansen *et al*. (2002) contend that CSS can be as secure as OSS because the security of software is dependent on the user and not on its openness or closedness.

The severity of vulnerabilities found between OSS and CSS is similar (Schryen 2009). Furthermore, more and faster patches are found in OSS whereas patches are not released as fast in CSS, thereby increasing the risk of using the system securely (Hoepman and Jacobs 2007). OSS improves the software transparency, security and trustworthiness howevver the security of CSS is dependent on the user and not necessarily its closedness or openness (Hansen *et al.* 2002; Circoria *et al.* 2012).

OSS is more secure than CSS as illustrated in this section. However, organisations have to consider security challenges when migrating from CSS to OSS because there are security challenges that have to be overcome when migrating from a closed system to an open system (Geetha 2012). Risk management systems and other security monitoring tools are part of the management framework developed in this thesis. Thus, the management framework developed in this thesis may assist in managing the security challenges during software migrations.

3.3 Security Challenges in OSS

While OSS offers a number of advantages, notably cost efficiency and reduced vendor lock-in, it does, however, raise a number of security concerns. Some security concerns regarding the migration from proprietary platforms to OSS platforms are phishing, stealing sensitive information, e.g. account details and cookies and getting hacked during the process (Mtsweni & Bierman 2008).

The problem of the security of OSS was highlighted by two events, namely, a report released by Fortify Software in July 2008 (Open Source Security Study Fortify Report 2008), claiming that necessary standards were not achieved by OSS developers, and that Linux kernel developers had covered up security vulnerabilities (Manfield-Devine 2008). It was recommended in Fortify's report that OSS should be viewed warily due to alleged high risks involved by government and commercial organizations. The report further recommended the conducting of risk analyses and code reviews on any OSS code that runs in business-critical applications.

According to Manfield-Devine (2008), the US Department of Homeland Security, which is part of the US government's Open Source Hardening Project, backed Coverity Software (Open Source Coverity Report 2008) to investigate security issues affecting OSS products, and they produced a report that disagreed with the Fortify findings. Coverity Software analyzed 55 million lines of code across 250 projects (amongst others Linux and Apache) and concluded that OSS quality and security are improving.

Some of the Security Challenges in OSS include:

(a) Linux Security Concerns: According to the US National Security Agency (NSA), Linux security has been enhanced to cater for access controls, but they acknowledge that more work is still required to make SE Linux a trusted operating system that meets requirements of governments or corporate users (NSA 2001). To enhance the security of Linux, the NSA informed Linus Torvalds to add backdoors into Linux (Engstrom 2013).

(b) Breaches of secrecy/unauthorised access: According to the Danish Board of Technology Working Group (Danish Board of Technology Working Group 2002), security in OSS for e-government includes protection against breaches of secrecy in the content of data communication (e.g. sensitive personal data, members of the public and companies' economic circumstances) and protection against unauthorised access to computers (e.g. destruction of data or hacking of websites).

(c) Lack of Linux security: From an analysis performed by Mi2g, it was found that Linux-based web server systems were increasingly attacked by system hackers and it was found that in the first six months of 2002, there was a 27% increase in successful system attacks (Mi2g Report 2002). Subsequently, Fitzgerald and Bassett (2003) have suggested that Open Source Software should not be used by highly security sensitive users and also not for critical systems. However, this view no longer holds since OSS quality has improved significantly since then (Silic & Back 2015; Pokarna *et al.* 2015)

(d) Software error corrections: Fitzgerald and Bassett (2003) have pointed out that much of the discussion around OSS security centers on software error fixes and is not about the security implications of the software architecture.

(e) Lack of security of OSS operating systems: Hussain *et al.* (2005) argue that operating systems (Windows, UNIX, Linux, for instance) do not protect sensitive information that is not captured on the screen. Security is a key aspect and an integral part of any software development (Vadalasetty 2009).

(f) Increase in cyber frauds and attacks on OSS users: The number of reported cyber frauds and attacks on OSS has increased and in 2007, the Federal Trade Commission (FTC) received 221,226 internet-related fraud complaints (Acello 2009).

(g) Lack of rigorous security level: Doinea (2010) stresses that due to the different types of open source software, many applications lack a rigorous security level and this might be a source of threats for OSS.

Aner and Cid (2010) claim that OSS should be evaluated from a security perspective to ascertain the level of security robustness or potential exposure to threats. They also stress that the increasing use of OSS may pose several security challenges to organizations.

Due to OSS source code ‘openness’, Vintila (2010) has pointed out that open source code is accessible to the users and such code may be enhanced for added functionality. Possible errors in the code can be corrected and overall improvements to the source code can be done.

According to Schryen (2011), few empirical studies and quantitative models on open source security appear in the literature, (e.g. Neuhaus *et al.* 2007; Alhazmi *et al.* 2007; Woo *et al.* 2006). However, after 2006, there has been an increase in using OSS as a research topic due to OSS adoption by organisations (Crowston *et al.* 2012). Schryen (2011) investigated empirically the published patches and vulnerabilities of open source and closed source software packages. He concluded that there are no significant distinctions between open source and closed source in terms of the severity of vulnerabilities, vendors' patching behaviour and the types of vulnerability disclosures over time.

3.4 Challenges during the Migration from Proprietary to OSS Platform

The migration from a proprietary software to an OSS software is frought with challenges which can be divided into technical and non-technical issues (Sarrab *et al.* 2013). The technical challenges include: security, performance, usability, technical infrastructure, integrity, support availability, data migration, information flow control, flexibility and ease of use, management and maintenance of OSS. The non-technical challenges include: organisational culture, human factors and legal issues (Sarrab *et al.* 2013). Some of the challenges during the migration from a proprietary platform to an OSS platform include:

- Usability: OSS's usability is regarded as one of the reasons that limit its use since most users use proprietary software (Sarrab *et al.* 2013). Usability is described using the following five characteristics: error frequency, efficiency of use, memorability, learnability, severity and subjective satisfaction (Nielsen 1993). OSS's usability should be advertised more widely to facilitate its use by many users (Sarrab *et al.* 2013).
- Performance: Software performance is one of the technical challenges for migration to OSS (Sarrab *et al.* 2013). OSS products tend to be performance competitive to the proprietary applications. The performance

of OSS should always be higher than proprietary applications (Metcalfe 2012).

- Lack of technical infrastructure: This is seen as one of the key challenges of OSS development because there is lack of good internet infrastructure and reliable broadband access for OSS development (Sarrab *et al.* 2013). Lack of software manuals in local languages also contributes to the quality of the technical infrastructure.
- Co-ordination and support availability: The challenge resides in communicating with large numbers of users and developers. All work has to be well co-ordinated and the development process should be transparent to these stakeholders (Bleek & Finck 2011). In addition, there is the need to train people on IT and programming concepts so that they can participate in the development of OSS products in an open source community (Sarrab *et al.* 2013). Users of the OSS products can then be taught on how to use them.
- Lack of technical support: The availability of very few OSS certification programmes for Information Technology support professionals leads to a lack of technical support (Van Belle *et al.* 2006).
- Interoperability and integration: The new OSS software may need to integrate with other, already installed, operational software and this might not be feasible due to vendor independence of OSS. The OSS implementation might not have taken into consideration the interoperability with other, already installed, operational software (ElHag & Abushama 2009). OSS development might not use user-centred design or established software engineering methods (ElHag & Abushama 2009).
- Organisational frame: In some OSS developments, developers are paid for their contributions while others are not paid. This has led to some ill-feeling amongst participating developers. They suggest that a new development rhythm should be found and communicated quickly enough to meet outside

expectations but still accommodate everybody willing to contribute (Bleek & Finck 2011).

- OSS code maintenance and service support: The main duties of software management and maintenance are error and flaw detection as well as correction and these should be performed during software development (Sarrab *et al.* 2013). There is the possibility of the fault detection and correction not being carried out and completed in the OSS development environment before the software is transferred to a live environment. This might lead to developers not being able to deliver higher quality products in a well-timed manner, making OSS code maintenance and management expensive. Organisations should invest in versioning and fine-grained comparison tools to trace changes carefully to facilitate knowing the impact of upgrading to a future release (ElHag & Abushama 2009). There is also the difficulty in getting qualified staff to support and maintain OSS (Van Belle *et al.* 2006).
- Integrity: The integrity of OSS is its ability to survive security attacks (Sarrab *et al.* 2013). There are attempts to reduce OSS vulnerabilities through secure development. However, externally malicious codes can be inserted using buffer overflow exploit during running of the OSS codes (Huda & Hisham 2009).
- Security: OSS are vulnerable to security flaws, errors and risks (Sarrab *et al.* 2013). Security errors and risks in OSS are detected rapidly and because the source code is available, the process of eliminating errors is also rapid (Huda & Hisham 2009). However, metrics for measuring software security for mission critical and real time software may be hard to come by (ElHag & Abushama 2009).
- Organisational culture: The changes that are required during the migration to OSS are easier with a centralised IT structure than a decentralised one (Sarrab *et al.* 2013). This is because the migration to OSS leads to physical and virtual organisational changes.

- Staff skills: The employees need to acquire new skills and knowledge about the OSS system. The employees might not want to accept the changes due to the steep learning curve (Sarab *et al.* 2013).
- Legal issues: There are about 80 OSS recognised licenses available since the beginning of 2010 (Sarrab *et al.* 2013). Different kinds of licenses are offered by suppliers of OSS products (Thomas 2005). Therefore organisations should understand the various license types.

3.5 A Model for Addressing the Security Challenges during Migration to OSS

An open source assessment framework and a threat modelling methodology, pioneered by Microsoft since 1999 (Shostack 2008), have been proposed by Aner and Cid (2010) to overcome the security challenges of OSS. The aim is to reduce the risks to confidentiality, integrity and availability and to identify and reduce threats, vulnerabilities and risks to an acceptable level. They mention that alternative methods to reduce risks include: (a) code auditing (b) penetration testing, and (c) using statistical analysis tools.

For Aner and Cid (2010), the threat modelling process consists of four stages, viz: (a) application analysis/diagramming (b) threat enumeration, (c) threat rating, and (d) mitigation options. They point out that the threat modelling approach with slight modifications can assist with the identification of security vulnerabilities, as well as investigating coding issues and implementation mistakes.

A rudimentary management framework to protect sensitive information during the migration to an open source system was developed in Chapter 2. The model that is proposed in this section for addressing the security challenges discussed in this chapter, is based in part on the threat-modelling framework in Aner & Cid (2010) and the sensitive information migration framework in Chapter 2.

The model is illustrated in Figure 3-1 and is discussed below:

(a) During the application analysis/diagramming phase (A), the applications are analysed from a flow of data perspective. All the aspects that make up the applications are catalogued and the relationships between the assets in terms of data exchange are identified through a UML class-oriented structure.

(b) The threat enumeration phase (B) consists of analysing each element in the class-oriented UML against a list of potential threats depending on the element type using the STRIDE Taxonomy (Swiderski & Snyder 2004). STRIDE is used as a classification schema to characterize known threats in accordance with the attacker motivation.

(c) The risk levels for each of the enumerated threats are determined and ratings of all threats are established during the threat rating phase (C).

(d) During the mitigation options phase (D), all functionality and patching are removed and other security controls are added and redesigned.

(e) The business rules and the data classification system are used to classify migrated data during the data categorisation phase (E).

(f) Data protection tools and privacy enhanced technologies are used to encrypt the data during the data encryption phase (F).

(g) The encrypted data is now migrated during the data migration phase (G).

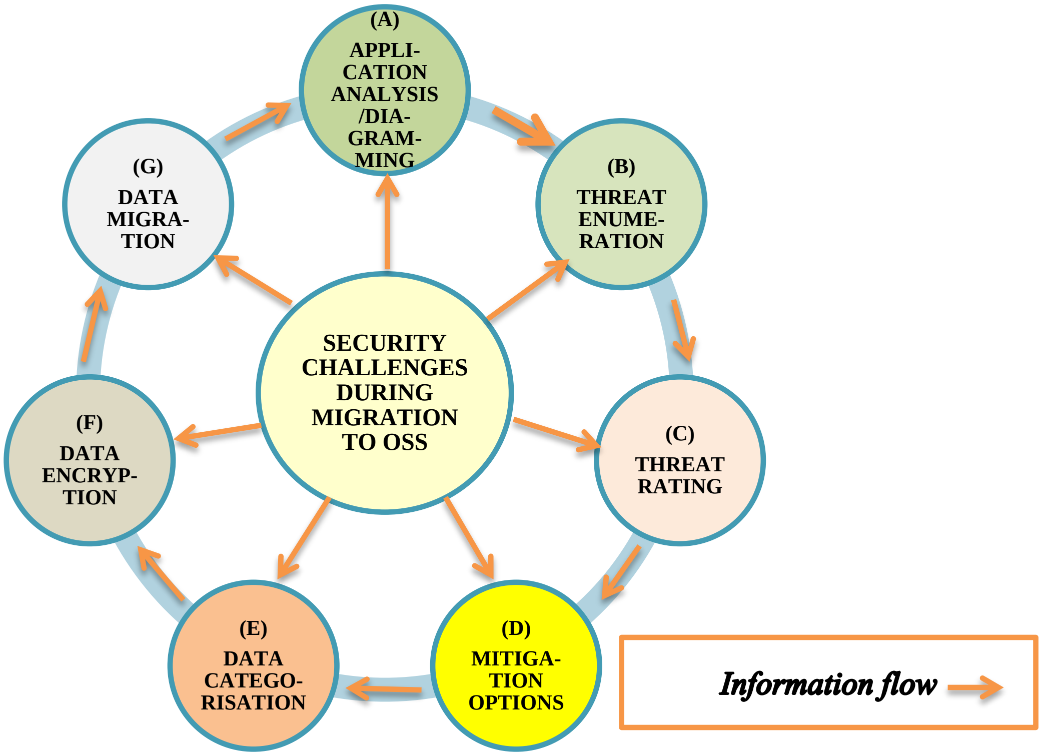


Figure 3-1: Modelling Security Challenges during OSS Migration

3.6 Implementing the Proposed Model

The following processes are proposed to implement the model in Figure 3-1:

Phase A: Application analysis/diagramming

(a) Identify security objectives – user identity protection, privacy and regulation, availability guarantees of applications.

(b) Catalogue all the applications.

(c) Analyse all the application designs and architectures to identify the components using data flows.

(d) Identify UML component diagrams.

(e) Identify the relationships between the assets using data exchange by using class-oriented UML structures.

Phase B: Threat enumeration

(a) Analyse each element in the class-oriented UML diagram against potential threats by using the STRIDE Taxonomy (Shostack 2014).
(b) Analyse data movement across trust boundaries (e.g. from Internet to Web tier).
(c) Identify the features and modules with a security impact that needs to be evaluated.
(d) Investigate how data enters modules, how modules validate and process the data, where the data flows to, how the data is stored and what fundamental decisions and assumptions are made by the modules.

Phase C: Threat Rating

(a) Identify threats using, e.g. Bugtraq tools and techniques. Bugtraq is a mailing list containing information on how to exploit and use intrusion detection systems vulnerabilities in defending networks (Vasa *et al.* 2015).
(b) Determine the risk levels of each threat.
(c) Establish the ratings of all the threats.
(d) Use either a threat graph or a structured list to write the threats.

Phase D: Migrations Options

(a) Remove the functionality and patching.
(b) Add other security controls.
(c) Redesign other security controls.

Phase E: Data Categorisation

(a) Develop business rules.
(b) Develop a data classification system.
(c) Classify data based on business rules and the above data classification system.

Phase F: Data Encryption

(a) Deploy data protection tools.

(b) Deploy privacy enhancement technologies.

(c) Use secure tools to encrypt the data.

Phase G: Data Migration

(a) Ensure that data to be migrated are encrypted by using verification techniques.

(b) Migrate the encrypted data.

3.7 Conclusion

In this chapter, two frameworks – one for threat modeling (Aner & Cid 2010) and another for protecting sensitive information during system migration (Ajigini *et al*. 2012) are integrated to propose a model for addressing the various security aspects in migrating from an open system to a closed system. The model is based on a seven-phase process as presented in Figure 3-1.

The rudimentary management framework proposed in section 2.14 for protecting sensitive information during system migration is further integrated with the security-protection model proposed in this chapter to develop the preliminary management framework on information sensitivity during software migrations in Chapter 5.

The next chapter relates to the research design and methodology, in which various concepts like research paradigms positivism and interpretivism, as well as research methods e.g. quantitative, qualitative and mixed methods, are explored.

Chapter 4

Research Design and Methodology

4.1 Introduction

The previous chapter focused on the review of literature on security challenges during OSS migration. A model for addressing the security challenges during migration to OSS was proposed.

This chapter details the research design and methodology employed to carry out the research work prior to the data analysis. An introduction to philosophical paradigms such as positivism, interpretivism, critical research and pragmatism is presented in section 4.2. Pragmatism is used as the underlying research philosophical assumptions underpinning this research and is discussed in sub-sections 4.2.5, 4.2.6 and 4.2.7. Quantitative, qualitative and mixed methods are explored in this chapter. The sequential explanatory mixed methods case study approach is used to carry out this research. Although, case studies usually focus on the current situation, the questionnaire placed emphasis on requirements for future developments based on the participants past and current experience of migrating sensitive data between platforms. Thus, the reasons why the focus of this research is on how organisations should manage rather than how they do manage sensitive information are explained in section 1.1.2. The data collection processes followed by the researcher are highlighted.

The layout of this chapter is as follows: section 4.3 highlights the various research methods used in IS research namely: qualitative, quantitative and mixed methods. A comparison of the three research methods is outlined and the reasons why the mixed methods approach was used in this work are provided in section 4.3.5.

Section 4.4 details the research methodology namely case study, while section 4.5 explains the data collection processes used to carry out this research. Section 4.6 highlights the data analysis procedures followed and section 4.7 concludes the chapter.

4.2 Philosophical Perspectives

The view of Khazanchi and Munkvold (2002) is that Information Systems (IS) research is guided by a research viewpoint or paradigm. This comprises ontological, epistemological and methodological assumptions framing the type of the research as well as the researcher’s role. Guba (1990) highlights that a paradigm comprises a pattern set of rules concerning the truth, knowledge of reality, and how the truth can be known. In the research literature, paradigms play an important role and it is an epistemological perspective as well as shared beliefs among the research community (Hall 2012). The view of Teddlie and Tashakkori (2009) on the paradigm is that it should be regarded as a worldview, combined with the various philosophical norms associated with the viewpoint. A worldview is a comprehensive structure of opinions and beliefs about the world while reality is everything that can be seen, smelt and touched (Gray 2004). Before starting on any research project, researchers are needed to reveal explicitly both their epistemological and ontological positions (Guba & Lincoln 1994).

Epistemology refers to the theory of knowledge that deals with the nature of knowledge, what constitutes valid knowledge, its scope and how we can obtain it. Ontology is the theory or study of existence (being), what constitutes reality while methodology is a procedure by which knowledge can be generated (Khazanchi & Munkvold 2002). Ontology refers to the nature of being or becoming existence while epistemology relates to the views of knowledge, where the knowledge is from, how it is acquired and what counts as knowledge (Klingner & Boardman

2011). Researchers' minds and beliefs on the nature of the phenomenon being investigated need to be clear, because the purposes and philosophies shape his or her views of the world to form his or her paradigms (Falconer & Mackay 1999).

Research methods (qualitative, quantitative and mixed-methods) are based on underlying philosophical assumptions on what can be regarded as valid research that use appropriate methods (Myers 1997). Therefore it is pertinent to know and comprehend the underlying research philosophical assumptions in a research method in order to conduct and/or evaluate the research method.

Three approaches are suggested by Orlikowski and Baroudi (1991) as well as Chua (1986) to study Information Technology (IT) in organisations – positivist, interpretive and critical. Goldkuhl (2012) highlights the importance of pragmatism and interpretivism as research paradigms for Information Systems qualitative research. Positivism, interpretivism, critical and pragmatist approaches will be discussed in the next subsequent sections that follow.

Pragmatism has been endorsed by Johnson and Onwuegbuzie (2004) as the underlying philosophy for mixed methods research. Another research paradigm is symbolic interactionism and it has been suggested by Benzies and Allen (2001) that symbolic interactionism could be integrated with other paradigms in multiple method designs. The world exists separate and apart from the individual's perception of it and hence the individual's perception of the world in which the researcher exists, does influence the researcher's behaviour. Hence according to symbolic interactionism, human beings should be regarded in the context of their environment. Symbolic interactionism postulates that each individual and his or her environment are inextricably linked through reciprocal relationships (Benzies & Allen 2001).

Other underlying paradigms include post-positivism and constructivism as suggested by Guba and Lincoln (1994). Constructivism is the cognitive processes by which people construct unique understandings and interpretations of the world (Leonardi & Barley 2008). They maintain that constructivism highlights subjectivity and foreground perception. Constructivism acknowledges that persons or organisations face local contigencies encouraging situated improvisations that lead to a unique pattern of practices and understandings (Leonardi & Barley 2008). Post-positivism will be discussed in the second section that follows.

Several scholars have discussed the question of positivism versus interpretivism (Lee 1989; Fitzgerald & Howcroft 1998; Weber 2004a; Orlikowski & Baroudi 1991; Walsham 1993, 1995). However, Johnson and Onwuegbuzie (2004) have pointed out that there are communalities among the traditional philosophical paradigms and they suggest the promotion of epistemological and methodological pluralism to perform more effective research. They state that researchers need to supplement one method with another by taking a non-purist or mixed position. Qualitative researchers should be able to use quantitative methods freely and vice versa. A mixed position which involves both quantitative and qualitative methods is employed in this research study.

4.2.1 Positivism

Auguste Comte (1798 – 1857), a French philosopher who founded the discipline of sociology endeavoured to combine empiricism and rationalism together within a new doctrine called ‘positivism’ (Bhattacherjee 2012). Positivism is regarded as the scientific method mostly used in the natural sciences and done by experiments to determine/discover universal laws. Positivism is also used in the social research and it is assumed that it can predict social trends and can be used to control events (Ryan 2006). The aim of positivist research is to test theory with a view of increasing the awareness of the phenomenon predictively (Myers 1997).

Positivists accept that there is an objective real world which can be known and described and it is beyond the individual’s body and that conclusions about reality are based on empirical observations that can be verified publicly (Schulze 2003). Reality can be explained using measurable properties that are free from the researcher’s instruments and the researcher (Myers 1997).

Positivism follows the universalist methods of science which disagrees that there is a fundamental difference between the social and the natural sciences (Myers & Klein 2011). Positivist researchers think that they can attain a full understanding based on experiment and observation (Ryan 2006), but this might not be the ideal way of doing IS research, because the human elements are not considered. IS research needs to incorporate the human element of people using IT systems (Pather & Remenyi 2004).

Positivism has the components of being reductionist, placing emphasis on empirical data collection, logical, based on a priori theories and positioned towards cause and effect (Creswell 2007). Positivism pursues the reduction of things to abstract and general principles by fragmenting human experience instead of treating it as a whole (Ryan 2006). This has led to opposition of positivistic epistemologies and developments in qualitative research, feminism, critical psychology, anthropology, post structuralism and ethnography. Critics of positivistic epistemologies believe that divisions between private and public knowledge or subjectivity and objectivity are constructed socially (Ryan 2006). Positivism is not used in this research because of its reductionism approach and also because of its extreme confidence in objectivity and empiricism which cannot stand up to inquiry when applied in both the social and natural sciences.

4.2.2 Post-positivism

Post-positivism is a slightly modified positivist paradigm (Devers 1999; Ryan 2006). Post-positivists agree that there are steady affliations among social phenomena. However, post-positivism recognises that research is guided by the theories used and values embraced by the researcher (Petter & Gallivan 2004). Post-positivism is also termed revised positivism due to the dissatisfaction with positivism. Post-positivists acknowledge that researchers are influenced by their own subjective selves in their work (Schulze 2003).

Today's quantitative researchers regard themselves as post-positivists since they believe that the truth can only be approximated and cannot be explained completely or perfectly (Onwuegbuzie *et al.* 2009; Ryan 2006). Quantitative researchers have adopted post-positivism in their work (Phillips & Burbules 2000).

In post-positivistic research, it is assumed that reality is composed of many simple elements (atomistic), distinct, visible events in which events are produced by predecessor attributes that function in a law-like fashion (Schulenberg 2007). They believe that a picture of reality can be outlined using linguistic, mathematical and graphic descriptions. Post-positivism is a replacement for positivism and it underpins empirical research methods in social sciences studies (Clark 1998).

Post-positivism does not have a well defined format prescribed to do research and this may make the replication of sound practices in further studies a challenging one (Morris *et al.* 2009). It is also not possible to separate the researcher's own bias or perspective from the research being done. The focus of post-positivism is on using multiple measurements in addition to observations and this may aid in the identification of bias found within interpretations (Trochim 2006). In post-positivism, it is not practical to separate the researcher's stance from manipulating the subject chosen for the research and this includes the interpretation of the results (Hutton 2009).

Post-positivism concerns reality that is socially constructed and not objectively determined (Noor 2008; Ryan 2006). Post-positivist criteria include dependability, transferability, credibility and confirmability (Devers 1999). Dependability is the degree of producing similar or consistent findings in the same research. Transferability is the degree of transferring findings to other settings. Credibility is the truth of the findings as viewed from the interviewed or observed and within the context in which the study is performed. Confirmability is providing evidence that supports the findings by the researchers. The post-positivist criteria of dependability, credibility and confirmability are used in this study. The view of Ryan (2006) is that post-positivist approaches are interpretive in which emphasis is placed on meaning, experience and knowledge and seeing the person as relational and multiple rather than being enclosed by reason. Post-positivism is not used in this research because post-positivism does not have a well defined format prescribed to conduct research and it is practically impossible to separate the researcher's own bias or perspective from the research being conducted.

4.2.3 Interpretivism

Interpretivism in IS research commenced in the 1970s with Boland (1979) highlighting the relevance of phenomenology and hermeneutics to IS research. Johari (2009) states that interpretive research has been used by many IS researchers because it encourages researchers to be more inductive and exploratory. It is also one of the theories of knowledge for studying IS research in organisations. IS researchers can comprehend human action and ideas in organisational and social settings by using interpretive research (Klein & Myers 1999). Interpretivism is more significant when IS is being studied from different cultural contexts and can provide guidance on how interviews should be conducted and case studies interpreted (Johari 2009).

Interpretive research involves understanding the phenomenon within cultural and contextual situations (Orlikowski & Baroudi 1991). They claim that interpretivism is a better paradigm than positivism when investigating IS in organisations, which has been acknowledged by different authors (Myers 1997; Yu 2003; Johnson & Onwuegbuzie 2004). Interpretive research is different from critical research but the two approaches have some overlapping areas (Myers & Klein 2011).

IS research can be regarded as being interpretive if based on the assumption that our knowledge of reality is gained from social constructions like documents, tools, language and other artifacts (Klein & Myers 1999). Klein and Myers (1999) propose seven principles for interpretive field research namely: the principle of multiple interpretations; the principle of abstraction and generalisation; the principle of interaction between the researchers and the subjects; the fundamental principle of the hermeneutic circle; the principle of dialogical reasoning; the principle of contextualisation and the principle of suspicion.

Interpretive researchers are informed by the social philosophies of phenomenology and solipsism (Falconer & Mackay 1999) and they attempt to comprehend the way research subjects hypothesise events, concepts and categories that influence individual behaviour (Kaplan & Duchon 1998). Solipsism is a philosophical idea indicating that only the mind exists and anything outside the mind is unjustified and individuals perceive concepts by abstraction from their inner experiences (Mastin 2008). Social constructs like language, shared meanings and consciousness are the approach to truth in interpretive research (Myers 1997).

Interpretive research often comprises using qualitative methods, however it is considered that research is not just interpretive if the type of data collected is qualitative. This is why interpretive research is not the same as qualitative research (Rowlands 2003). Myers (1997) also adds that qualitative research cannot be said to be interpretive unless the researcher's underlying philosophical assumptions are taken into consideration. Falconer and Mackay (1999) have listed interpretive research methods to include: case studies, Action Research, hermeneutic analysis, narrative analysis, Ethnography, use of metaphors, semiotic analysis and Grounded Theory.

Interpretive research in IS has gained momentum and from a study carried out by Mingers (2003), it was found that 17% of papers in six well-known US and European-based journals from 1993 to 2000 were considered to be interpretive research. This is also the view of Walsham (2006), that interpretive research is being widely used in IS research. The knowledge of interpretivism is important for an IS researcher and this prompt the researcher to understand interpretivism in this study. Interpretivism is not used in this research because interpretivists are subjective in nature, believing that people are not connected to the laws of science or nature and making conclusions which are personal, in depth and which cannot be necessarily generalised. Primary data obtained in interpretivist research cannot be generalised because the data is influenced by the researcher's values and viewpoints and thus the reliability and representativeness of data are weakened to a certain level.

4.2.4 Critical Research

The critical research group was started at the Frankfurt school by Felix Weil in 1923 and the group was influenced by Max Weber, Sigmund Freud and Herbert Marcuse (Pather & Remenyi 2004). The aim of the group is to overcome the limitations of positivism and phenomenology. Jurgen Habermas is regarded as one of the best advocates of critical social theory (Myers 1997). Critical research relies on the premise that the current social systems are based on history and can develop standpoints where current social practice can be analysed in order to replace it with other structures and norms (Falconer & Mackay 1999).

Critical researchers assume that social reality is produced/reproduced by people as well as historically composed (Myers 1997). They perform their research using the context of feminism, corporate power structures, Marxism, anti-colonialism and anti-racism (Pather & Remenyi 2004). Consequently, Critical research can be sub-divided as marxism, feminism and queer research (Oates 2006). Critical theorists believe that the world is not a universe of facts that exists independently of the observer (Khazanchi & Munkvold 2002). Moreover, the focus of critical research includes: oppositions, conflicts and contradictions in contemporary society and eliminates the causes of alienation and domination because it is emancipatory (Myers 1997).

The ability for people to change their social and economic circumstances consciously is inhibited by cultural, social, and political supremacy forms. Because of this, Critical researchers should declare their biases and interests in their research. In IS literature, critical research is gaining momentum and includes the work of Mingers (2003) and Pozzebon (2003). However, Falconer (2008) maintains that critical research is under-represented in the IS research literature.

A set of principles for performing critical research in information systems has been recommended by Myers and Klein (2011). They maintain that critical research is becoming an important stream in information systems. This is evident in the four special issues of IS journals that have been devoted to critical research (Cecez-Kecmanovic *et al*. 2008; Kvasny & Richardson 2006). Critical research in information systems comprises of social issues such as freedom, power, social control and values with regard to the use, development and impact of information technology (Myers & Klein 2011). Critical research is not used in this research because the purpose of this study is not to deliver social critic about society but rather to suggest a practical solution on a managerial level regarding sensitive information resources.

4.2.5 Pragmatism

Pragmatism is a philosophy of science based on actions (Taatila & Raij 2012; Dewey 1929; Peirce 1992; Blosch 2001). Action is needed to change existence and action needs to be guided by purpose and knowledge (Goldkuhl 2004; Blosch 2001). Although other matters may be of interest to the study, they have to be positioned around actions as the primary unit of analysis. Pragmatism affirms that reality possesses practical character (Dewey 1960). The world is seen as a set of practical actions that are born from thinking (Taatila & Raij 2012). Truth is considered to be the end result of an inquiry (Kelemen & Rumens 2012; Haack 1976).

Pragmatism has been advocated by Johnson and Onwuegbuzie (2004) as the 'philosophical paradigm' for mixed methods research. Goldkuhl (2012) states that IS research has been influenced greatly by pragmatist thinking. Pragmatism involves action and change, as well as the interplay between knowledge and action (Taatila & Raij 2012). This implies that the world is intervened with actions and pragmatists do not just merely observe the world. Pragmatism involves two-way interaction and alternative views have to be considered for new views to be acquired (Taatila & Raij 2012). Moreover, scientific work is created from experimentation and flexibility. Flexibility is regarded as the most important factor in pragmatism.

Many authors have followed the pursuit for pragmatism in IS research (Goldkuhl 2004, 2008; Goles & Hirschheim 2000; Marshal *et al.* 2005). Information Systems is regarded as a rational discipline that emphasises practical propositions and applied research as well as theoretical inferences (Agerfalk 2010). Pragmatic thinking could support IS researchers in the following areas: conceptual modelling, open source, e-government, information infrastructuring, social media and open innovation processes (Agerfalk 2010).

In pragmatism, the meaning of an idea or a concept can be regarded as the practical outcomes of the idea/concept. This implies that the different actions conducted during the research process are based on the belief of the concept (Goldkuhl 2012). Within pragmatism, the knowledge character is not restricted to understanding (interpretivism) and explanations (positivism), but includes other knowledge forms such as normative knowledge (exhibiting values), prescriptive knowledge (giving guidelines) and prospective knowledge (suggesting possibilities). The view of Goldkuhl (2012) is that all these different forms of knowledge within a pragmatist theory of knowledge should be regarded as constructive knowledge including explanatory and descriptive knowledge.

Pragmatism combines practical consequences and real effects as parts of meaning and truth (Venkatesh *et al.* 2013). That is, the value of an idea can be obtained from the practical consequences of accepting it (Hannes & Lockwood 2011). The concept of truth is very important and it is a key area in pragmatic thinking (Taatila & Raij 2012). Furthermore, pragmatism warns against conceptualisations that are not based on the empirical and practical human world (Goldkuhl 2012).

Hannes and Lockwood (2011) conclude that pragmatism is based on abduction reasoning and this supports using both quantitative and qualitative methods in a research study. According to Morgan (2007), pragmatism relies on abductive reasoning that fits in between deduction and induction and vice versa. Observations are firstly converted into theories and the theories are assessed through action. This results to a research process where quantitative and qualitative methods are sequentially merged with the inductive results from a qualitative approach serving as inputs to the deductive goals of a quantitative approach, and vice versa (Morgan 2007).

What implications can be inferred from practical IS research? There is a need to be careful and meticulous when developing concepts in IS research. Empirical data should allow to be subjected to traceability. Concepts should also permeate practical reality (Goldkuhl 2012). Simple concepts should be developed first and these can then lead to more complex ones. In pragmatism, the aim of the research is towards change, action and knowledge which contributes to IS practices improvement. Therefore a pragmatist position can explain and justify the development of a senstive information management framework, since the management framework is aiming at formulating and trying out what would be better in practice.

Pragmatists are not only interested in recording informants' conceptions but also are interested in actions and these include actions that are successful as well as those that are not (Marton 1981). A pragmatist asks people not only about their perceptions of the world but also what the people do, that is, there is the avoidance of narrow interpretivism that does not take into consideration change and improvement (Goldkuhl 2012).

Goldkuhl (2008) has explained three types of pragmatism, namely:

- Referential pragmatism
- Functional pragmatism
- Methodological pragmatism

Referential pragmatism allows actions, action-objects, actors, activities and practices to become the primary studied objects (knowledge about actions) while functional pragmatism regards knowledge as a basis for action (constructive knowledge). Methodological pragmatism deals with the creation of knowledge since pragmatism highlights the active role of the researcher in producing theories from data.

The view of Goles and Hirshheim (2000) is that pragmatism adopts a pluralist attitude and not taking a dogmatic (i.e. rigid) position concerning different methods. Pragmatist researchers make use of combinations of methods and methods that are suitable to the research purpose and current empirical situation. However, Goldkuhl (2012) emphasises that pragmatism implies pluralism since pragmatism adopts pluralist attitudes, but not all pluralism can be pragmatic. This is because pragmatism does not take a rigid position on using specific methods (Goldkuhl 2012), but the reverse is not the same.

Pragmatists also argue that research should not only aim to represent reality accurately but also to be useful and aim at how it is being utilised (Rorty 1999). This notion of utility has called for reflexive research practice (Feilzer 2010). Researchers should consider questions such as what is it for; who is it for; how do their values influence the study? Pragmatism aims to discover if the research has assisted to find out what the researcher wants to know (Hanson 2008b). Moreover pragmatists are not too concerned with the research methods used, as long as the method used can answer the question of what the researcher wants to know. However, Denscombe (2008) stresses that this is not a reason for poor research but needs a good understanding of the research methods that are transparent and replicable.

Pragmatism has been used in Design Research (DR) and Action Research (AR) in IS research by the following researchers: Hevner *et al.* (2004); Baskerville and Myers (2004); Cole *et al.* (2005); Lee and Nickerson (2010); and Mingers (2001) and, therefore, the IS community is becoming more aware of the pragmatism paradigm. Pragmatism has been suggested by some mixed methodologists as the best paradigm for validating using mixed methods research (Datta 1994; Howe 1988; Teddlie & Tashakkori 2003; Venkatesh *et al.* 2013). Pragmatism is the underlying philosophical paradigm used to carry out this research.

4.2.6 Why adopt Pragmatism?

Pragmatism focuses on the interests of both quantitative and qualitative researchers by indicating that all human inquiry comprises intention, values, imagination and interpretation but must also be based on pragmatic life experience (Yardley & Bishop 2008). Pragmatism is positioned toward solving practical problems in the real world (Feilzer 2010) and it is not based on the postulations about the nature of knowledge. Some authors (Patton 1990; Datta 1994; Howe 1988; Teddlie & Tashakkori 2003) have indicated that pragmatism is the best philosophical approach with specific research questions and issues in a mixed method research.

Pragmatism is seen as a philosophy of science embracing a multiplicity of methods and many method philosophies (Maxcy 2003). It allows many research projects to be carried out without identifying unchanged prior knowledges, rules or laws governing what is regarded as true or valid (Maxcy 2003). Hall (2012) has proposed three mixed methods positions for mixed methods researchers to support their research:

- a-paradigmatic stance;
- the multiple paradigm and

- the single paradigm stance. Evidence supports this stance has been most favoured by the majority of researchers since the research only adopts a single paradigm (e.g. pragmatism or transformative approach) to include the quantitative and qualitative research methods.

A research can be considered to be of value if it can be studied in different ways and the results can be used in ways that can enhance positive consequences within the researcher’s value system (Goles & Hirschheim 2000). IT research should appreciate practicality in its research by complementing theoretical research with thorough research that evaluates and describes the developments in IT practice (Goles & Hirschheim 2000). Pragmatism is the philosophical approach that provides the theoretical basis for research that combines IT theory and practice and this is line with the suggestion by Ormerod (1996) that pragmatism should be used to combine consulting and academic research in IS. This research may be regarded as one that combines both IT theory and practice paving way for the use of pragmatism in the study.

Pragmatism is a practical approach to solving a research problem and is strongly associated with mixed methods approach (Cameron 2011). Communication among researchers from different paradigms can be improved if a pragmatic and balanced or pluralist position is taken (Maxcy 2003). However, mixing research approaches should be done in the best opportunistic ways for answering important research questions. This is why pragmatism is a suitable paradigm underlying mixed methods research.

Research questions are very important to pragmatists before consideration is given to methods and paradigms that fit the research questions (Venkatesh *et al.* 2013). Research questions must be articulated firstly when using pragmatism. The pragmatic approach gives significance to the research question due to the mixing of quantitative and qualitative research (Bryman 2006b). This is also the view of Teddlie and Tashakkori (2003) that pragmatist researchers regard the research question to be more important than either the paradigm or the method used to carry out the research.

Research questions have been given the greatest value and importance by pragmatists. Pragmatists opt for a paradigm and method that are appropriate to answering the research questions (Venkatesh *et al.* 2013). This research addresses a practical problem and it is not based on the norms about the nature of knowledge. The main research question RQ, and the four sub research questions (SQ1 to SQ4) in section 1.2 are developed before the qualitative and quantitative research in this study. The pragmatist stand to develop the management framework on sensitive information migrations supports the decision to use a mixed methods approach as the best appropriate research method. The research questions of this study as well as the purpose of the research have been articulated during the commencement of this research.

Pragmatism presents a method for choosing methodological combinations that presents researchers to offer better solutions to many of their research questions (Johnson & Onwuegbuzie 2004). Pragmatists accept that there are philosophical differences between various paradigms but these philosophical assumptions can be merged and coordinated in conjuction with the methods chosen to carry out the research (Greene & Caracelli 2003). Researchers should be able to utilise what will work best for their study in terms of methodological decisions without philosophical paradigms' limitations (Patton 1988). Therefore, the most obvious underlying philosophical assumption to be chosen for this research is pragmatism since it uses both quantitative and qualitative methods in a mixed methods setting (cf. Venkatesh *et al.* 2013).

4.2.7 Limitations of Pragmatism

Pragmatism does not postulate that objects possess universal essences that define the object but allows the precise status of the object to be open and interpretively flexible. Researchers, however, have to avoid extreme subjectivity to erode the ability to explain the real world, most especially in relation to technical capacities or principles (Blosch 2001). Thus pragmatism leads to weak essentialism, however this does not suggest eternal unchanging essences, since the contexts of action may change and cause the essences to be redefined.

The problem for pragmatists is that a thing can be known in so many different perspectives in order to facilitate its understandings. The implication of this is that one thing can function in many different objects, and also many different things can represent one object (Blosch 2001). The interests, assumptions and the practical skills of the researcher determines the role that a thing is given in a situation. Therefore, the interactions between the object and the other objects determine the status of the object under the prevailing circumstances (Dewey 1960). Because of this, there are consequences for the understanding of technological artifacts (Blosch 2001).

Pragmatism does not recognise the organism/environment dualism and is therefore based on realism with a small ‘r’ (Putnam 1987). There is no knowledge on one side and real world on the other side, rather the real world is effected through knowledge and vice versa (Blosch 2001). Consequently, a theory is regarded to be true if it can forecast real-world problems.

According to Johnson and Onwuegbuzie (2004), the weaknesses of pragmatism include:

- There might be less focus on basic research than applied research because applied research may be seen to produce more practical and immediate results.
- Incremental change may be produced by pragmatism instead of more structural, fundamental, or revolutionary change in the society.
- Pragmatic researchers have failed to answer satisfactorily the question of who should benefit from pragmatic solution.
- The meaning of usefulness or workability can be unclear unless it is explicitly addressed by the researcher.
- It is difficult to deal with the useful and non-useful cases when using pragmatic theories of truth.
- Pragmatism has been rejected by many philosophers because of its logical failure to many philosophical disputes.

However, despite these limitations, pragmatism is still the best underlying philosophy for this research because the aim of this study is towards change, action and knowledge. Therefore, a pragmatist position will explain and justify the development of the management framework since the framework is aiming to improve what would be better in practice.

4.3 Research Methods

4.3.1 Induction, Deduction and Abduction Methods

There are three principles of reasoning in research and these are inductive, deductive and abductive (Johansson 2003). Deduction is regarded as the approach that makes the researcher to work with a plain theoretical framework (Rowlands 2003). In the induction approach, the researcher is not compelled by preceding theory but aims to develop relevant theory, propositions, and concepts in the research work. However, the process of the abductive approach is from rule to result to case (Danermark 2001). In abduction, the case does not present the logically necessary conclusion but rather a reasonable decision. Abduction is described by Johansson (2003) as ‘the process of facing an unexpected fact, applying some rule (known already or created for the occasion), and, as a result, positing a case that may be’. This implies that in abductive reasoning, behaviour is accounted for rather than predicted (Svennevig 2001).

Abduction was originated in 1866 by Charles Sanders Peirce who firstly called it ‘reasoning by hypothesis’ (Plutynski 2011). Induction extrapolates from one set of facts whereas abduction concludes from facts of one kind, to facts of another (Atkinson 2015). Induction and abduction reasoning differs from deductive reasoning because inductive and abductive reasoning make claims that do not follow logically from the premises (Atkinson 2015).

The abductive approach is different from induction and deduction in its research process (Kovacs & Spens 2005). Abductive reasoning interprets individual occurrences within an appropriate framework by using the perspective of a new conceptual framework to understand something in a new way (Danermark 2001). During theory building, the data is collected simultaneously in abductive reasoning and is similar to the methods of action research and case study research (Taylor *et al.* 2002).

Abduction is contrasted with deduction and induction by Plutynski (2011) in that abduction is considered as a way to study facts and then create a theory to explain the facts (Cunningham 1998), but deduction extract out the consequences of a rule and case (i.e. the ‘necessary reasoning’), whereas induction is a way of generating new theory emerging from data. Deductive research examines theory (for example in a literature review) and then obtains logical conclusions from this theory by using hypotheses and propositions to present them and then test them in an empirical setting so as to come up with general conclusions basing it on the proof or falsification of its self-generated hypotheses and propositions (Kovacs & Spens 2005). Inductive research is different from deduction research because the world is observed to obtain emerging suggestions and their simplifications in a theoretical setting (Kovacs & Spens 2005).

The focus of abduction is finding explanations for observed facts. These are propositions that are added to observed facts that make them appropriate in other situations than those observed (Peirce 1955). Peirce is also regarded as the founding father of pragmatism due to his abduction concept that serves a unique role in conceptualising pragmatism (Richardson & Kramer 2006). Therefore, abduction is the process of developing useful explanations which is essentially 'an inference' from observed facts and is an essential concept within pragmatism.

Philosophically, Johnson and Onwuegbuzie (2004) state that mixed methods research utilises the pragmatic approach and system of philosophy which includes induction, deduction and abduction. Venkatesh *et al.* (2013) point out that the pragmatic approach is grounded on abductive reasoning that interchanges between deduction (quantitative approach) and induction (qualitative approach). In this research, the abductive approach is used to synthesise specific suggestions in the literature as well as other observations to develop a general management framework.

4.3.2 Qualitative Research Method

Qualitative research is using the data (e.g. documents, interviews, and participant observation data) to explain and comprehend social phenomena (Sidi *et al.* 2009). They further highlighted the shift from technological research to research on managerial and organisational issues using qualitative research methods. Qualitative research is a process where a human or social problem is resolved in a natural setting by looking at the complete picture of the problem and reporting on it in its natural state (Creswell 1994, 2002).

Qualitative methods include interviews and questionnaires, documents and texts, observation and participant observation (fieldwork) and the researcher’s limitations and responses (Myers & Newman 2006). Methodologies used to carry out qualitative research include Ethnography, Action Research, Case Study Research, and Grounded Theory (Myers & Newman 2006). Qualitative research is used to explore meanings of social phenomena experienced by individuals themselves in their natural setting (Malterud 2001). Qualitative research is built on inductive reasoning rather than deductive reasoning (Williams 2007).

Qualitative research is regarded as a research method that can be positivistic or interpretive depending on the researcher’s philosophical expectations (Cavaye 1996; Oates 2006). Qualitative research is used as one of the mixed methods approaches in this research in a pragmatist setting as explained in Chapter 6. It is used to validate the results obtained in the quantitative analysis. Johnson and Onwuegbuzie (2004) maintain that researchers must first consider all the characteristics of the qualitative research method when using mixed methods approach, hence the reason for exploring qualitative research method in this section. Qualitative research method is used in this study as one of approaches in the mixed methods research.

4.3.3 Quantitative Research Method

Quantitative research methods were created originally to understand natural phenomena in the natural sciences while focusing on objective measures (Sidi *et al.* 2009). It emerged in approximately 1250 A.D. due to the quest for data quantification (Williams 2007). It is research employing measurement procedures and other techniques to understand truly quantitative attributes (Michell 2011; Westerman & Yanchar 2011). Quantitative methods include numerical methods (e.g. mathematical modelling), laboratory experiments and formal methods (Leedy & Ormrod 2012).

It is always presumed that quantitative research is associated with positivism and qualitative methods with interpretivism (Johari 2009). The quantitative paradigm ontological position is that there is only one truth, and the researcher and what is being investigated are independent entities epistemologically (Sale *et al.* 2002). That is, the research is autonomous of the researcher and data is used to determine truth objectively. This view no longer holds as some authors (Klein & Myers 1999; Weterman & Yanchar 2011) have given examples of quantitative-interpretive research in their work.

Quantitative research is good for theory testing and developing universal statements since a general picture of the situation can be provided (Schulze 2003). A quantitative research method is an efficient method that can be used to gather data from many respondents. However, it does not include an interpretation of the reasoning behind the questions, thus leaving the researcher to interprete the results of the quantitative study.

Quantitative research commences with a statement of the problem to be investigated, then a literature review, and a quantitative data analysis (statistical analysis). It has origins in the physical sciences, notably physics and chemistry (Creswell 2002). There are three quantitative research categories: (a) descriptive; (b) experimental; and (c) causal comparative (Leedy & Ormrod 2012). Research methods used to conduct quantitative research include: developmental design, survey research, correlational and observational studies (Williams 2007). The results from quantitative research might be confirming, explanatory and predictive.

The descriptive research examines the current situation as it exists and involves identifying the attributes of an occurrence based on an examination. The experimental research involves the mathematical models in the data analysis and a systematic approach to data collection. The researcher examines how the independent variables are altered by the dependent variables in the fundamental comparative research and it also involves the origin and consequences of the relationships between the variables (Leedy & Ormrod 2012). Johnson and Onwuegbuzie (2004) maintain that researchers must first consider all the characteristics of the quantitative research method when using mixed methods approach, hence the reason for exploring quantitative research method in this section. Quantitative research method is used in this study as one of approaches in the mixed methods research.

IS research is classified by Orlikowski and Baroudi (1991) as positivist when there exist hypothesis testing, quantifiable measures of variables, formal propositions and the drawing of inferences. In IS research, there is a strong positivist bias in analysing and understanding systems (Falconer & Mackay 2000). This view is supported by Orlikowski and Baroudi (1991) in their research findings that 97 per cent of American IS research used a positivist approach. Roode (2003) supports this view by acknowledging that the pioneers of IS research came from disciplines such as physics and mathematics, who had been doing research by using a quantitative method.

4.3.4 Mixed Methods Research

Some authors (Lee 1999; Robey 1996; Sidorova *et al.* 2008) consider diversity to be a major strength of Information Systems. Many IS researchers are advocating using different methods and approaches for conducting IS research (Stockdale & Standing 2006). This has led to researchers combining qualitative and quantitative research in a single research project (Bryman 2006a). Mixed methods research uses many research methods or more than one worldview in a research work (Tashakkori & Teddlie 2003a; 2003b). Mixed methods research is the combination of quantitative and qualitative methods and exploits the respective strengths of quantitative and qualitative approaches (Ostlund *et al.* 2011; Peng *et al.* 2011). This integration of quantitative and qualitative approaches has continued to be an interesting issue (Morgan 2007; Onwuegbuzie & Leech 2005). There are many definitions of mixed methods research in the literature and Table 4-1 contains some selected definitions of mixed methods research by different authors.

TABLE 4-1
Definitions of Mixed Methods Research by Different Authors

Authors	**Definitions of mixed methods research by each author**
Johnson and Onwuegbuzie (2004)	Mixed methods approach is a type of research where the researcher uses both quantitative and qualitative techniques, approaches, concepts, methods or language in a single study.
Venkatesh *et al.* (2013)	Mixed methods research uses both the qualitative and quantitative methods within the same study.
Tashakkori and Teddlie (2003a, 2003b)	Mixed methods approach is a research method involving more than one method, both quantitative or qualitative research approach.
Mingers (2001, 2003)	Mixed methods approach uses more than a single research method within a single study which can include two quantitative methods or two qualitative methods.
Sawyer (2000, 2001)	Mixed methods approach uses many data collection methods to obtain multiple data results.

Tashakkori and Creswell (2007)	In a mixed methods approach, the researcher collects, analyses data and then uses both qualitative and quantitative approaches to generate conclusions.
Johnson *et al*. (2007)	A mixed methods approach comprises using the combination of qualitative and quantitative elements (e.g. use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) by the researcher to carry out the study.

In Table 4-1, it is clear that the qualitative or quantitative techniques do not replace each other in mixed methods research, instead, they complement or integrate with each other to give better results from the complete research process. Moreover, almost all the authors with the exception of Sawyer (2000, 2001) define mixed methods research as an approach that uses both the quantitative and qualitative methods within the same study. However, Sawyer (2000, 2001) still recognises the fact that mixed methods approach uses many data collection methods to obtain multiple results within a single study.

In this study, the researcher restricts the definition of mixed methods research to include research combining qualitative and quantitative methods in a single research inquiry to gather, analyse and describe both types of data.

Mixed methods approach is different from multi-methods approach. The term ‘mixed methods’ is used as a reference to using two or more methods in a research project resulting in quantitative and qualitative data (Teddlie & Tashakkori 2009; Creswell & Clark 2007; Greene 2007). The term ‘multi-methods’ is used to refer to the use of amalgamations of methods which yield similar data (Teddlie & Tashakkori 2009). The difference between multi-methods research and mixed methods research is that in multi-methods research, the researcher can use either two quantitative methods or two qualitative methods whereas in mixed methods research, the researcher has to use both qualitative and quantitative methods within a single study (Teddlie & Tashakkori 2003, 2009; Rocco *et al.* 2003; Petter & Gallivan 2004; Fidel 2008).

It was Jick (1979) who first introduced the concept of mixed methods research for merging qualitative and quantitative methods in social science research (Ostlund *et al.* 2011). Methodological pluralism (i.e. multiple methods) in the IS literature has not been employed within IS research (Mingers 2001, 2003). However, a mixed methods research approach has been employed recently in IS research (Mingers 2001, 2003; Weber 2004; Lee & Hubona 2009). Mixed methods research is described by some authors (Teddlie & Tashakkori 2003, 2009; Ridenour & Newman 2008) as the third methodological paradigm after quantitative and qualitative methods that represent the first and the second paradigms respectively.

There have been requests for using mixed methods research, but it has not gained much ground in IS research, where about five per cent of IS research studies published in AIS Journals between 2001 and 2007 have made use of mixed methods research (Venkatesh *et al.* 2013). There are no guidelines in the IS literature for performing and assessing mixed methods research in IS (Venkatesh *et al.* 2013). However, guidelines to conduct mixed-methods research in IS has been given by Venkatesh *et al.* (2013).

Mixed methods research uses multiple methods (quantitative and qualitative research approaches) in a research setting (Venkatesh *et al.* 2013). Although there is much confusion in social sciences research on how mixed methods research should be embarked on, in IS research this debate regarding mixed methods research deployment appears to be largely resolved (Petter & Gallivan 2004). Mixed methods research has been used by many researchers since the concept was first introduced (Bloch *et al.* 2014; Ostlund *et al.* 2011; Mingers 2003; Johnson & Onwuegbuzie 2004; Gulati & Taneja 2013; Lee *et al.* 2013).

Venkatesh *et al.* (2013) have proposed that the decision of conducting mixed methods research relies on the research question, reason and setting. Mixed methods research utilises quantitative and qualitative research methods either concurrently or sequentially (Mingers 2001; Mingers 2003; Sawyer 2000; Sawyer 2001; Petter & Gallivan 2004; Venkatesh *et al.* 2013). An example is the use of interviews (a qualitative data collection approach) and questionnaires (as a quantitative data collection method) to gather data about an IS research. This is the approach used in this study. It should be noted that questionnaires can also be used in qualitative research.

IS has accepted mixing the two paradigms in mixed methods research (Gallivan 1997; Kaplan & Duchon 1988; Lee 1999; Mingers 2001; Sawyer 2001; Venkatesh *et al.* 2013). The combination of qualitative and quantitative research in a mixed methods setting can be done at different stages of the research process, e.g. during the formulation of the research questions, sampling of data, data collection and data analysis (Bryman 2006a).

There are four types of mixed methods research as suggested by Creswell and Clark (2007):

- triangulation – combining qualitative and quantitative data,
- embedded – the use of both qualitative or quantitative data to solve a research question, where one form of data is entrenched within the other,
- explanatory – the collection and analysis of quantitative data followed by the collection and analysis of qualitative data,
- exploratory – the collection and analysis of qualitative data followed by the collection and analysis of quantitative data.

Mixed methods research have three major strengths (Venkatesh *et al.* 2013). Mixed methods research can:

- address confirmatory and exploratory research questions simultaneously,
- provide stronger inferences (meta-inferences) than a single method,
- can provide for a greater range of different and/or corresponding views.

Venkatesh *et al.* (2013) review mixed methods research articles in IS from 2001 to 2007 and find that development and completeness are the most dominant purposes for performing mixed methods research in IS (32% and 26% respectively). They find that diversity (3%) and compensation (3%) are the least occurring reasons for conducting mixed methods research in IS. They find that expansion, corroboration/confirmation and complementarity are the remaining purposes for doing mixed methods research with 36%. They also find that 55% of the mixed methods research dominant method used is quantitative while 45% is qualitative. Moreover, they find that 65% of the articles discussed meta-inferences (quantitative and qualitative integrative findings). They suggest that IS researchers adopting mixed methods in their work can integrate both qualitative and quantitative inferences (meta-inferences) so that substantive theory can be discovered.

Bryman (2006a) performs a study on an investigation of articles combining quantitative and qualitative research published between 1994 and 2003, and he finds that qualitative interviews and quantitative survey methods are the vast majority of approaches employed in the articles. In his study, 82.4% of all articles coded used a survey instrument while 57.3% of all articles are a grouping of a survey instrument and qualitative interviewing. The data collected by either a structured interview or a questionnaire predominates on the quantitative side while the data collected by either a semi-structured or unstructured interview predominates on the qualitative side.

There are two design strategies for conducting mixed methods research: (a) concurrent design – which involves the gathering and analysis of quantitative and qualitative data in parallel and then the merging or comparing the results, (b) sequential design – which involves the collection and analysis of quantitative and qualitative data in different phases and then the integration in a separate phase (Creswell *et al*. 2003; Tashakkori & Teddlie 1998). Table 4-2 illustrates the six types of mixed-method designs by Creswell (2002). In this thesis the sequential explanatory design is used as the type of the mixed methods approach.

TABLE 4-2
Six Types of Mixed-methods Designs (source: Creswell 2005)

Mixed-methods design	**Characteristics**
A. **Sequential designs:**	
Sequential explanatory design	This design comprises of two phases and it involves the gathering and analysis of quantitative data which is then followed by the collection and analysis of qualitative data. Priority is given to the quantitative part. Qualitative results are used to further explore and explain the results of a primarily quantitative study.
Sequential exploratory design	This design begins with the qualitative data collection and analysis phase. This qualitative phase is then followed by a quantitative data collection and analysis phase with the aim of increasing generalisability of the findings. The priority is given to the qualitative aspect of the study. The results of the quantitative analysis are used to inform the follow-up qualitative data collection.

Sequential transformative design	This design includes two different data collection phases and any of the two methods may be utilised first when collecting data. The priority can be given to either the quantitative or the qualitative phase, or even to both depending on the availability of resources. The researcher builds the research within a transformative theoretical perspective. Qualitative findings are made to provide an enhanced understanding of the quantitative findings in order to explore inequalities.
B. Concurrent designs	
Concurrent triangulation design	The quantitative and qualitative methods are used simultaneously in one phase in this design. The aim is to cross-authenticate, check or corrobate results within a single study. Both the quantitative and qualitative methods are considered to be equally important.
Concurrent nested design	This design comprises of only one data gathering phase, during which both qualitative and quantitative data are simultaneously collected. However, one method (either qualitative or quantitative) must take the predominant position, and the other method should be entrenched within the predominant method to seek information in a different level or to address a different question.
Concurrent transformative design	The features of both concurrent triangulation and concurrent nested designs are combined in this design. This may involve a combination of qualitative and quantitative components that are equally important. It is then also entrenched with a supplement method to further explore the issue. However, all data are collected during the same period in one data gathering phase.

Table 4-2 illustrates the differences between the sequential and concurrent designs of the mixed methods research. Sequential design comprises of three types: sequential explanatory design; sequential exploratory design and sequential transformative design. The sequential explanatory design commences with the quantitative data gathering and analysis followed by the qualitative data study (Clark *et al.* 2010). Qualitative study results are used to further investigate and clarify the results of a primarily quantitative study. The sequential explanatory design is used in this research.

The sequential exploratory design begins with the qualitative study and then followed by the quantitative analysis with the aim of increasing the generalisability of the qualitative findings. The researcher generally stresses the importance of the qualitative method because the design commences with this part (Clark *et al.* 2010).

Concurrent design also comprises of three types: concurrent triangulation design; concurrent nested design and concurrent transformative design. Both qualitative and quantitative methods are used simultaneously in the concurrent triangulation design. Concurrent nested design involves using both methods to gather data, however, one method must be predominant. The features of both concurrent triangulation design and concurrent nested design are combined in the concurrent transformative design (Cresswell 2005).

Mixed methods have been used by many researchers (Bloch *et al.* 2014; Mingers 2003; Johnson & Onwuegbuzie 2004; Kaplan & Duchon 1988; Trauth & Jessup 2000; Markus 1994; Ostlund *et al.* 2011; Gulati & Taneja 2013, Lee *et al.* 2013). However, mixed methods research has not been used frequently in the IS field despite strong and continuous support from IS researchers (Mingers 2003; Fidel 2008; Peng *et al.* 2011). Harrison III (2013) has pointed out that guidance is needed when performing mixed methods research and for evaluating the thoroughness of data gathering and analysis of both types of data in mixed method research. The sequential explanatory mixed methods approach is used to carry out this research since it lends itself to stronger interpretations based on the results (Clark *et al*. 2010; Schulenberg 2007).

In a survey carried out by Chen and Hirschheim (2004) in which they examined 1893 papers published in the American or European journals between 1991 and 2001 revealed that 71% of US research used quantitative methods while 49% of European journals used qualitative methods at the methodological level. Azorin and Cameron (2010) indicate that in the articles published in the Strategic Management Journal between 2003 and 2009, 73.3% used quantitative methods, 10.6% used non-empirical methods and 12.9% used mixed methods approach while 3.2% used qualitative methods. This implies that there is a paradigm shift towards the use of mixed methods approach in IS research. The sequential explanatory mixed methods case study approach is used to carry out this research.

4.3.5 Comparison of the three Research Methods

Table 4-3 depicts the comparison of the three research methods by different authors.

TABLE 4-3
Comparison of the Three Research Methods by Different Authors

Characteristics	**Mixed Methods Research**	**Quantitative Research**	**Qualitative Research**	**Authors**
Major characteristics	It is an extensive and innovative research method that is diverse, inclusive and complement ary.	It centres on validation, clarification, inference, theory/hypo-thesis assessment, projection, statistical analysis and consistent data collection.	It centers on investigation, theory/hypo-thesis creation, induction, discovery, qualitative analysis and the researcher as the primary channel of data collection.	Johnson & Onwuegbuzie (2004)

Strengths	Numbers are used to add precision to words, while pictures and narrative are used to give meaning to numbers; pictures and narrative provides quantitative and qualitative research rigour; researchers are able to produce and examine a grounded theory; can resolve many research questions because the researcher is not limited to one method or approach; can provide strong proof for a deduction through combination and validation of findings.	Data collection is relatively quick; provides detailed, numerical, quantitative data; data analysis does not consume much time by using statistical software; research results are relatively unbiased of the researcher; useful for studying large numbers of people.	Data are grounded on the researcher's own categories of sense; It can be utilized for studying a partial number of cases in depth; can perform cross-case evaluations and analysis; offers awareness and explanation of people's personal occurrences; makes use of grounded theory to generate inductively an exploratory theory about an occurrence; Data are usually collected in realistic situations.	Johnson & Onwuegbuzie (2004)

Weaknesses	It can be challenging for a researcher to perform both qualitative and quantitative research concurrently and this may require using a team of researchers; the researcher has to be educated about using multiple methods and approaches (steep learning curve); it can be more expensive; it is more time consuming.	Researcher's categories as well as researcher's theories may not affect local regions' perceptions; researcher may omit the occurrences because of the emphasis on hypothesis examination or theory rather than on hypothesis creation or theory; the knowledge may be too general and conceptual.	Knowledge produced may not take a broad view of other people or other situations; it is difficult to predict quantitatively; it is difficult to examine theories and hypotheses; it takes more time to collect the data relatively; it's time consuming to perform data analysis; the researcher's personal biases and idiosyncrasies can easily influence the results.	Johnson & Onwuegbuzie (2004)

Validation in the research method	It assesses the quality of results and/or interpretation from all the quantitative and qualitative data in the research study; it provides a clear analysis and evaluation of how results are integrated from both qualitative and quantitative studies and the quality of this combination.	It recognises the importance of reliability and validity; uses measurement validity, design validity and inferential validity; it has generally recognized and undoubted rules for validation	It uses design validity, analytical validity and inferential validity; does not have rules or assessment conditions for validation that are generally agreed for validation and/or widely used; validation is ambiguous and contentious.	Venkatesh *et al.* (2013) Nunnally & Bernstein (1994) Lee & Hubona (2009) Ridenour & Newman (2008) Whittemore *et al.* (2001)

Methodologies	Observation. survey, experiment, simulation, case study, interview, action research, grounded theory, content analysis, participant observation, critical theory, ethnography consultancy	(Passive) Observation, measurement and statistical analysis; survey, questionnaire, or instrument; experiments; simulation; case study	Qualitative content analysis, interviews; ethnography/ hermeneutics; participant observation; grounded theory	Mingers (2003)
Research reasoning	It leans towards abduction reasoning	It leans towards deductive reasoning	It leans towards inductive reasoning	Kovac & Spens (2005) Johansson (2003)

In Table 4-3, the three research methods namely mixed methods research; quantitative research and qualitative research are compared with respect to major characteristics, strengths, weaknesses, validation, methodologies and research reasoning. The strength of the mixed methods research lies in the use of numbers to add precision to words, while pictures and narratives are used to give meaning to numbers; pictures thus providing quantitative and qualitative research rigour. The strength of the quantitative research lies in its high reliability due to the rigorous data collection and critical analysis while the strength of qualitative research is its ability to investigate for underlying values, assumptions and beliefs thus making the inquiry to be broad and open-ended while allowing participants to raise issues that matter most to them.

The weakness of the mixed methods research lies in the time that the researcher has to be educated about using multiple methods and approaches (steep learning curve); thus making it more expensive and more time consuming while the weakness of the quantitative research is that the knowledge may be too general and conceptual. The weakness of the qualitative research is that it takes more time to collect the data relatively and it is more time consuming to perform data analysis. The researcher's personal biases and idiosyncrasies can easily influence the results. The table indicates that mixed methods approach combines the strengths, methodologies and validation techniques of both the qualitative and quantitative methods. This makes this approach an attractive option for a study where relevance and rigour should carry equal weight. The next paragraph will motivate why mixed methods is the chosen approach for this thesis.

4.3.6 Why Mixed Methods Research?

The present research world is one that is complex, inter-disciplinary and dynamic, therefore, researchers require understanding of multiple methods to provide superior research, facilitate communication and collaboration (Johnson & Onwuegbuzie 2004). Mixed methods research enables the demonstration of divergent views to a larger range since different views can cause a re-examination of the framework that is developed (Teddlie & Tashakkori 2003). Mixed methods research is to gain from the strengths and reduce the weakness of both in research studies but it is not a replacement for either quantitative or qualitative research approach (Johnson & Onwuegbuzie 2004). Mixed methods research offers triangulation by allowing for greater validity in a study using both quantitative and qualitative data (Doyle *et al.* 2009). This is one of the reasons why mixed methods approach is used to carry out this research.

Researchers use mixed methods approach in their studies because a better understanding of the problem can be obtained (Clark *et al.* 2010). The complementary strength of mixed methods research obtained from quantitative and qualitative methods is the major reason of using mixed methods research in this study. Mixed methods approach can increase the building of knowledge most especially if the two methods are applied in a sequence (Clark *et al.* 2010). Mixed methods approach is used in this research because of its complementary strength from both quantitative and qualitative methods.

Quantitative and qualitative tools like questionnaires and surveys are an efficient and economical way for collecting data from many respondents (Bryman 2004). An interview questionnaire is a qualitative tool used to efficiently gather and investigate in-depth human insights and views on complex social phenomena (Saunders *et al.* 2003; Bryman 2004). The use of both quantitative and qualitative tools in a study improves the strength of the research and provide more complete and comprehensive features of the research (Doyle *et al.* 2009). Therefore mixed methods emerged to resolve the limitations of quantitative and qualitative approaches (Peng *et al.* 2011). Mixed methods research gives more convincing explanations of the study because of the combination of the statistical results with the qualitative narratives and a broader audience may be interested in the study (Clark *et al.* 2010). Mixed methods approach is well accepted by the practitioners because it supports the ways that problems are resolved in practise.

Mixed methods integrates a level of flexibility that results to an emergent design by enabling researchers to build on initial findings without changing the overall design (e.g. the conducting of qualitative interviews to explain specific quantitative results) (Clark *et al.* 2010). In this research, the researcher qualitatively observed, interviewed and supplemented the quantitative analysis using interviews to validate the rudimentary framework constructs developed from literature. Triangulating one set of results with another leads to better understanding and enhancement of the validity of inferences (Azorin & Cameron 2010). The results can be said to be valid, if several different methods for investigating a phenomenon of interest are used and the results offer mutual confirmation (Niglas 2004). Mixed methods research also increases the methodological rigour of a study because multiple forms of validity/validation are used (Clark *et al.* 2010). This is one of the reasons why mixed methods is used in this research.

The mixed methods approach also allows the researcher to simultaneously answer affirmative and investigative questions and then generate and validate theory in the same research study (Teddlie & Tashakkori 2003). Mixed methods research assists to answer research questions that quantitative and qualitative methods alone cannot answer (Creswell & Clark 2007). According to Johnson and Onwuegbuzie (2004), the most fundamental aspect in a mixed methods research study is the research question and the researcher has to understand research questions so that useful answers can be obtained by considering all of the pertinent features of traditional quantitative and qualitative research. Moreover, the choice of using a mixed methods research design is based on the purpose of the research as well as the research questions (Hall 2012). Mixed methods researchers suggest that the research question within each stage of the research circle is more important than the method used or the paradigm underlying the method (Teddlie & Tashakkori 2003).

The aim of this research is to develop a framework to manage sensitive information between migration of platforms. As Venkatesh *et al.* (2013) have pointed out, illustrating meta-inferences is a significant and vital characteristic of mixed methods research and this process is conceptually similar to theory development from observations where they are regarded as the results from the qualitative and quantitative analyses. The framework that is developed in this research work can be regarded as the basic theory of the phenomenon of interest which is obtained from the merged meta-inferences obtained from the findings of the qualitative and quantitative analyses. The quantitative analysis of the research is used to develop the structure of the framework while the qualitative analysis is used to validate the framework, hence the importance of meta-inferencing in this study.

In this research, some questions are exploratory while others are assenting, that is, some questions are used to develop theory inductively while others are used to confirm the framework using a deductive approach, therefore a mixed methods approach is very suitable in this research. A mixed methods approach lends itself to stronger interpretations based on the results (Schulenberg 2007). The framework on information sensitivity is developed from both quantitative data and qualitative data, thereby allowing for stronger inferences since the qualitative phase changes and amends the quantitative findings. The researcher ensures that all bias and preconceptions are eliminated in this work by not being emotionally involved with or have a particular attitude toward the research and also moved beyond common-sense beliefs. Mixed methods approach is used in this research due to the above-mentioned reasons.

4.3.7 Limitations of Mixed Methods Approach

There are challenges in using mixed methods research: (a) mixed methods can be influenced by social factors (Petter & Gallivan 2004), (b) interpretation of mixed methods data, because the analyses of the methods may oppose each other. In the second case, Jick (1979) has pointed out that the researchers must resolve the differences by looking for the reasons causing the inconsistent data. Venkatesh *et al.* (2013) emphasise that mixed methods research should not replace thoroughly performed single methods studies in IS but rather should be used as an additional approach to gain knowledge on phenomena of interest to IS researchers.

Mixed methods approach can be compromised/weakened if both quantitative and qualitative approaches are not used properly or not designed well (Trotter II 2012).

Designing and implementing mixed methods research can be difficult in actual practices (Ivankova *et al.* 2006; Fidel 2008). Therefore, there must be careful consideration on how researchers decide on how the quantitative and qualitative methods should be combined in a study and in what order (e.g. concurrently or sequentially) as well as the priority of these methods and the objectives that each method will achieve (Creswell 2009; Fidel 2008; Ivankova *et al.* 2006). These decisions have to be made in response to the predefined research questions and research setting (Peng *et al.* 2011). Thus, making inappropriate decisions too will affect negatively the rigorousness and reliability of the mixed methods research design (Peng *et al.* 2011). This will also weaken/undermine the richness and significance of the resulting research findings.

It might be difficult for inexperienced researchers to mix the methods in one single study because the different approaches are different in terms of underlying epistemologies, data analysis techniques and data collection methods (Bryman 2007; Peng *et al.* 2011; Small 2011). Senior researchers might have preference for using one approach and may not want to use the other approach due to lack of skills to use the other approach. Additionally, a team of researchers using mixed methods research might make the situation complicated because each of them might have a strong position on their own single approach and might be unwilling to use the other approach (Patton 2002; Fidel 2008; Bryman 2007). Therefore mixed methods approach requires that researchers have a broader skills set that includes both the quantitative and the qualitative methods (Azorin & Cameron 2010).

There are challenges when conducting mixed methods research because more work, more time and financial resources are needed (Creswell *et al.* 2007; Niglas 2004). Mixed methods researchers have challenges of publishing their studies due to the limitation of word and page in journals (Clark 2005; Bryman 2007). According to Collins *et al.* (2007), there are four challenges in mixed methods research: (i) the challenge of representing the lived experience using words and numbers; (ii) the challenge of legitimation (validity) – which refers to the obstacle in getting findings and/or making transferable, credible, trustworthy, dependable and confirmable inferences; (iii) integration of the quantitative and qualitative approaches challenges and (iv) the challenge of politics – tensions resulting from the combination of the quantitative and qualitative approaches. Despite all these challenges, the mixed methods approach offers the best approach for this research by gaining the strengths and reducing the weakness of both the quantitative and qualitative approaches, allowing for greater validity, and offering completeness of the study.

4.3.8 Validation in Mixed Methods Research

Validity in research refers to the extent that a research answers the study question or the strength of the research conclusions accurately (Sullivan 2011). Hanson (2008a) regards validity as the connection between method and theory which is the vital criteria for assessing the legitimacy of a method. Validity is the truthfulness of the findings. Venkatesh *et al.* (2013) highlight the importance of validation in mixed methods research. There are many types of validity, however, the three major categories are: (a) content validity; (b) criterion-related validity and (c) construct validity (Long & Johnson 2000).

Content validity refers to the degree that a test measures the content domain knowledge. It refers to the extent to which the phenomenon under investigation is addressed in its entirety (Long & Johnson 2000). It shows how the test items adequately and representatively sample the test content area to be measured. Content validity is determined by using expert judgement (not statistics).

Criterion-related validity is the degree of a test’s usefulness for predicting a person’s behaviour in a specified situation. It involves comparing the instrument and findings with an established standard in order to establish the correlation between measured performance and actual performance (Long & Johnson 2000). It is otherwise known as predictive validity.

Construct validity is the extent to which a test measures a non-observable construct, theoretical or an intended hypothethical construct. It is the process of validating the analysis about abstract attributes or constructs and it is worked over a time period based on accumulation of evidence. Factor analysis is used to establish construct validity in this study.

Validation is a major concern in mixed methods research with the the term data quality referring to reliability while the term inference quality refers to validity in mixed methods research (Venkatesh *et al.* 2013; Luyt 2012; Coaley 2010). IS researchers need to check for threats to validity during the data gathering and analysis in both quantitative and qualitative components of mixed methods research (Venkatesh *et al.* 2013). This should be done in order to improve the validity of their research.

IS researchers are encouraged to discuss explicitly validation for the mixed methods part of their work (Venkatesh *et al.* 2013). IS researchers should also review the potential threats to validity during data gathering and analysis for both quantitative and qualitative components of their mixed methods research. A researcher needs to focus on the theoretical drive and be methodologically coherent in order to avoid serious threats to validity in mixed methods research (Morse *et al.* 2006). They emphasise that the entire research design needs to be considered in order to maintain validity.

Creswell and Clark (2007) has raised the following issues concerning validation in mixed methods research: (a) How validity should be conceptualised in mixed methods research, (b) The period and how to report and discuss quantitative and qualitative validity aspects of mixed methods research, (c) The possibility of researchers following the traditional validity guidelines and expectations, (d) How potential threats to the validity can be reduced during data collection and analysis phases in mixed methods research.

Some authors (Bamberger 2007; Creswell & Clark 2007) have argued that validation in mixed methods research should include validations in quantitative and qualitative research. However, Teddlie and Tashakkori (2003; 2009) stress that validity in mixed methods research should be given another new terminology – inference quality because validity has lost its meaning. This will assist in differentiating between mixed methods validation from qualitative and quantitative validation (Venkatesh et al. 2013). Inference quality is the accuracy of conclusions that are inductively or deductively derived in a research and it refers to validity (Teddlie & Tashakkori 2003). They suggest that inference quality comprises of design quality and interpretive rigor (or explanation quality). Design quality checks if a mixed methods research abide by commonly accepted best practices while interpretive rigor are the standards used to evaluate the accuracy of the conclusion.

The validation of the qualitative and quantitative analysis of this research is discussed in sections 5.11 and 6.2 respectively. The qualitative validation includes credibility (internal validity), confirmability and dependability. Credibility, confirmability and dependability are all confirmed in this study. The quantitative validation includes construct validity (Exploratory Factor Analysis and Spearman's correlations), reliability (Cronbach Alpha) and internal validity (triangulation of results).

4.3.9 The Mixed Methods Approach Processes Followed in this Research

The sequential explanatory mixed methods design was used to carry out this research. Figure 4-1 is the visual model of the mixed methods design: sequential explanatory design procedures developed by Ivankova *et al.* (2006). The quantitative analysis was performed first and this was followed by the qualitative analysis. The qualitative analysis was performed in order to serve as a confirmatory and completeness process for the results of the quantitative study. Qualitative results were used to further explore and explain the results of the primarily quantitative study.

The relationship between qualitative and quantitative approaches is an important issue in the design of a mixed methods approach or which approach should be given priority (Kelle 2005). This triangulation of approaches necessitates that the different approaches are brought into line to validate results. The implication of this is that qualitative approach is used as a surbodinate to quantitative approach, with the main aim of validating and illustrating quantitative results (Bloch *et al.* 2014). The qualitative analysis was used to validate and explain the results of the quantitative analysis in the research study.

(a) The quantitative analysis design process was carried out using descriptive statistics, Spearman’s correlations and exploratory factor analysis. The measuring instrument (Appendix B) was developed from an in-depth analysis of the rudimentary management framework (Figure 2-1) and the security challenges model during OSS migrations (Figure 3-1). A pilot study was carried out to improve the measuring instruments. Sampling was done by focusing and obtaining data from six government organisations in Pretoria.

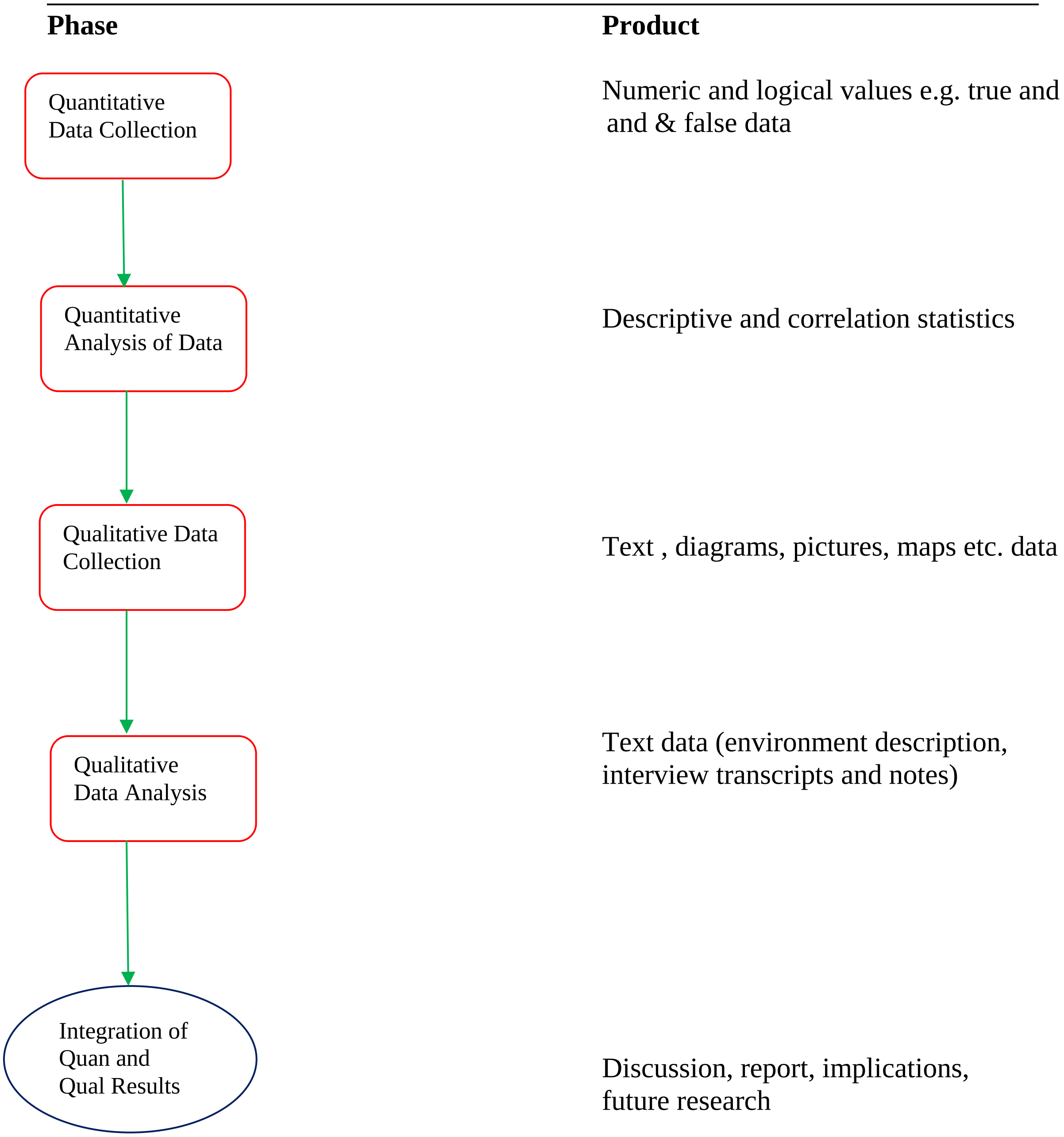

Phase
Product
Quantitative Data Collection
Numeric and logical values e.g. true and and & false data
Quantitative Analysis of Data
Descriptive and correlation statistics
Qualitative Data Collection
Text , diagrams, pictures, maps etc. data
Qualitative Data Analysis
Text data (environment description, interview transcripts and notes)
Integration of Quan and Qual Results
Discussion, report, implications, future research

Figure 4-1: Visual Model of Sequential Explanatory Design Procedures (from Ivankova *et al.* 2006)

Two hundred and fifty questionnaires were distributed to IT staff members in the six government organisations mentioned in section 1.1.1. Ninety respondents completed their questionnaires, giving a response rate of 36%. Measuring instruments were designed and developed as shown in Appendix B. Exploratory Factor Analysis (EFA) was used to test the validity of the constructs. Item analysis was used to test the reliability of the constructs in the measuring instruments. The constructs were then subjected to correlational analysis to explore the prelimary management framework. The resulting preliminary management framework was validated using qualitative analysis.

(b) The qualitative analysis design process was carried out after the completion of the quantitative analysis design. This is used to explore further the quantitative findings and also to achieve triangulation. Ten participants from six South African government organisations were selected and interviewed using purposive sampling. The main questions asked each participant were shown in Appendix C and were used to validate the results obtained in the quantitative analysis. The researcher used gatekeepers to obtain permission to interview participants and the letter requesting permission from gatekeepers of participating organisations is shown in Appendix D(a).

The interviews were transcribed and the resulting transcripts were imported into the NVIVO software for coding analysis. The coding analysis resulted in the generation of categories, sub-categories, and sub sub-categories which are used to develop themes and then combined with the preliminary management framework to form the final management framework.

The researcher avoided the following ethical considerations: bias; inappropriate research methodology; invalid reporting; and using information inappropriately. Validity of the qualitative study was ensured because the researcher obtained data from different sources (data triangulation).

4.4 Research Methodology

This section focuses on the research methodology used in the study. The difference between a research method and a research methodology is that the research method refers to the processes used to carry out the research methodology (Oates 2006).

4.4.1 Case Study Research

A case study is used to produce a comprehensive understanding of multifaceted issues in its real-life setting (Crowe *et al.* 2011). The origin of case study is from human and social sciences including evaluative research (Creswell 2007). Case study is a research strategy that allows researchers to understand the dynamics that are within single settings (Johari 2009). A case study can be positivist, interpretive or critical depending on the researcher’s philosophical stance (Yin 2009; Walsham 1993; Jarvensivu & Tornroos 2010).

The case study research method is one of the most common qualitative research methods used in information systems (Myers 2005). Case studies investigate problems in their natural settings, especially when there is no distinct boundary between the phenomenon and the context being investigated (Yin 2003). Yin (2003) also reiterates that case studies can be quantitative or qualitative and it involves the use of ‘how’ and ‘why’ types of research questions.

Case study research can utilise a single case or many case designs (Yin 2003). According to Yin (2003), single (lone) case studies are appropriate if: (a) it is a condition not previously accessible to scientific inquiry; (b) it characterises a precarious case to test a well-articulated theory and (c) it is an extreme or unique situation. By contrast, multiple (many) case designs are used when the research involves explanation, theory testing or theory building.

Case study research is not associated with one underlying research philosophy, but rather it can be used in research that has an underlying philosophy of positivism, interpretivism or critical thinking (Oates 2006). Case studies are conducted in real world situations and therefore, have a high degree of realism (Runeson & Host 2009). Case study research has been used in research into the design, implementation and usage of information systems since it allows for studying managers' behaviour, users' perceptions, developers, technology, legislation, group dynamics, power and politics (Oates 2006).

Data collected in a case study can be quantitative or qualitative (Runeson & Host 2009). Quantitative data is analysed using statistics while qualitative data is analysed by using categorisation and sorting. Although most case studies are based on qualitative data, however, there are combinations of both quantitative and qualitative data within case studies and these are sometimes referred to as mixed methods (Robson 2011). This supports the use of mixed methods case study in this study.

Triangulation in case studies is very important in order to increase the precision of the research, especially when relying on qualitative data, since this is broader and richer but less precise than quantitative data (Runeson & Host 2009). There are four types of triangulation (Stake 1995):

- Data triangulation: involves more than one data source where collection takes place in the study.
- Observer triangulation: using more than one observer in the case study.
- Methodological triangulation: using more than one method e.g. qualitative and quantitative methods in the study.
- Theory triangulation: involves using alternate theories or viewpoints in the study.

Pan and Tan (2011) have proposed a structured pragmatic situational (SPS) approach to perform case studies and it involves eight steps. The steps are:

- Step 1: Access negotiation – gaining access to the organisation where the case study will be performed;
- Step 2: Conceptualising the phenomenon – gathering information about the organisation and the phenomenon;
- Step 3: Collecting and organising the initial data-use open coding by breaking data into themes examined, comparing for similarities and differences, and categorising;
- Step 4: Constructing and extending the theoretical lens;
- Step 5: Confirming and validating data;
- Step 6: Perform selective coding;
- Step 7: Ensure theory-data model alignment;
- Step 8: Write the case report.

An alternative approach is proposed by Runeson and Host (2009), and they proposed five major processes to conduct case study research:

- Case study design – defining the objectives and planning the case study;
- Preparation for data collection – defining the procedures and protocols for data collection;
- Collecting evidence – collecting data on the phenomenon;
- Analysis of the collected data;
- Reporting.

The main difference between the case study processes proposed by Pan and Tan (2011) and Runeson and Host (2009) is that Pan and Tan (2011) use the Grounded Theory of Strauss and Corbin (1998) to collect and analyse the data whereas Runeson and Host (2009) use open coding to analyse the data. The researcher based the case study methodology utilised in this study on the one proposed by Runeson and Host (2009). This is because the steps of the case study processes proposed by Runeson and Host (2009) can easily be achieved within a mixed methods setting. Therefore, it is used to carry out the case study in this research.

Cepeda and Martin (2005) have highlighted that a sound case study should have three main elements:

- conceptual framework;
- research circle;
- literature-based examination of resulted theory.

They maintain that researchers must bring ideas (or frameworks) about appropriate notions in their area of interest by surveying relevant literature relating to their research topics and identifying gaps in their research area. This will allow researchers to make their views exciting in their area of study and gain exposure to a range of concepts, ideas and theories. These three main elements are further explained in the next four paragraphs that follow.

Cepeda and Martin (2005) maintain that the conceptual framework is developed from the research propositions, existing knowledge found in the literature and explained by a researcher's theoretical underpinnings. The conceptual framework should be defined at the begining of the research project and should be critically examined during the research cycles in order to understand the research topic and theme. The conceptual framework is a series of changing models that undergoes continuous review and refinement throughout the lifecycle of the research project (Cepeda & Martin 2005). One of the inputs of high-quality research is to represent this abstract composition, challenge its underlying norms and explain it (Cepeda & Martin 2005). Miles and Huberman (1994) regard this as the conceptual framework that explains the main things to be studied in a graphic or narrative form as well as the input factors, constructs or variables and the assumed relationships between them.

The research cycle involves planning the research; data gathering and data analysis. The organisations to obtain data from should be identified and methods for gathering, recording, processing and analysing data (as well as related criteria for rigour and validity) and method to report the findings should be identified. The researcher collects and analyses data and takes field notes during the data collection stage. Data gathering takes place first, before data analysis in quantitative research whereas they may overlap in qualitative research (Cepeda & Martin 2005). There are many approaches to analyse qualitative data, among them: (a) Toulmin's 1988 approach; (b) coding process; (c) use of Grounded Theory. A bottom-up approach grounded in data is used to analyse the qualitative data in this research.

Coding as a qualitative research data analysis approach involves using the concepts of the conceptual framework as the initial codes that direct the analysis along with other codes to combine new premises (Caroll & Swatman 2000). This should be done iteratively in order to gain a deep understanding of the data and its fundamental patterns which can then result in new concepts and themes. The researcher should then reflect after the data analysis is performed with the view to challenge, confirm or revise and update the conceptual framework to include knowledge acquired through the research cycle.

The last stage is the theory building phase and this is the result of the completion of the reflection phase. This involves the clarification and categorisation of concepts as well as specifying the relationships between categories in order to generate theories from the research theme. This process will be performed many times (iteratively) as part of a hermeneutic cycle to build the theory. Cepeda and Martin (2005) point out that case studies build theory from many (multiple) cases to improve the conceptual framework.

Case study research in IS has increased and its validity as a research method is no longer deliberated by researchers (Mingers 2003; Lee & Hubona 2009). The number of interpretive studies in leading Information Systems journals has been steadily increasing and most of them make use of case study research methodology (Mingers 2003). Case study research has been used extensively in IS research (Benbasat *et al.* 1987; Smit 1990; Gable 1994; Renken & Moswetsi 2006; Runeson & Host 2009; Pan & Tan 2011; Jurisch *et al.* 2013). In IS research, case studies have been described as a research method ideal for understanding the interactions between IT and organisational contexts because they allow for a multitude of sources of data collection in order to understand the phenomenon that is being investigated (Drake *et al.* 1998). Case study is used in this research, and the data were gathered in seven government organisations (data triangulation) as mentioned in section 1.1.1.

4.4.2 Justification for Using Case Study Research

Case studies are appropriate for examining complex phenomena (Klein & Meyers 1999), processes (Gephart 2004) and are also good for addressing research questions involving the ‘how’ and ‘why’ questions (Walsham 1995). Case study research allows understanding of the problem by the researcher, the complexity and nature of the process taking place, obtaining valuable perceptions that can be acquired and contributing to knowledge by using findings to generalise theory (Yin 2003). Case studies are used to support developing theory related to poorly understood occurrences (Cepeda & Martin 2005; Welch *et al.* 2011).

The strengths of the case study research method in Information Systems include: (a) IS can be studied in a normal location, (b) the learning of the state of the art and the generation of theories from practice; (c) the environment and difficulty of the process taking place can be understood by the researcher; (d) insights into the new topics emerging from IS field can be gained (Cepeda & Martin 2005).

Case studies can assist to develop or refine theory and it can be approached in different ways subject to the epistemological standpoint of the researcher (Crowe *et al.* 2011). That is, the researcher might take a critical, interpretivist or positivist approach, although it may be better to use more than one approach in any case study. A management framework is developed in this research and case study is the preferred methodology since it can assist to develop a theory.

The aim of this research is to develop a management framework to manage sensitive information during software platforms migration. Case study is an effective research method to study this phenomenon by allowing the researcher to comprehend the problem, the complexity and nature of the development taking place and finally the conceptualisation of the management framework (theory generalisation). A mixed methods case study approach was used in this work that involved multiple data collection sources (data triangulation) in different organisations to develop the management framework.

4.4.3 Limitations of Case Study

The case study is like other research approaches not without limitations. Researchers should avoid the temptation of collecting too much data in a case study as it might take too much time to complete the study. Researchers should also set aside some time for data analysis and interpretation of the resulting complex data sets (Crowe *et al.* 2011).

Yin (2009) has criticised case study to lack scientific rigour and does not provide much basis for generalisation (that is, producing transferable findings). These concerns can be overcome by using transparency; respondent validation and theoretical sampling throughout the research process (Crowe *et al.* 2011). Transparency can be realised by the researcher describing in detail the reasons for the methods chosen, data collection, all the steps in the case selection, and the background of the researcher as well as his or her level of involvement during data collection, and data interpretation (Crowe *et al.* 2011). Transparency and respondent validation are used in this research.

Weaknesses of case study research are lack of deducibility, controllability, generalisability and repeatability (Lee 1989). However, Lee (1989) stresses that these problems are not common and that they are also present to some degree in other research methods. Despite these weaknesses, case study is the preferred methodology to carry out this research since it offers data triangulation which contributes to the validity of the research.

4.5 Data Collection

The quantitative and qualitative data were collected consecutively, with the quantitative data collected firstly and then the qualitative data. This research is a multi-strand study in the sense that more than one source of data and research method were used. The qualitative design methodology used in this research is both correlative and descriptive. Firstly, questionnaires were distributed to CSIR IT staff, a government organisation based in Pretoria, South Africa. This was done in order to perform item analysis and also to finalise the questionnaire for distribution. Twenty five (25) respondents completed the first pilot version of the questionnaire that was used for item analysis. After the item analysis was performed some questions were removed while others were re-worded and the final version of the questionnaire was developed. During this stage, the Exploratory Factor Analysis could not be performed due to the limited number of respondents that participated in the pilot phase.

The reason for conducting Item Analysis is to find the items that form the internally consistent constructs and to remove those that do not and this reflects the extent of inter-correlation among the items. The items that do not inter-correlate imply that they do not represent a common underlying construct, while the ones that inter-correlate show that they share a common variance or they indicate that they share the same underlying construct (Wiid & Diggines 2013). The internal consistency of a scale is measured by the Cronbach Coefficient Alpha. Item analysis is used to assist in building reliability and validity in the test measuring instrument and can be both qualitative and quantitative.

The final version of the questionnaires was then distributed to 250 respondents in the government organisations mentioned in section 1.1.1 and based in Pretoria, South Africa, for completion in order to obtain statistical (quantitative) data. Ninety completed questionnaires were received back from the respondents. The primary data collection method in IS research are surveys and controlled experiments as well as inferential statistics (Johari 2009).

4.5.1 The Measuring Instrument

The key variables in this research were measured by a self-report questionnaire. The measuring instrument used in the research is divided into six sections (i.e.) sections A to F. Section A is about the biographical data of the respondents and these include: the respondent’s organisation; race; age; respondent’s type of work; and if the respondent has been part of a migration project. The rest of the questionnaire assessed the variables in the research hypotheses. The questions in sections B to E are based on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The Likert scale is also called the summated rating scale and it is based on the premise that each statement within the scale has the same significance in answering the research question (Kumar 2005). Section B is sub-divided into three sections (sections B1, B2 and B3) and is about the awareness of the sensitive nature of data. Section B1 statements concern the employee behaviour: staff awareness of the sensitive nature of company data.

Section B2 statements are related to training: awareness of the sensitive nature of data. Section B3 statements cover employee accountability: awareness of the sensitive nature of data.

Section C is sub-divided into four sections (sections C1, C2, C3 and C4). Section C1 statements are about organisational strategy: handling sensitive nature of data. Section C2 statements relate to organisational policies and procedures: handling sensitive nature of data. Section C3 statements involve the organisational data: preparation towards ensuring sensitive data management. Section C4 statements are about the organisational standards (processes, hardware & software): enforcement will ensure proper handling of sensitive data.

Section D consists of five parts (sections D1, D2, D3, D4 and D5). Section D1 statements are about data categories and business rules: providing a basis for data classification during migration of sensitive data. Section D2 statements involve data classification system: addressing security issues when handling sensitive data during migration of platforms. Section D3 statements focus on data protection tools: ensuring sensitive data protection during migration of platforms. Section D4 statements are about data sensitivity assessment: identifying different protection needs for information. Section D5 statements involve security models: ensuring protection of sensitive data during migration.

Section E is sub-divided into five parts (sections E1, E2, E3, E4 and E5). Section E1 statements are about data migration planning: protecting sensitive data during migration. Section E2 statements involve data migration process: protecting sensitive data during migration. Section E3 statements centre on data migration tools: protecting sensitive data during migration. Section E4 statements are about data migration controls: protecting sensitive data during migration. Section E5 statements focus on data migration monitoring: protecting sensitive data during migration. Section F is the last section in the questionnaire and is used to provide for further comments from the respondents.

Each respondent was assigned a number on each variable that ranged from one to ninety. The respondent’s responses were then collated in an Excel spreadsheet. Scores on each item that measure the variable were averaged for further data analysis. The measuring instrument is included in the Appendix B.

4.5.2 The Qualitative Interview Process

Interviews are used extensively in qualitative research (Petty *et al.* 2012) and they may be structured, unstructured or semi-structured (Robson 2011). Structured interviews are like the questionnaire type which yield shallow response level. In unstructured interviews, the researcher is led by the direction of the participant. There are few pre-determined areas in semi-structured interviews with possible occasions where the researcher has to guide the conversation (Petty *et al.* 2012). Interviews can be done either face to face, by the internet or via the telephone and they can last between 30 and 90 minutes (Peng *et al.* 2011). Interviews are audio-taped and later transcribed.

Qualitative data was collected from interviewees from different South African organisation employees, mostly working in the IT security section of their organisations by using memos (direct observation/participant observation) and audio tapes to perform tape-recording of the interviews. The purpose of interviews in qualitative research is to explain and depict people’s experiential life (Schwandt 2001). The interviewees were allowed to talk freely during the interview session and they were not interrupted while answering the questions. The intervieewes were subjected to direct observation by the researcher during the interviewing sessions. Ten (10) IT specialists were interviewed by the researcher to gather the qualitative research data. All the interviews took place in offices of the individual interviewees and the only people present in the rooms were the researcher and the interviewee.

Most of the interviewees used their hands to gesticulate their key points during the interview sessions. The researcher showed empathy, listened to the interviewees in a relaxed manner and responded to answers at times by nodding, smiling and shrugging of the shoulders. The interviewees listened to the questions and answered them appropriately. According to Rowley (2012), researchers can conduct 12 interviews within 30 minutes per interview or six to eight interviews within one hour per interview in a study. Collins *et al.* (2007 p. 273) come up with a table on the minimum sample size recommendations for most common quantitative and qualitative research designs. According to Collins *et al.*'s table, the minimum sample size suggestion is 3 – 5 participants (Cresswell 2002) for case study research design; the minimum sample size suggestion is 10 interviews (Cresswell 1998) and 6 (Morse 1994) for phenomenological research design. Therefore these reasons suggest that the use of 10 interviewees is adequate for this study.

The research questions asked in this study during the qualitative data collection process are listed in Appendix C. Additionally, the researcher asked follow-up questions from the interviewees depending on their responses to the questions asked from them. This implies that the researcher used an incomplete script and thus allowed for flexibility, openness and improvisation during the interview sessions (Myers & Newman 2006).

4.6 Data Analysis

It is anticipated that the quantitative and qualitative data analysis will inform each other during the data analysis phase. The quantitative data were analysed using exploratory factor analysis, reliability analysis, descriptive analysis and Spearman's correlation. The new constructs identified during the quantitative data analysis were then mapped to the initial preliminary management framework. The set of questions asked during the qualitative interviews were drawn up from the results of the quantitative analysis. The qualitative study is aimed to validate the results obtained from the quantitative analysis.

The qualitative interviews were transcribed word for word and the transcripts were read many times so as to understand the data and to be able to code and gather new categories using NVIVO. The analysis of the qualitative data was informed by the quantitative data so that pattern expanding, confirming or contradicting the quantitative results can be discovered. Moreover, the qualitative data brought new variables that need to be confirmed using the quantitative data and this movement between the two analyses enhances the researcher's reflexivity. New variables were discovered during the qualitative analysis while most of the identified variables during the quantitative analysis were validated.

4.7 Conclusion

In this chapter, the various philosophical perspectives are outlined and the reason why pragmatism is chosen to be the philosophical paradigm underpinning the research philosophy is explained. The research questions in section 1.2 were formulated on the basis of the researcher's pragmatic stand. The research is not grounded in notions about the nature of knowledge. The research methods – qualitative, quantitative and mixed methods, are all explained and the reason for using mixed methods research is also clarified in this chapter. The steps followed when conducting the mixed methods research were explained.

The various research methodologies used in IS research are all described and the reason for adopting the case research methodology is clarified. The mixed methods data collection processes and how quantitative and qualitative data were obtained are highlighted.

The next chapter relates to the quantitative data analysis, in which the preliminary management framework on information sensitivity during migrations is presented to be validated by using the qualitative method in Chapter Six.

Chapter 5

Preliminary Management Framework: Quantitative Data Analysis

5.1 Introduction

The previous chapter focused on the research design and methodology and explained the mixed methods research method as well as the data collection and the data analysis processes employed in this research. Questionnaires will be used to collect quantitative data while semi-structured interviews will be used to collect qualitative data. The quantitative data will be analysed using exploratory factor analysis, reliability analysis, descriptive analysis and Spearman's correlation.

In this chapter, the results of the quantitative data analysis used in the research are explored to develop the preliminary management framework. The quantitative data analysis is carried out using the JMP Version 11.0 software from the Statistical Analysis System (SAS). JMP Software is a statistical discovery software that combines powerful statistics with dynamic graphics, in memory and on the desktop (SAS 2014). The responses from the respondents were captured in a Microsoft Excel document and then imported into the JMP software for analysis.

Two hundred and fifty (250) questionnaires were distributed to IT specialists working in the six South African Government organisations mentioned in section 1.1.1. Ninety (90) questionnaires were completed and received by the researcher and this gives a response rate of 36%. The questionnaire used to carry out the quantitative part of the research is shown in Appendix B. The content of the questionnaire incorporates both the rudimentary management framework in section

2.14 and the model resulting from the security challenges during OSS migration in section 3.5.

This chapter elaborates on the statistical analysis performed in this research. These include biographical data distributions, factor analysis and reliability analysis. A preliminary management framework from quantitative analysis is developed in section 5.11.

5.2 Biographical Data Distributions

5.2.1 Type/Nature of Respondent Employment

The Biographical Data is the first component in the questionnaire called component A as depicted in Appendix B. The Biographical Data Distribution is shown in figures 5-1 to 5-10. These figures were derived from the Tables E5-1 to E5-10 respectively as shown in Appendix E. Figure 5-1 describes the type/nature of respondent employment. The majority of the respondents were from three government organisations, namely: SITA, South African Department of Public Works and South African Department of Social Development.

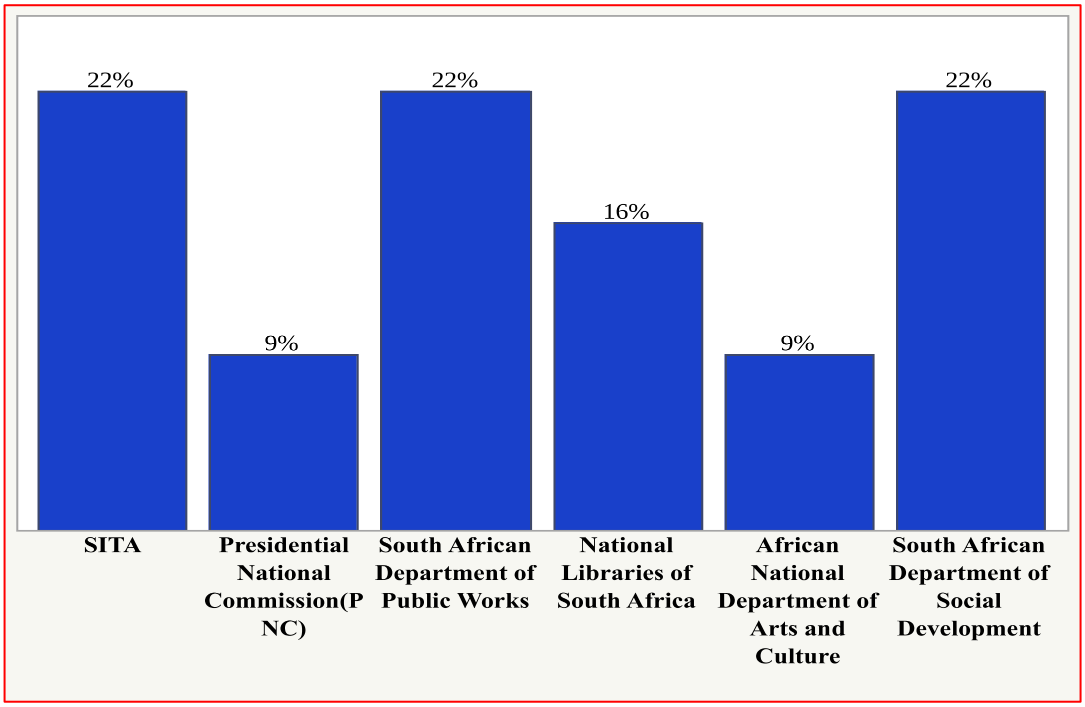


Figure 5-1: Type/Nature of Respondent Employment

5.2.2 Respondent’s Post Levels (IT Specialists)

Figure 5-2 shows the respondents’ post levels for the IT specialists. The figure shows that the majority of the respondents fall into the developers and the junior developers categories (49% and 28% respectively).

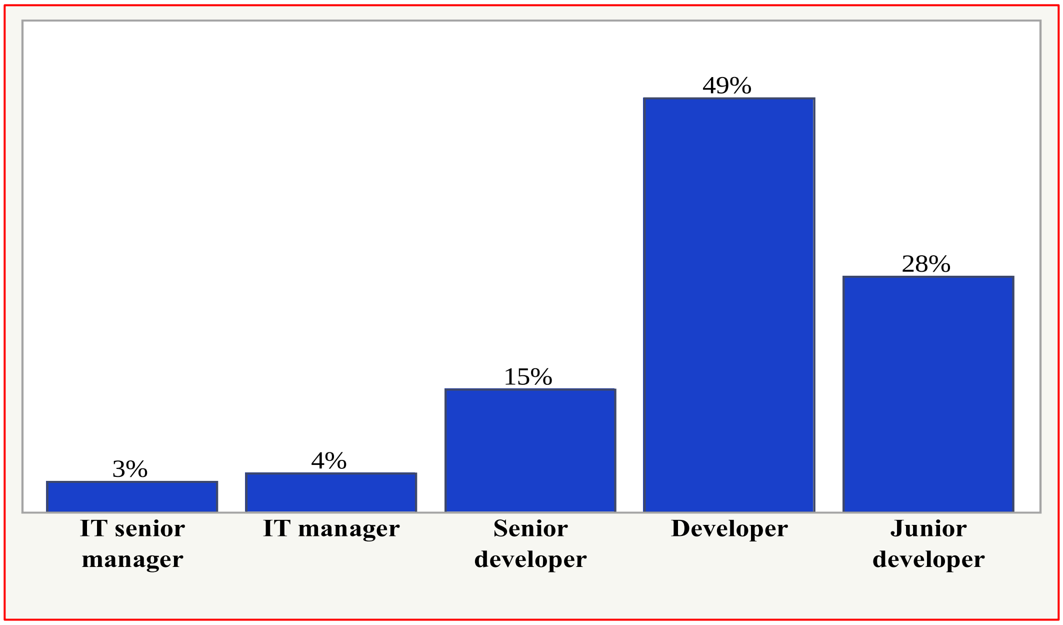


Figure 5-2: Respondent Post Level (IT Specialists)

5.2.3 Population of Respondents by Gender

Figure 5-3 depicts the population of the respondents in terms of their gender – male or female. The figure reveals that almost equal numbers of females and males completed the quantitative questionnaire (52% and 48%).

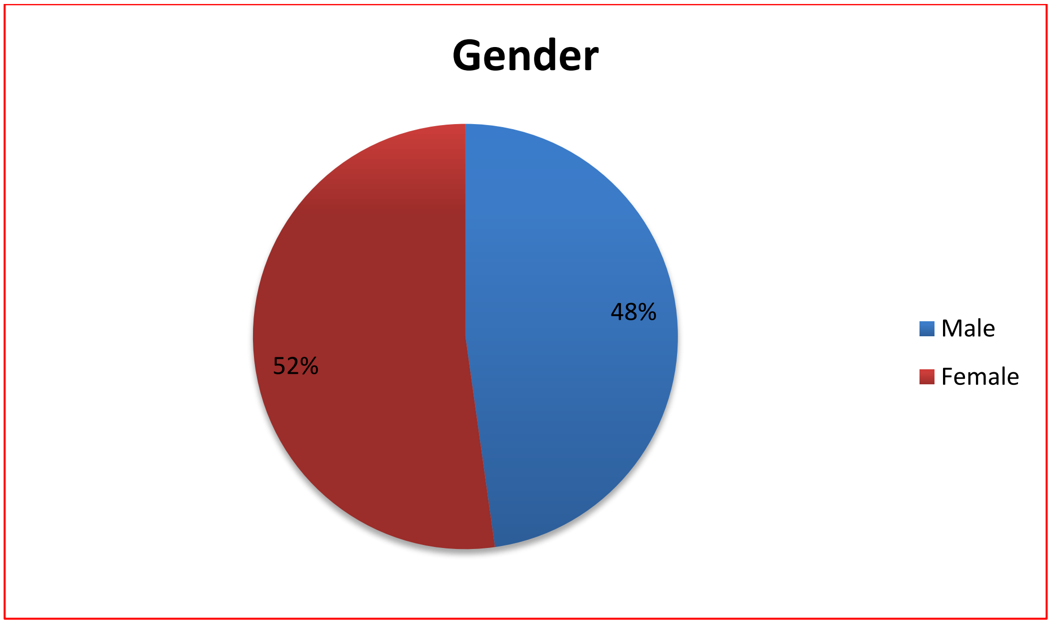


Figure 5-3: Population of Respondents by Gender

5.2.4 Population of Respondents by Race

The population of the respondents by race is depicted in Figure 5-4. It shows that most of the respondents are black (91%). This suggests that most IT specialists working in South African government departments based in Pretoria are black people since it could be that the demographics of the people living in Pretoria are mostly black (STATISTICS SA 2013).

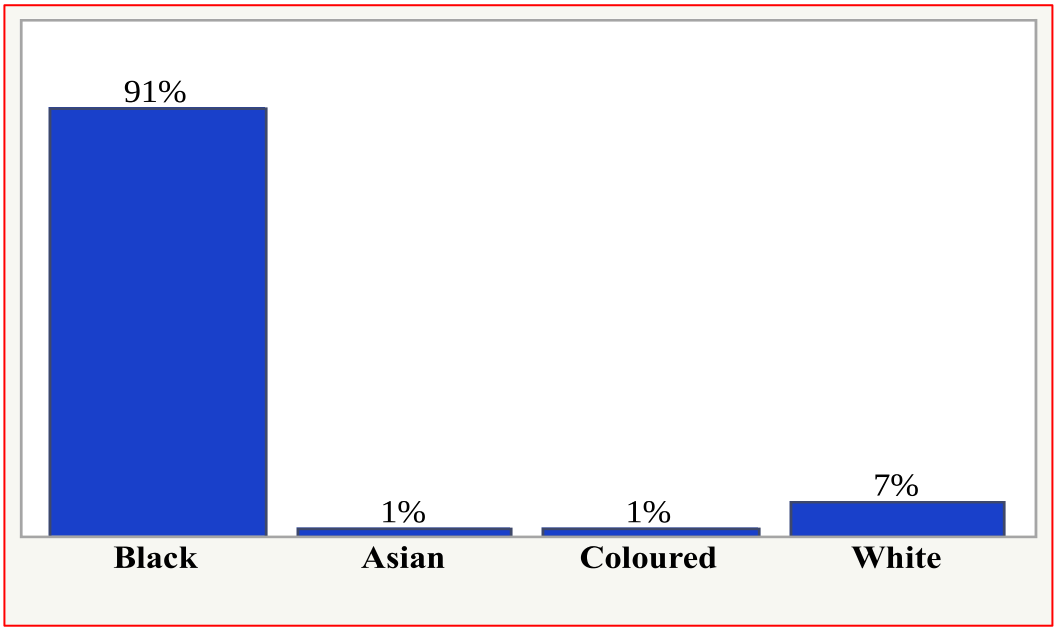


Figure 5-4: Population of Respondents by Race

5.2.5 Age Distribution of Respondents

The Age distribution of all the respondents is shown in Figure 5-5. It can be inferred from the figure that most respondents are in the age group between 26 and 35 and 36 to 45 (60% and 34% respectively). This suggests that most IT specialists in South African government departments in Pretoria are still youthful.

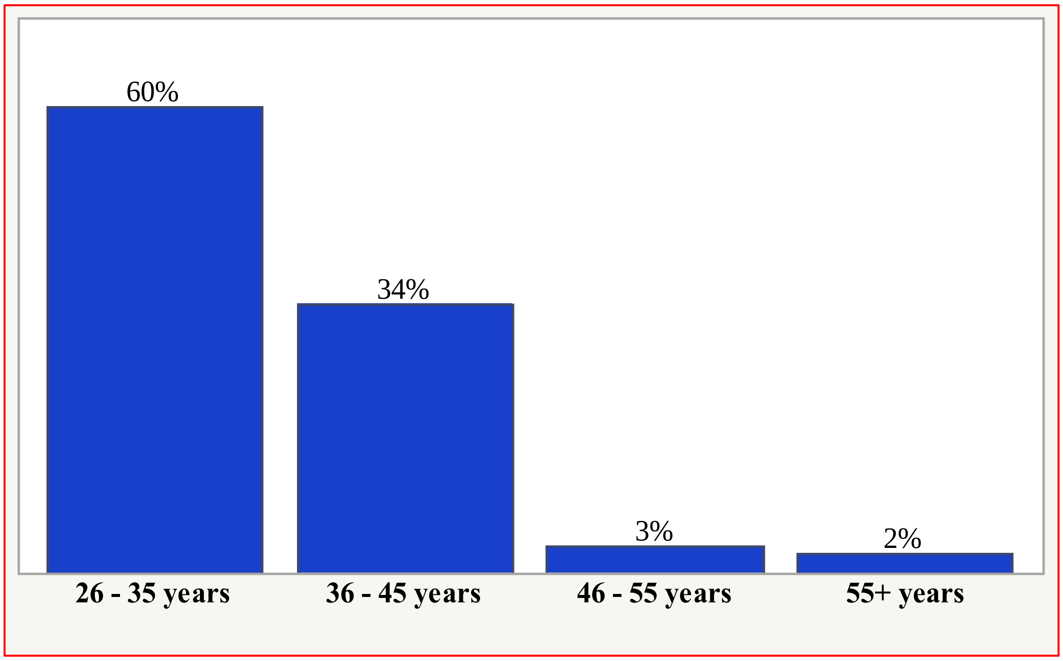


Figure 5-5: Age Distribution of Respondents

5.2.6 Respondent's Type of Work

Figure 5-6 depicts the respondents' type of work in their organisations. It shows that most of the respondents work at transferring and loading data/ETL migration and data security/IT security (21% and 34% respectively). This suggests that the majority of the IT respondents are from the data/IT security domain.

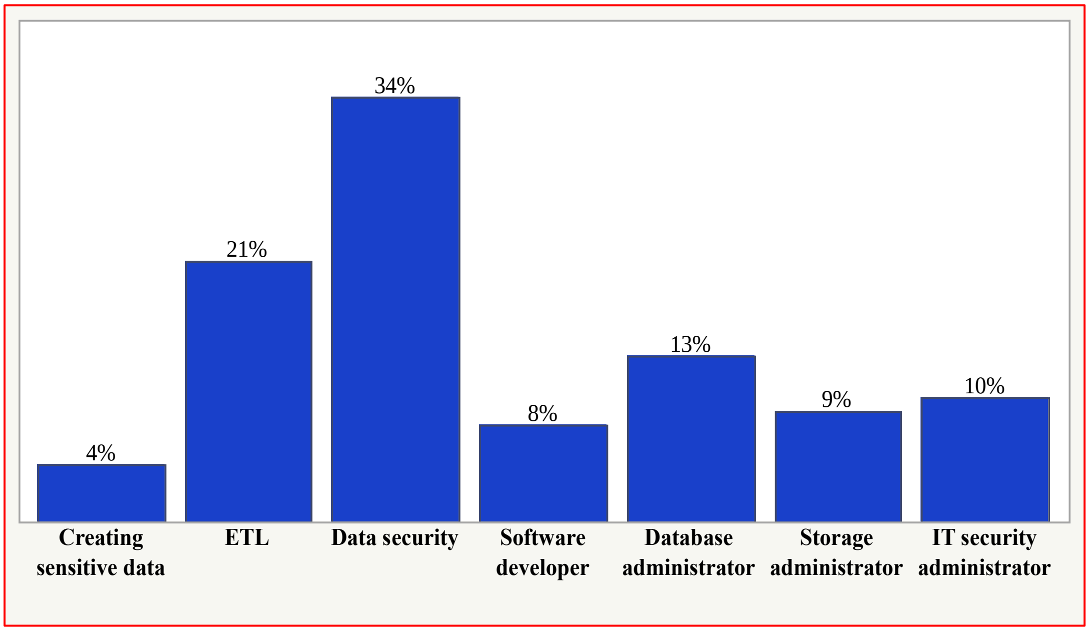


Figure 5-6: Respondents' Type of Work

5.2.7 Respondents' Employment Category

The respondents' employment category is highlighted in Figure 5-7. The figure reveals that most of the respondents are permanently employed (89%). This suggests that most IT specialist working in South African government departments in Pretoria are permanently employed and could, therefore, be more responsible in the handling of sensitive data.

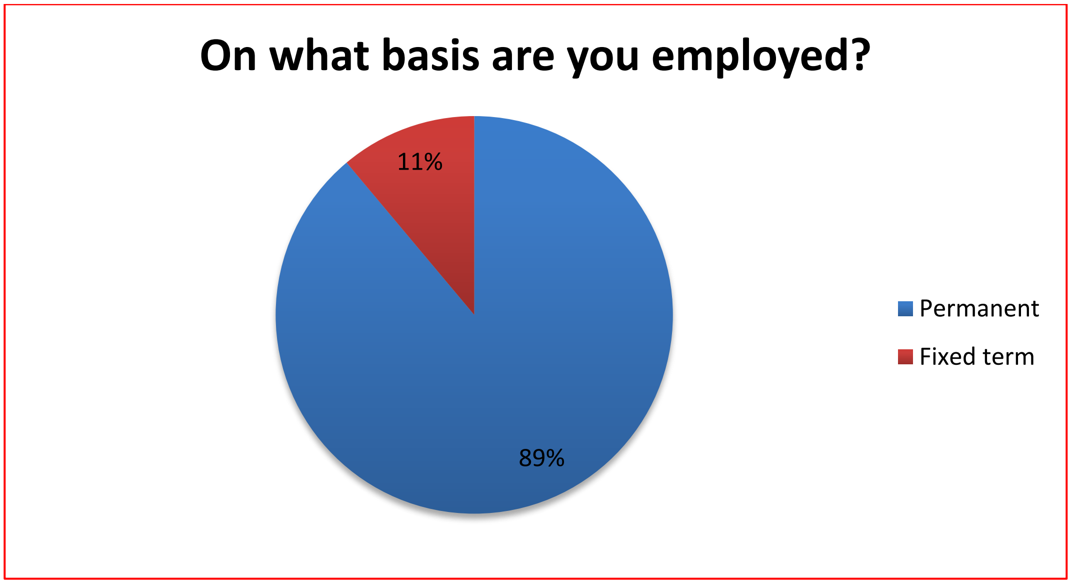


Figure 5-7: Respondent’s Employment Category

5.2.8 Respondents’ Awareness of Sensitive Data Management Policy

The respondent’s awareness of sensitive data management policy in organisations is depicted in Figure 5-8. It shows that most of the respondents are aware of sensitive data management policy in organisations (92%). This shows that there could be an awareness of sensitive data management policy among the IT respondents.

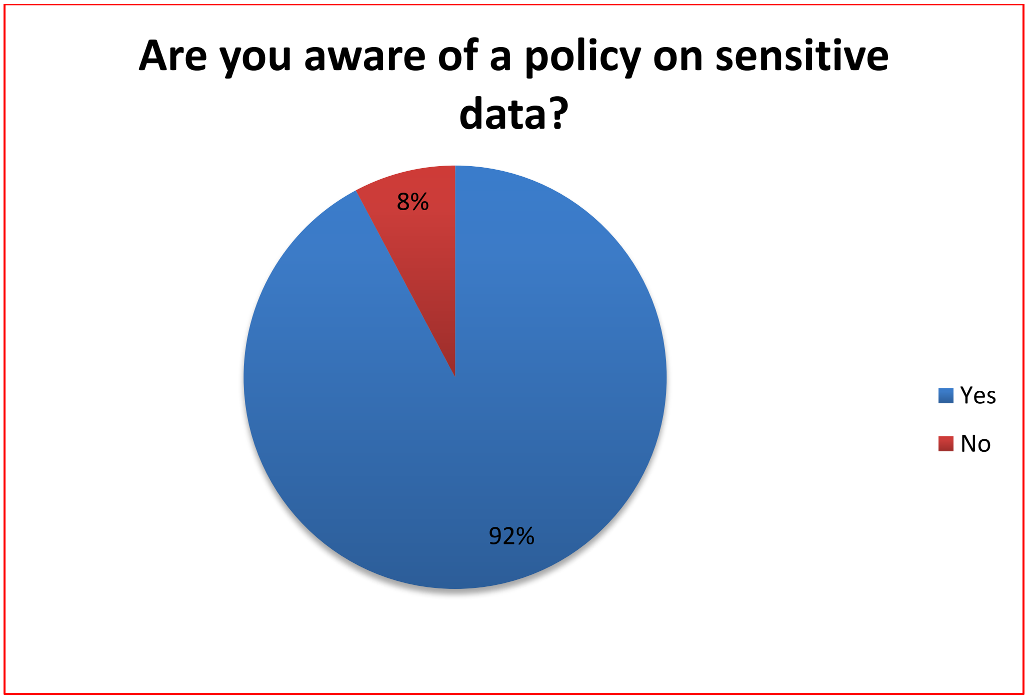


Figure 5-8: Respondents' Awareness of Sensitive Data Management Policy

5.2.9 Respondents' Years in Service with Company

The respondents' number of years in service with their companies is illustrated in Figure 5-9. From this figure, most of the respondents have worked in their organisations for 3 to 5 years and 6 to 10 years (54% and 22%).

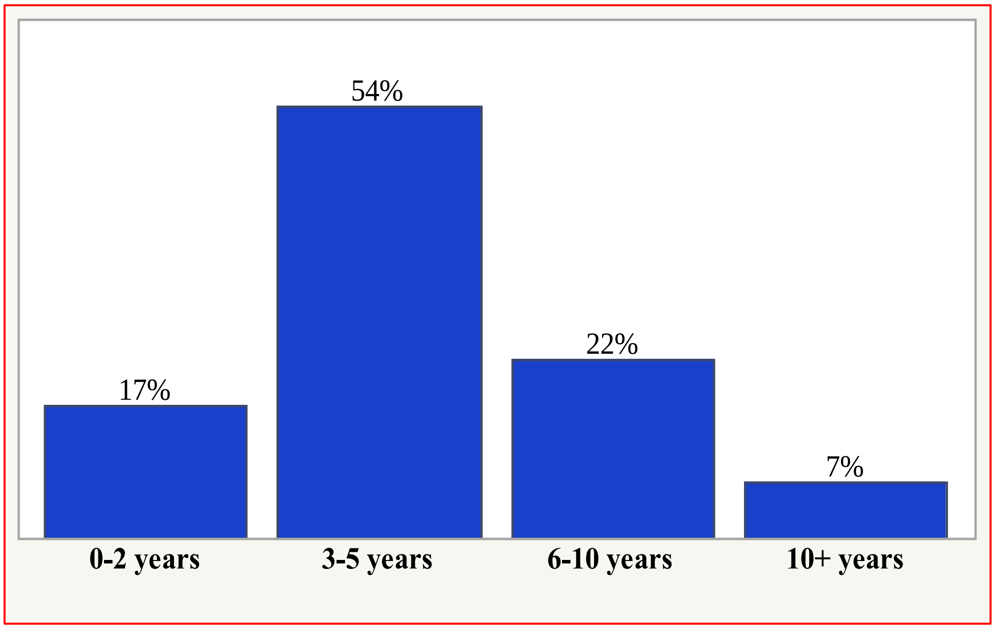


Figure 5-9: Respondents' Number of Years in Service with Company

5.2.10 Respondents' Participation on Platform Migration Projects

The respondents' participation in platform migration projects is shown in Figure 5-10. This figure reveals that the majority of the respondents have participated in migration projects (94%). This means that most respondents have been part of migration projects and their contributions would be valuable in the research due to their knowledge in this area.

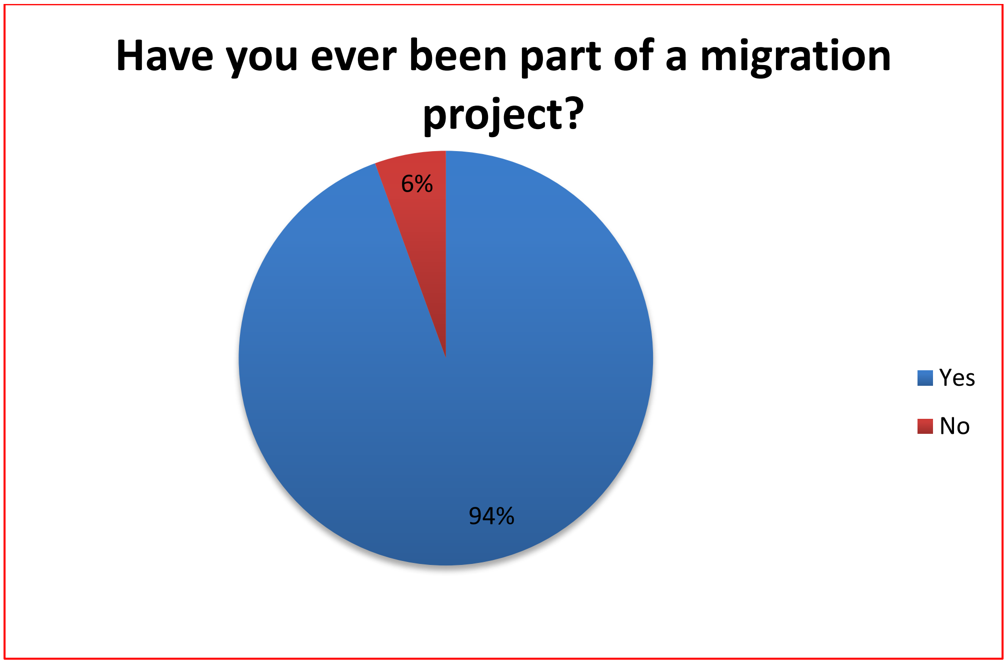


Figure 5-10: Respondents' Participation on Platform Migration Projects

5.3 Identifying Constructs using Factor Analysis

Factor analysis is used to identify and shorten the number of constructs from a large number of items (Worthington & Whittaker 2006). It is also used to authenticate the validity of newly developed measuring instruments, that is, to check if the newly developed measuring instrument is measuring the intended constructs (Worthington & Whittaker 2006). It is used to investigate the integrity of the measuring instrument and further aids in theory refinement (Hendrick & Hendrick 1986). Validity is a measure of the accuracy of a measuring instrument, that is, to which extent the measuring instrument measures what is intended to be measured (the intended concept) (Linn & Grondlung 2000; Stewart 2009; Ritter 2010; Tavakol & Dennick 2011). The two major types of factor analysis are: (a) Exploratory Factor Analysis (EFA) and (b) Confirmatory Factor Analysis (CFA) (Thompson 1992; Kahn 2006).

Exploratory Factor Analysis (EFA) is used to produce theory (Henson & Roberts 2006), and it is also used to evaluate the construct validity during the early development of a measuring instrument. The idea is to identify and eliminate the items that do not measure an intended construct or measure multiple constructs that could be poor indicators of the desired construct (Worthington & Whittaker 2006). Construct validity has been described by Wiid and Diggines (2013) as the degree to which a construct measures what it was designed to measure in order to ensure that the overall scale consists of the correct constructs. Researchers have been employed to utilise inductive reasoning when they use EFA in order for them to deliver meaningful research outcomes. In IS, EFA is used as a pre-study to verify the validity of the measuring instruments used (Treblmaier & Filzmoser 2009).

Confirmatory Factor Analysis (CFA) is used to examine the theory after the constructs have been identified and the variables describing each construct are recognised (Henson & Roberts 2006). CFA is mostly performed by using Structural Equation Modelling (SEM). It is necessary to have prior knowledge of the expected relationships between items and constructs before performing CFA (Worthington & Whittaker 2006). CFA is often used, after a measuring instrument has been evaluated using EFA, in order to understand whether the construct structure created by the EFA corresponds to the data from the new sample (Worthington & Whittaker 2006).

In this research, only the EFA was performed to explore and identify the constructs as the first phase of the validity process. The CFA was not performed because the data was not sufficient to perform SEM and there was not enough theoretical backing. Moreover, hypotheses could not be used to support the research questions because there was no theory to be confirmed in this research. The EFA is used to evaluate the construct validity of the constructs in the questionnaire following the following steps:

Step 1:

The Kaiser-Meyer-Olkin Measure of Sampling adequacy value (KMO value) was measured on all the items (or questions) in the questionnaire. This was done in order to determine the viability of conducting an EFA on the questions in the questionnaire so as to provide a measure of the correlation structure of the questions on which the EFA analysis was performed. If there is a strong correlation structure, this implies that the individual items correlate well with each other and the questions can be grouped together into factors, thus allowing the questions to correlate well with each other to form the factors/constructs (Wiid & Diggines 2013). However, if the correlation structure is weak, it implies that the factors or constructs cannot be formed. Ranges of KMO values are from 0 to 1 with a KMO value higher than 0.5 indicating a strong enough correlation structure to perform an EFA while a KMO value below 0.5 indicates a weak correlation structure implying that it is not viable to conduct an EFA (Wiid & Diggines 2013).

Step 2:

An inspection of the communalities (common variance) of the individual questions was performed. This is to determine whether the questions can be part of the overall scale or ‘fit in’ with the rest of the questions. It is used to evaluate specific items that should be deleted or retained because a communality is an indication of the proportion of an item’s variance that is shared with the other items (construct structure) (Tabachnick & Fidell 2001). Communality refers to common variance (the variance that is shared with other items) as opposed to unique variance that is unique to that item (Worthington & Whittaker 2006). A communality value near 1 implies that a question correlates highly with the rest of the questions while a communality value below 0.4 means that the questions should be reconsidered (Tabachnick & Fidell 2001).

EFA is used to answer these two basic questions:

- How many factors or constructs are there in the scale or component (e.g. Section B)? and
- What are such factors or constructs/sub-constructs (e.g. sub-construct B1)?

Therefore, the first process in the EFA is to determine the number of factors (or constructs/sub-constructs) in the scale (component) while the second process in the EFA is the determination of the factors (or constructs/sub-constructs) that are in the scale (component), and which questions/statements (items) constitute the factors (or constructs/sub-constructs).

The criteria used to determine the number of factors (the first process in the EFA) are as follows (Treblmaier & Filzmoser 2009; Zwick & Velicer 1986; Thompson & Daniel 1996; Wiid & Diggines 2013), however, the researcher used only the first three criteria in this study because they are sufficient to determine the factors and ensure their validity:

- Cumulative percentage explained by the factors ≥ 60%.
- Eigenvalues ≥ 1 (this rule is also known as the Kaiser Guttman rule). Eigenvalues represents the sum of squared loadings for a factor and all factors with eingenvalues greater than 1 should be retained. This is the most frequently used criterion.
- Inspection of the Scree Plot. The Scree Plot describes a graph of the eigenvalues used to understand the importance of each factor. The Scree Plot is used by a researcher to inspect the descending values of eigenvalues in order to find a break in the size of eigenvalues, after which the remaining values tend to appear to level off horizontally. The Scree Plot should reveal a distinct break between the steep slope of the large constructs and the gradual trailing off of the remaining constructs.
- Minimum average partial correlation.
- Bartlett's Chi Square Test.

- Parallel analysis.

Step 3:

The factor loading of a variable on a construct is an indication of the weight quantity that is assigned to the construct. A factor loading value near 1 shows that the question loads highly on the specific factor while a factor loading of 0.40 on a specific factor is also considered meaningful (Wiid & Diggines 2013). A factor loading of 0.30 should be considered when a construct is being interpreted according to Tinsley and Tinsley (1987), while Tabachnick and Fidell (2001) state that the minimum acceptable factor loading for a construct is 0.32.

The original questionnaire is made up of four scales or Sections (B, C, D and E). Table 5-1 illustrates the grouping of questions in all the Sections of the questionnaire. The Sub-sections are described as follows: Section B1 = Employee Behaviour, Section B2 = Employee Training, Section B3 = Employee Accountability, Section C1 = Organisational Strategy, Section C2 = Organisational Policies & Procedures, Section C3 = Organisational Data, Section C4 = Organisational Standards, Section D1 = Data Categories & Business Rules, Section D2 = Data Classification System, Section D3 = Data Protection Tools , Section D4 = Data Sensitivity Assessment, Section D5 = Security Models, Section E1 = Data Migration Planning, Section E2 = Data Migration Process, Section E3 = Data Migration Tools, Section E4 = Data Migration Controls, and Section E5 = Data Migration Monitoring.

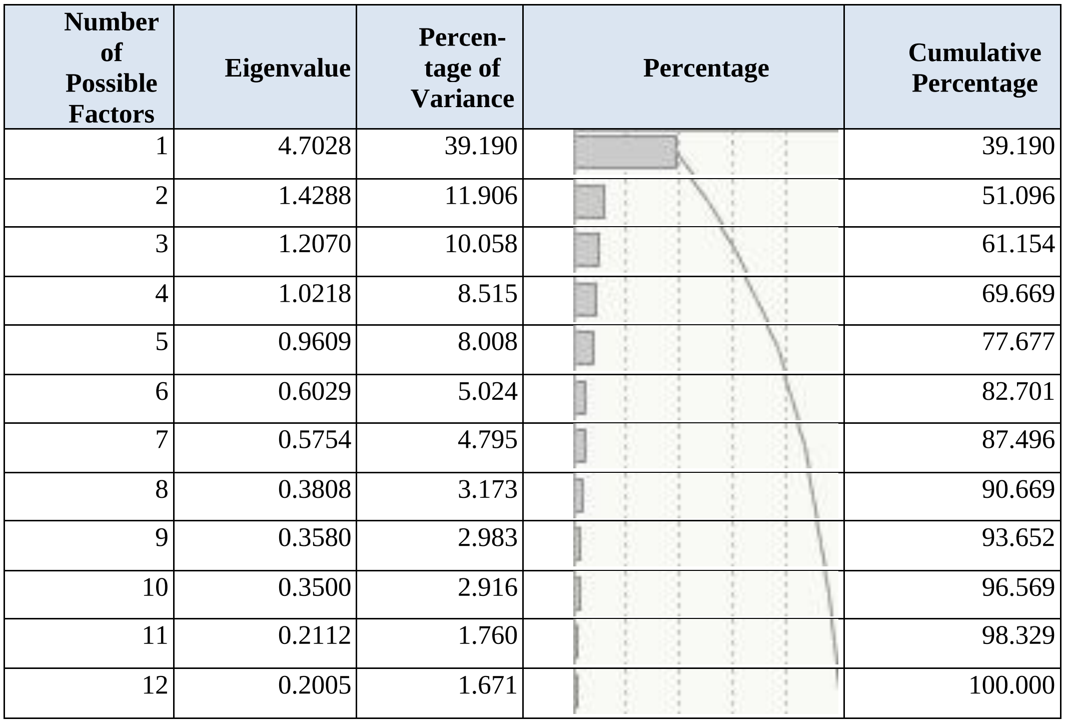
TABLE 5-1
Grouping of Questions in all the Sections of the Questionnaire

SECTION B = EMPLOYEE			SECTION C = ORGANISATION				SECTION D = DATA				SECTION E = DATA MIGRATION				
B1	B2	B3	C1	C2	C3	C4	D1	D2	D3	D4	E1	E2	E3	E4	E5
B1.1	B2.1	B3.1	C1.1	C2.1	C3.1	C4.1	D1.1	D2.1	D3.1	D4.1	E1.1	E2.1	E3.1	E4.1	E5.1
B1.2	B2.2	B3.2	C1.2	C2.2	C3.2	C4.2	D1.2	D2.2	D3.2	D4.2	E1.2	E2.2	E3.2	E4.2	E5.2
B1.3	B2.3	B3.3	C1.3	C2.3	C3.3	C4.3	D1.3	D2.3	D3.3	D4.3	E1.3	E2.3	E3.3	E4.3	E5.3
B1.4	B2.4	B3.4	C1.4	C2.4	C3.4	C4.4	D1.4	D2.4	D3.4	D4.4	E1.4	E2.4	E3.4	E4.4	E5.4
							D1.5		D3.5			E2.5			

An Exploratory Factor Analysis is performed on each section separately to identify sub-constructs or dimensions that are valid for analysis. The sub-constructs are re-grouped by using Exploratory Factor Analysis to obtain the valid sub-constructs. The EFA has been thoroughly explained in the previous sections and, from now onwards, only the summaries of how the EFA is obtained will be provided.

5.3.1 Exploratory Factor Analysis of Section B of the Questionnaire

Figure 5-11 shows the output of the EFA for section B (the eigenvalues and the % of variance declared by the factors). According to Wiid and Diggines (2013), if the value of the cumulative percentage is more than 60%, then it is considered to be adequate as the cut-off cumulative percentage to determine the number of factors. The first three factors (sub-constructs) show a cumulative percentage of 61.154%. Therefore, these three factors explain 61.154% of the variance in the original 12 items, and this is considered to be adequate enough to decide on the number of factors. Consequently, these three factors were extracted as shown in Figure 5-11.



Number of Possible Factors	Eigenvalue	Percen-tage of Variance	Percentage	Cumulative Percentage
1	4.7028	39.190		39.190
2	1.4288	11.906		51.096
3	1.2070	10.058		61.154
4	1.0218	8.515		69.669
5	0.9609	8.008		77.677
6	0.6029	5.024		82.701
7	0.5754	4.795		87.496
8	0.3808	3.173		90.669
9	0.3580	2.983		93.652
10	0.3500	2.916		96.569
11	0.2112	1.760		98.329
12	0.2005	1.671		100.000

Figure 5-11: Output of the EFA for Section B: The Eigenvalues and % of Variance Declared by the Factors

Figure 5-12 shows the Scree Plot. In this figure, it can be seen that the first two factors decline most steeply, but three factors were taken and this implies that the decision to keep the three factors is sufficient. This is considered to be adequate enough to decide on the number of factors, and consequently, these three factors were extracted as shown in Figure 5-13.

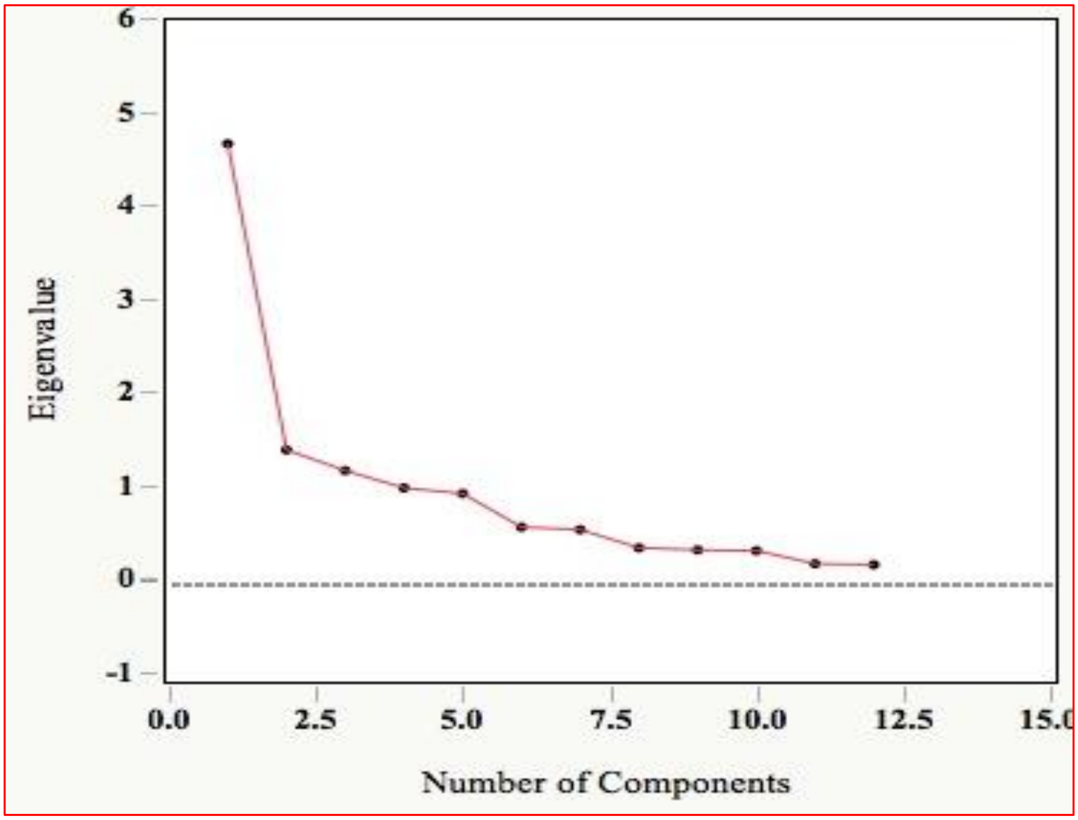


Figure 5-12: Output of the EFA on Section B: Scree Plot

Figure 5-13 depicts the output of the EFA on section B with emphasis on the rotated matrix with factor loadings. This figure was used to determine the composition of the factors. The maximum likelihood method was used to extract the factors, and this was followed by a varimax (orthogonal) rotation. For interpreting the factor loadings, an item is said to load on a given section if the factor loading was 0.40 or greater for that section and less than 0.40 for the other. The factor loadings in Figure 5-13 which are beneath each column of the figure for each factor were inspected. The colour of each item (question) in the first column indicates the sub-construct that load on the factor with the same colour.

Statements	Factor 1	Factor 2	Factor 3
B1.1	0.70	0.08	0.13
B1.2	0.73	0.12	0.25
B1.3	0.36	0.22	0.60
B1.4	0.16	0.09	0.53
B2.1	0.55	0.56	0.22
B2.2	0.14	0.27	0.41
B2.3	0.15	0.10	0.42
B2.4	0.53	0.72	0.03
B3.1	0.48	0.28	0.26
B3.2	0.03	0.66	0.54
B3.3	0.38	0.05	0.15
B3.4	0.06	0.76	0.29

Figure 5-13: Output of the EFA on Section B – The Rotated Matrix with Factor Loadings

Table 5-2 indicates how the questions were re-grouped in section B after performing EFA on section B of the questionnaire in order to ensure the validity of the identified sub-constructs. The following names for the factors are proposed that make logical and theoretical sense derived from the set of the statements (items) that make up the Factor: Factor 1: Awareness accountability score; Factor 2: Training handling; Factor 3: Consequences of sensitive data.

TABLE 5-2
Re-Grouping of Questions in Section B of the Questionnaire

Factor 1: Awareness accountability score	Factor 2: Training handling	Factor 3: Consequences of sensitive data
Question B1.1	Question B2.1	Question B1.3
Question B1.2	Question B2.4	Question B1.4
Question B3.1	Question B3.2	Question B2.2
	Question B3.4	Question B2.3

5.3.2 Exploratory Factor Analysis of Section C of the Questionnaire

Figure 5-14 shows the output of the EFA for section C (the eigenvalues and the % of variance declared by the factors). If the value of the cumulative percentage is more than 60%, then it is considered to be adequate. The first three factors show a cumulative percentage of 64.671%, therefore, these three factors explain 64.671% of the variance in the original sixteen items, and this is considered to be good. Consequently, these three factors were extracted as shown in Figure 5-14.

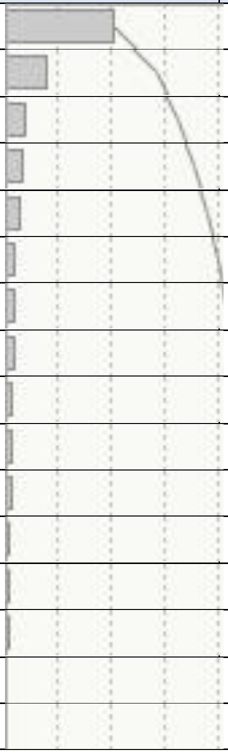
Number of Possible Factors	Eigenvalue	Percentage of Variance	Percentage	Cumulative Percentage
1	6.6039	41.274		41.274
2	2.4932	15.582		56.856
3	1.2504	7.815		64.671
4	0.9838	6.149		70.820
5	0.8861	5.538		76.358
6	0.5995	3.747		80.106
7	0.5462	3.414		83.520
8	0.5089	3.181		86.700
9	0.4498	2.811		89.511
10	0.3571	2.232		91.744
11	0.3354	2.096		93.840
12	0.2852	1.783		95.622
13	0.2451	1.532		97.154
14	0.1947	1.217		98.371
15	0.1596	0.998		99.369
16	0.1010	0.631		100.000

Figure 5-14: Output of the EFA on Section C: The Eigenvalues and % of Variance Explained by the Factors

Figure 5-15 shows the Scree Plot. In this figure, it indicates that the first three factors decline most steeply and this implies that the decision to keep the three factors is sufficient.

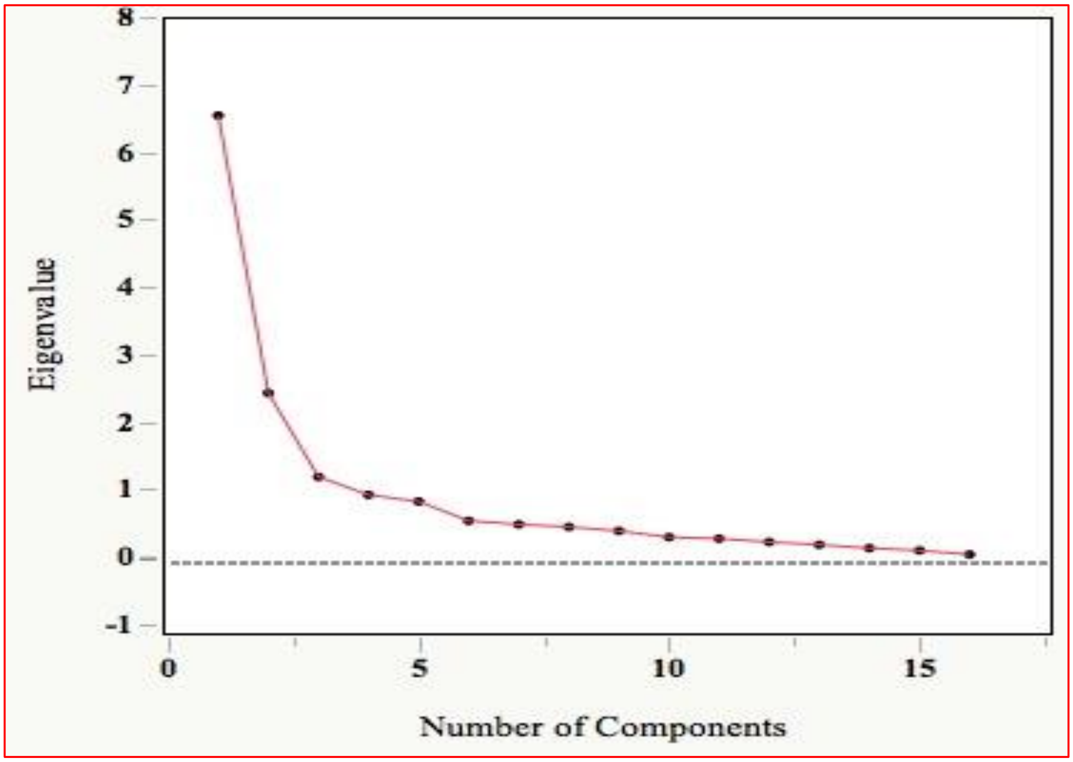


Figure 5-15: Output of the EFA on Section C: Scree Plot

Figure 5-16 is the output of the EFA on section C with emphasis on the rotated matrix with factor loadings. The factor loadings in Figure 5-16 which are beneath each column of the figure for each factor were inspected.

Statements	Factor 1	Factor 2	Factor 3
C1.1	0.31	0.20	0.18
C1.2	0.97	-0.1	-0.1
C1.3	0.64	0.35	0.13
C1.4	0.09	0.64	0.29
C2.1	0.43	0.53	0.17
C2.2	0.73	-0.0	0.05
C2.3	0.16	0.51	0.14
C2.4	0.04	0.63	0.18
C3.1	0.75	0.09	0.31

C3.2	0.20	0.34	0.41
C3.3	0.62	0.49	0.22
C3.4	0.04	0.34	0.56
C4.1	0.02	0.92	0.19
C4.2	0.60	0.23	0.43
C4.3	0.28	0.23	0.83
C4.4	0.61	0.40	0.36

Figure 5-16: Output of the EFA on Section C – The Rotated Matrix with Factor Loadings

Table 5-3 indicates how the questions are re-grouped in section C after performing EFA on section C of the questionnaire to ensure the validity of the sub-constructs. The following names for the factors are proposed that make logical and theoretical sense derived from the set of the statements (items) that make up the Factor: Factor 1: General data policies; Factor 2: Specific sensitive data policy; Factor 3: Access to sensitive data.

TABLE 5-3
Re-Grouping of Questions in Section C of the Questionnaire

Factor 1: General data policies	Factor 2: Specific sensitive data policy	Factor 3: Access to sensitive data
Question C1.2	Question C1.4	Question C3.2
Question C1.3	Question C2.1	Question C3.4
Question C2.2	Question C2.3	Question C4.3
Question C3.1	Question C2.4	
Question C3.3	Question C4.1	
Question C4.2		
Question C4.4		

5.3.3 Exploratory Factor Analysis of Section D of the Questionnaire

Figure 5-17 shows the output of the EFA for section D (the eigenvalues and the % of variance declared by the factors). If the value of the cumulative percentage is more than 60%, then this is considered to be good enough. The first two factors show a cumulative percentage of 64.419% therefore these two factors explain 64.419% of the variance in the original 22 items, and this is considered to be good. Therefore, the two factors were extracted as shown in Figure 5-19.

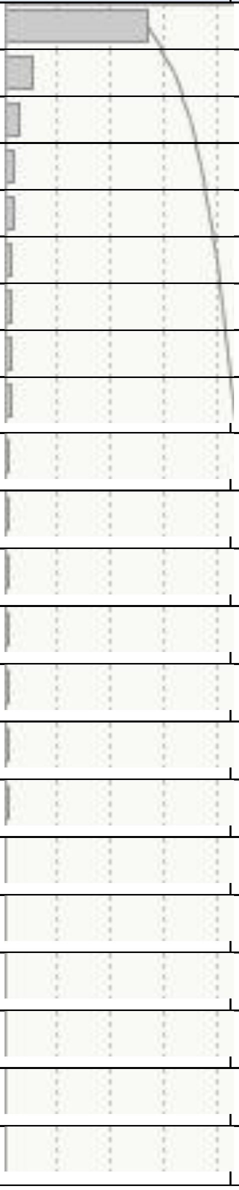
Number of Possible Factors	Eigenvalue	Percentage of Variance	Percentange	Cumulative Percentage
1	11.8417	53.826		53.826
2	2.3304	10.593		64.419
3	1.3315	6.052		70.471
4	0.7420	3.373		73.844
5	0.7340	3.336		77.180
6	0.6419	2.918		80.098
7	0.5298	2.408		82.506
8	0.4839	2.200		84.706
9	0.4529	2.059		86.764
10	0.4044	1.838		88.603
11	0.3968	1.804		90.406
12	0.3485	1.584		91.991
13	0.2727	1.239		93.230
14	0.2620	1.191		94.421
15	0.2424	1.102		95.523
16	0.2337	1.062		96.585
17	0.1763	0.801		97.386
18	0.1592	0.724		98.110
19	0.1326	0.603		98.712
20	0.1110	0.504		99.217
21	0.0869	0.395		99.612
22	0.0854	0.388		100.000

Figure 5-17: Output of the EFA for Section D: The Eigenvalues and % of Variance Declared by the Factors

Figure 5-18 illustrates the Scree Plot. In this figure, it can be observed that the first two factors decline the most steeply. This means that the decision will be to keep the two factors.

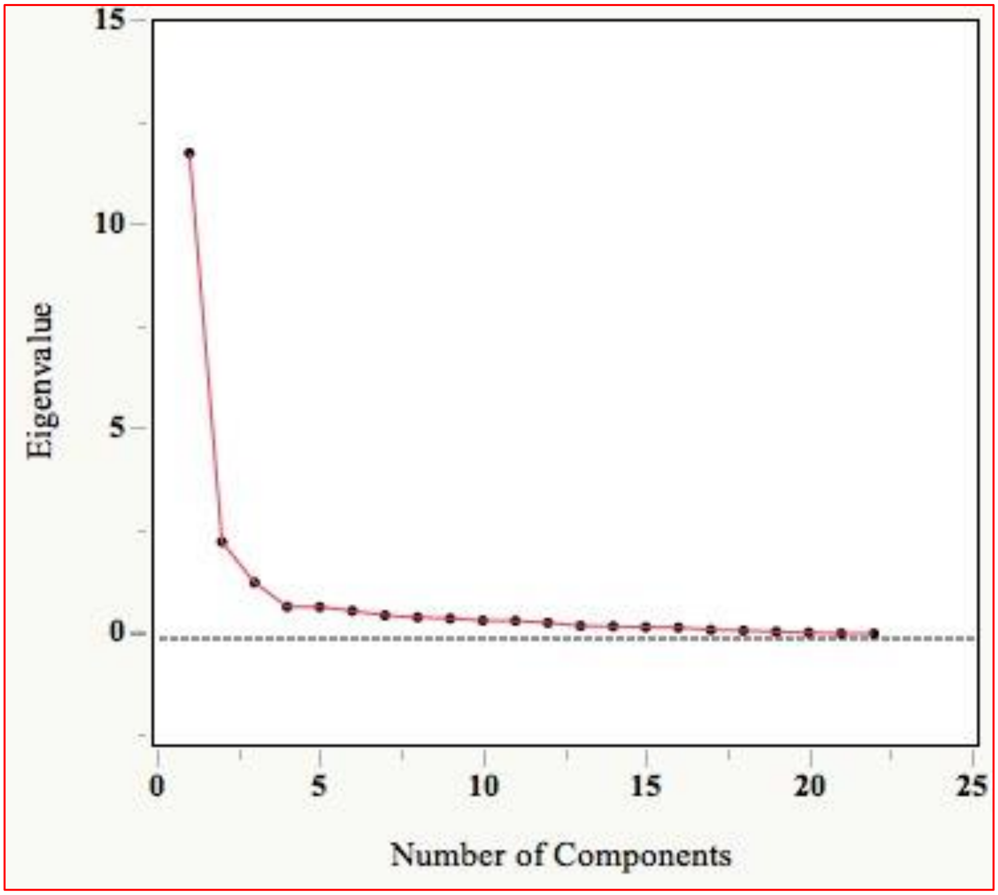
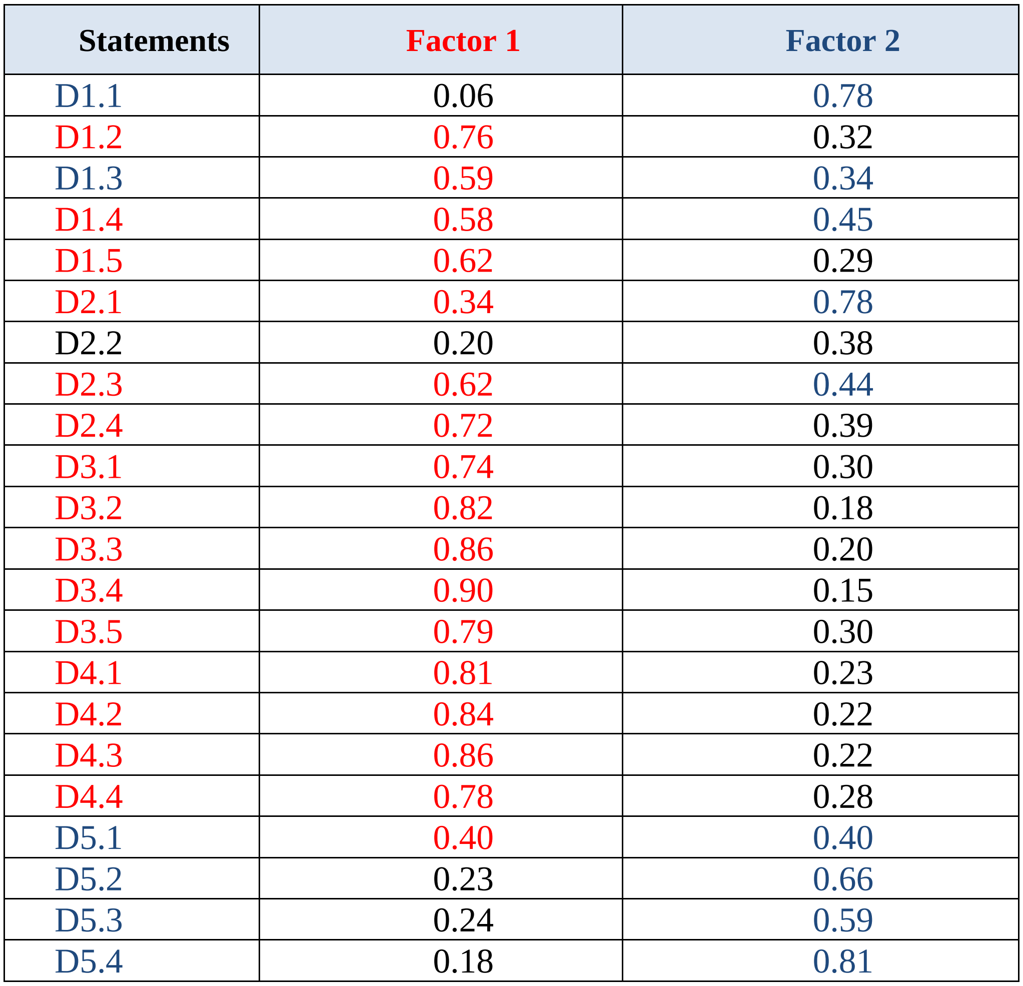


Figure 5-18: Output of the EFA on Section D: Scree Plot

Figure 5-19 represents the output of the EFA on section D with emphasis on the rotated matrix with factor loadings. This figure was used to determine the composition of the factors. The factor loadings in Figure 5-19 which are beneath each column of the figure for each factor were inspected. Question D2.2 has a score below 0.4 in both factors 1 and 2 and was not considered to be part of the re-grouping as depicted in Table 5-4.



Statements	Factor 1	Factor 2
D1.1	0.06	0.78
D1.2	0.76	0.32
D1.3	0.59	0.34
D1.4	0.58	0.45
D1.5	0.62	0.29
D2.1	0.34	0.78
D2.2	0.20	0.38
D2.3	0.62	0.44
D2.4	0.72	0.39
D3.1	0.74	0.30
D3.2	0.82	0.18
D3.3	0.86	0.20
D3.4	0.90	0.15
D3.5	0.79	0.30
D4.1	0.81	0.23
D4.2	0.84	0.22
D4.3	0.86	0.22
D4.4	0.78	0.28
D5.1	0.40	0.40
D5.2	0.23	0.66
D5.3	0.24	0.59
D5.4	0.18	0.81

Figure 5-19: Output of the EFA on Section D – The Rotated Matrix with Factor Loadings

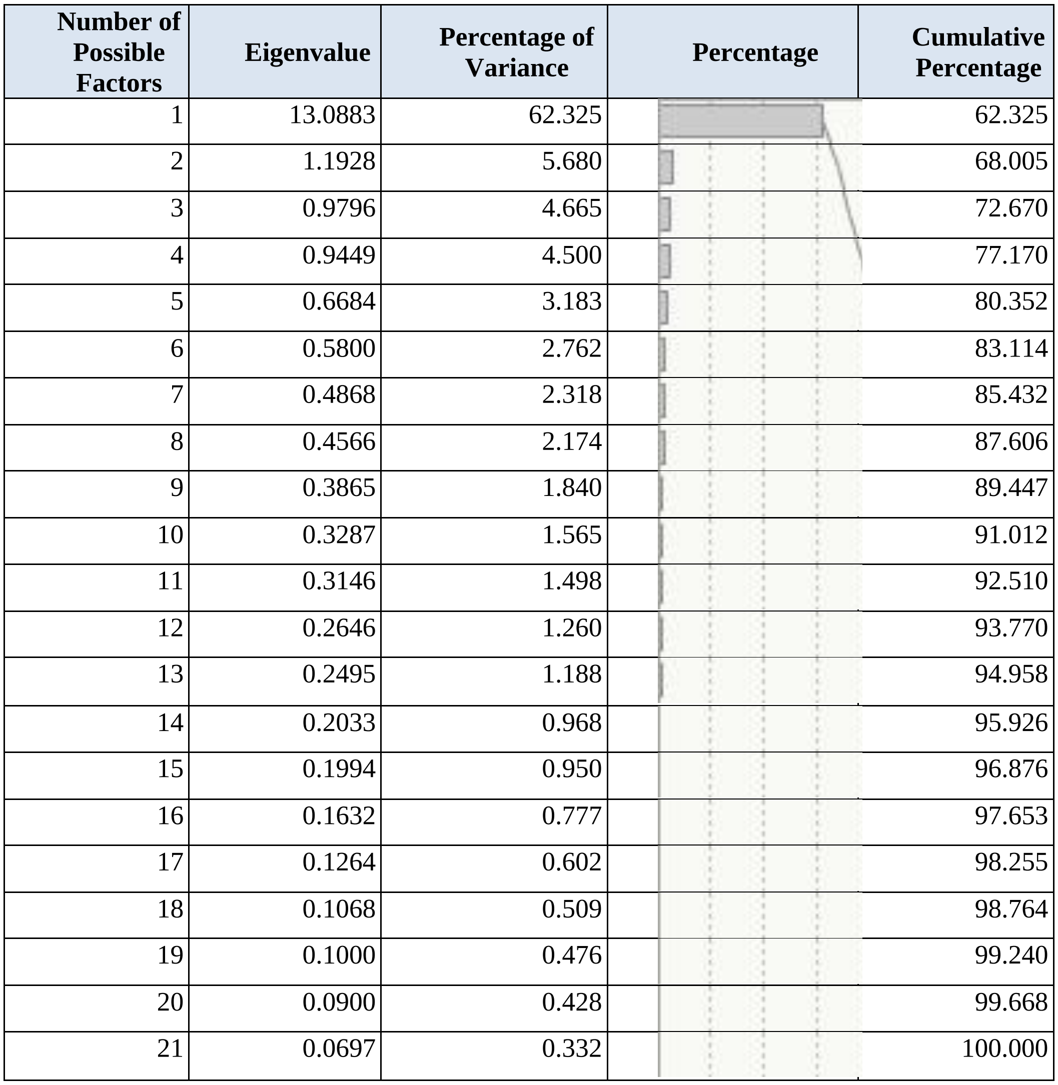
Table 5-4 illustrates how the questions were re-grouped in section D after performing an EFA on section D of the questionnaire to ensure validity of the sub-constructs. The following names for the factors are proposed that make logical and theoretical sense derived from the set of the statements (items) that make up the Factor: Factor 1: General data issues; Factor 2: Data security model.

TABLE 5-4
Re-Grouping of Questions in Section D of the Questionnaire

Factor 1: General data issues	Factor 2: Data security model
Question D1.2	Question D1.1
Question D1.3	Question D2.1
Question D1.4	Question D5.1
Question D1.5	Question D5.2
Question D2.3	Question D5.3
Question D2.4	Question D5.4
Question D3.1	
Question D3.2	
Question D3.3	
Question D3.4	
Question D3.5	
Question D4.1	
Question D4.2	
Question D4.3	
Question D4.4	

5.3.4 Exploratory Factor Analysis of Section E of the Questionnaire

Figure 5-20 shows the output of the EFA for section E (the eigenvalues and the % of variance declared by the factors). If the value of the cumulative percentage is more than 60%, then this is considered to be meaningful enough. The first two factors show a cumulative percentage of 68.005%, therefore these two factors explain 68.005% of the variance in the original 21 items and this is considered to be good. Therefore, these two factors were extracted as shown in Figure 5-22.



Number of Possible Factors	Eigenvalue	Percentage of Variance	Percentage	Cumulative Percentage
1	13.0883	62.325		62.325
2	1.1928	5.680		68.005
3	0.9796	4.665		72.670
4	0.9449	4.500		77.170
5	0.6684	3.183		80.352
6	0.5800	2.762		83.114
7	0.4868	2.318		85.432
8	0.4566	2.174		87.606
9	0.3865	1.840		89.447
10	0.3287	1.565		91.012
11	0.3146	1.498		92.510
12	0.2646	1.260		93.770
13	0.2495	1.188		94.958
14	0.2033	0.968		95.926
15	0.1994	0.950		96.876
16	0.1632	0.777		97.653
17	0.1264	0.602		98.255
18	0.1068	0.509		98.764
19	0.1000	0.476		99.240
20	0.0900	0.428		99.668
21	0.0697	0.332		100.000

Figure 5-20: Output of the EFA for Section E – The Eigenvalues and % of Variance Explained by the Factors

Figure 5-21 illustrates the Scree Plot. In this figure, it can be seen that the first two factors decline most steeply. This means that the decision will be to keep the two factors.

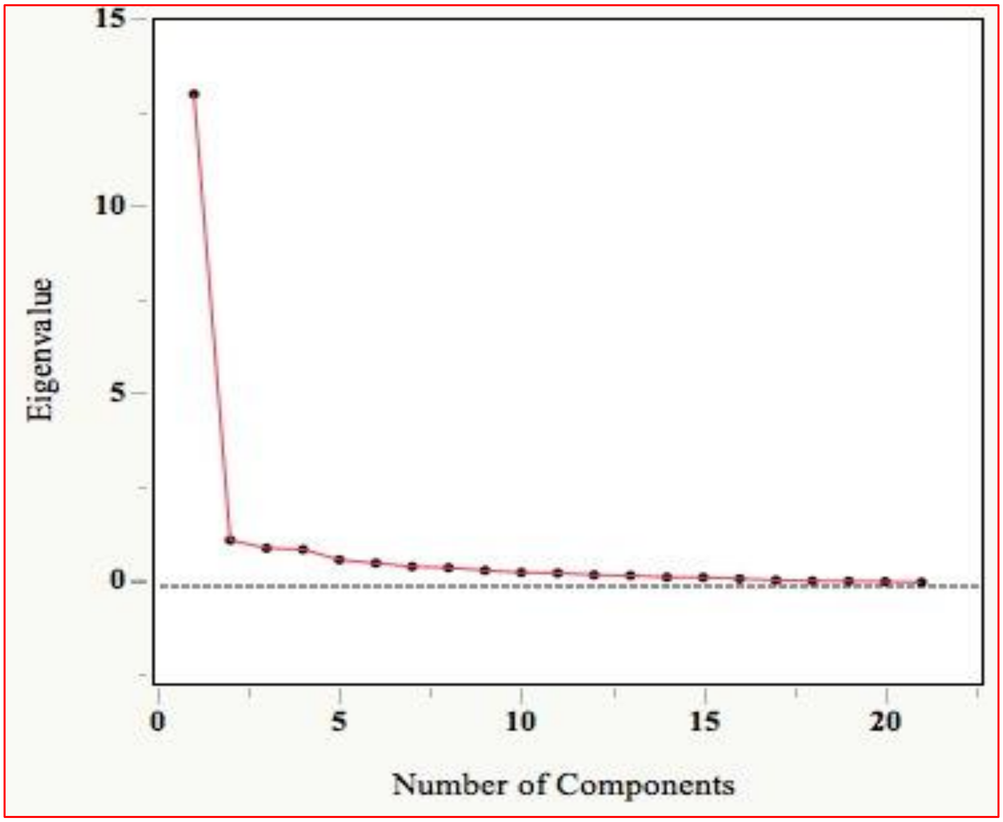


Figure 5-21: Output of the EFA on Section E – Scree Plot

Figure 5-22 represents the output of the EFA on section E with emphasis on the rotated matrix with factor loadings. This figure is used to determine the composition of the factors. The factor loadings in Figure 5-22 which are beneath each column of the figure for each factor were inspected.

Statements	Factor 1	Factor 2
E1.1	0.23	0.84
E1.2	0.46	0.77
E1.3	0.50	0.67
E1.4	0.38	0.75
E2.1	0.37	0.29
E2.2	0.58	0.51
E2.3	0.54	0.52
E2.4	0.74	0.39
E2.5	0.61	0.55

E3.1	0.74	0.45
E3.2	0.59	0.45
E3.3	0.68	0.46
E3.4	0.73	0.44
E4.1	0.79	0.32
E4.2	0.31	0.55
E4.3	0.55	0.50
E4.4	0.80	0.40
E5.1	0.86	0.28
E5.2	0.78	0.32
E5.3	0.59	0.47
E5.4	0.66	0.50

Figure 5-22: Output of the EFA on Section E: The Rotated Matrix with Factor Loadings

Table 5-5 shows how the questions are re-grouped in section E after performing the EFA on section E of the questionnaire to ensure validity of the sub-constructs. The following names for the factors are proposed that make logical and theoretical sense derived from the set of the statements (items) that make up the Factor: Factor 1: General control; Factor 2: Migration planning.

TABLE 5-5
Re-Grouping of Questions in Section E of the Questionnaire

Factor 1: General control	**Factor 2: Migration planning**
Question E2.2	Question E1.1
Question E2.3	Question E1.2
Question E2.4	Question E1.3
Question E2.5	Question E1.4
Question E3.1	Question E4.2
Question E3.2	
Question E3.3	
Question E3.4	
Question E4.1	
Question E4.3	
Question E4.4	
Question E5.1	
Question E5.2	

Question E5.3	
Question E5.4	

5.4 Reliability Analysis of the Ten Sub-Constructs

Reliability is the degree of consistency of a measuring instrument, that is, the ability of a measuring instrument to assess consistently and determines whether the instrument measures anything (Ritter 2010; Tavakol & Dennick 2011; Kerlinger & Lee 2000). Reliability is defined as the consistency of the constructs of a measuring instrument (Wiid & Diggines 2013). Reliability analysis is the method of testing the reliability of the constructs in the questionnaire. The difference between validity and reliability is that a measuring instrument cannot be regarded as valid unless it is reliable, however, reliability of a measuring instrument does not hinge on its validity (Tavakol & Dennick 2011).

There are various types of reliability coefficients but the Cronbach's Alpha is one of the most widely used reliability coefficients (Hogan *et al.* 2000). It was Lee Cronbach in 1951 who developed the Cronbach's Alpha in order to provide a measure of the internal consistency of a measuring instrument and its value is expressed as a number between 0 and 1 (Cronbach 1951). The researcher ensured that all the questions of each construct in the questionnaire were positively stated, otherwise, they were recoded.

The researcher used the SAS Software to perform a reliability analysis on the results of the questionnaire, and the analysis produced a Cronbach's Alpha value (α) which is interpreted in accordance with Wiid and Diggines (2013). For an α value:

- above 0.8, reliability is considered to be good
- between 0.6 and 0.8, reliability is considered to be acceptable

- below 0.6, reliability is considered unacceptable.

According to Wiid and Diggines (2013), a reliable Cronbach's Coefficient Alpha value validates that the individual items of a construct measured the same concept in the same manner (or consistently). The results of the reliability analysis of the new constructs obtained as a result of the factor analysis on the original questionnaire follow in Table 5-6 below:

TABLE 5-6
Reliability Analysis Results of the Sub-Constructs

Variables	**Items**	**Cronbach Alpha**	**Reliability**
Sub-Construct 1: Awareness Accountability score or (Employee_awareness/ information handling/accountability)	B1.1;B1.2;B3.1	0.7033	Acceptable
Sub-Construct 2: Training handling or (Employee_course type/sensitivity classification)	B2.1;B2.4;B3.2; B3.4	0.8443	Good
Sub-Construct 3: Consequences of sensitive data or (Employee_Training/Info Non-protection consequences)	B1.3;B1.4;B2.2; B2.3	0.6265	Acceptable
Sub-Construct 4: General data policies or (Organisation_strategy/culture/ communication/data)	C1.2;C1.3;C2.2; C3.1;C4.2;C4.4	0.8922	Good
Sub-Construct 5: Specific sensitive data policy or (Organisation_data security policy/sensitive info identification)	C1.4;C2.1;C2.3; C2.4;C4.1	0.8342	Good
Sub-Construct 6: Access to sensitive data or (Data_access/controls/standards enforcement)	C3.2;C3.4;C4.3	0.7046	Acceptable

Sub-Construct 7: General data issues or (Employee_roles/responsibilities)	D1.2;D1.3;D1.4; D1.5;D2.4;D3.1; D3.2;D3.4;D3.5; D4.1;D4.2;D4.3; D4.4	0.9658	Good
Sub-Construct 8: Data security model or (Organisation_security models)	D1.1;D2.1;D5.1; D5.3;D5.4	0.8630	Good
Sub-Construct 9: General control or (Monitor/control_tools/migration issues/risk assessment/migration duration/network bandwidth)	E2.2;E2.3;E2.4; E2.5;E3.1;E3.3; E3.4;E4.1;E4.3; E5.1;E5.2;E5.3; E5.4	0.9647	Good
Sub-Construct 10: Migration planning or (Migration processes_application identification/time management/servers de-staging/source data backup/data quality)	E1.1;E1.2;E1.3; E1.4;E4.2	0.8975	Good

Table 5-6 illustrates the summarised items for the ten sub-constructs. Estimates of internal consistency as measured by Cronbach’s alpha all exceeded 0.80 with the exception of three constructs that are less than 0.70 and are reported in Table 5-6. This indicates good reliability for the seven constructs that have Cronbach’s alpha exceeding 0.80.

5.5 Description of the Sub-Constructs

In this section, the descriptive statistics starts with the description of the sub-constructs. The analysis of the responses to questions in sub-constructs 1 to 10 are illustrated in Table 5-7 to Table 5-16 respectively.

TABLE 5-7

Analyses of Responses to Questions in Sub-Construct 1

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
B1.2	1.11%	1	2.22%	2	12.22%	11	24.44%	22	60.00%	54
B1.1	1.11%	1	0.00%	0	4.44%	4	25.56%	23	68.89%	62
B3.1	0.00%	0	4.44%	4	10.00%	9	20.00%	18	65.56%	59

Sub-Construct 1 applies to awareness accountability score or

awareness accountability score or employee_awareness/information

handling/accountability with three questions (B1.2; B1.1 and B3.1). The questions relate to the employee awareness; employee information handling and employee accountability. The analyses of responses to questions in sub-construct 1 are shown in Table 5-7. Statement B1.1 is the most important one because it has the highest percentage of Agree and Strongly Agree from the responses of the respondents.

TABLE 5-8

Analyses of Responses to Questions in Sub-Construct 2

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
B2.1	4.44%	4	6.67%	6	14.44%	13	16.67%	15	57.78%	52
B2.4	3.33%	3	8.89%	8	14.44%	13	25.56%	23	47.78%	43
B3.2	1.11%	1	7.78%	7	13.33%	12	16.67%	15	61.11%	55
B3.4	2.22%	2	2.22%	2	11.11%	10	31.11%	28	53.33%	48

Sub-Construct 2 refers to training handling or employee_course type/sensitivity classification with four questions (B2.1; B2.4; B3.2 and B3.4). The questions relate to the employee course type; employee sensitivity classification. The analyses of responses to questions in sub-construct 2 are revealed in Table 5-8. Statement B3.2 is the most important one because it has the highest percentage of Agree and Strongly Agree from the responses of the respondents.

TABLE 5-9
Analyses of Responses to Questions in Sub-Construct 3

	Strongly Disagree		**Disagree**		**Neutral**		**Agree**		**Strongly Agree**	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
B1.3	0.00%	0	2.22%	2	12.22%	11	25.56%	23	60.00%	54
B1.4	0.00%	0	1.11%	1	6.67%	6	17.78%	16	74.44%	67
B2.2	0.00%	0	0.00%	0	4.44%	4	18.89%	17	76.67%	69
B2.3	0.00%	0	0.00%	0	1.11%	1	14.44%	13	84.44%	76

Sub-Construct 3 relates to consequences of sensitive data or

employee_training/info non-protection consequences with four questions (B1.3; B1.4; B2.2 and B2.3). The questions relate to the employee training/employee information non-protection consequences. The analyses of responses to questions in sub-construct 3 are shown in Table 5-9. Statement B2.3 is the most important one because it has the highest percentage of Agree and Strongly Agree from the responses of the respondents.

TABLE 5-10
Analyses of Responses to Questions in Sub-Construct 4

	Strongly Disagree		**Disagree**		**Neutral**		**Agree**		**Strongly Agree**	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
C1.2	0.00%	0	2.22%	2	5.56%	5	15.56%	14	76.67%	69
C1.3	1.11%	1	3.33%	3	5.56%	5	17.78%	16	72.22%	65
C2.2	0.00%	0	1.11%	1	4.44%	4	22.22%	20	72.22%	65

C3.1	0.00%	0	1.11%	1	8.89%	8	22.22%	20	67.78%	61
C3.3	1.11%	1	7.78%	7	20.00%	18	18.89%	17	52.22%	47
C4.2	0.00%	0	4.44%	4	11.11%	10	16.67%	15	67.78%	61
C4.4	1.11%	1	4.44%	4	12.22%	11	16.67%	15	65.56%	59

Sub-Construct 4 refers to general data policies or organisation_strategy/culture/ communication/data with seven questions (C1.2; C1.3; C2.2; C3.1; C4.2 and C4.4). The questions relate to the organisational strategy; organisational culture and organisational data. The analyses of responses to questions in sub-construct 4 are shown in Table 5-10. Statement C2.2 is the most important one because it has the highest percentage of Agree and Strongly Agree from the responses of the respondents.

TABLE 5-11
Analyses of Responses to Questions in Sub-Construct 5

	Strongly Disagree		**Disagree**		**Neutral**		**Agree**		**Strongly Agree**	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
C1.4	3.33%	3	1.11%	1	12.22%	11	18.89%	17	64.44%	58
C2.1	0.00%	0	6.67%	6	11.11%	10	27.78%	25	54.44%	49
C2.3	1.11%	1	3.33%	3	20.00%	18	20.00%	18	55.56%	50
C2.4	1.11%	1	5.56%	5	15.56%	14	22.22%	20	55.56%	50
C4.1	3.33%	3	6.67%	6	10.00%	9	25.56%	23	54.44%	49

Sub-Construct 5 relates to specific sensitive data policy or organisation_data security policy/sensitive info identification with five questions (C1.4; C2.1; C2.3; C2.4 and C4.1). The questions relate to the organisation data security policy and organisation sensitive information identification. The analyses of responses to questions in sub-construct 5 are summarised in Table 5-11. Statement C1.4 is the most important one because it has the highest percentage of Agree and Strongly Agree from the responses of the respondents.

TABLE 5-12
Analyses of Responses to Questions in Sub-Construct 6

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
C3.2	0.00%	0	0.00%	0	6.67%	6	16.67%	15	76.67%	69
C3.4	2.22%	2	7.78%	7	10.00%	9	16.67%	15	63.33%	57
C4.3	0.00%	0	2.22%	2	7.78%	7	24.44%	22	65.56%	59

Sub-Construct 6 refers to access to sensitive data or data_access/controls/standards enforcement with three questions (C3.2; C3.4 and C4.3). The questions relate to the data access, data controls and data standards enforcement. The analyses of responses to questions in sub-construct 6 are revealed Table 5-12. Statement C3.2 is the most important one because it has the highest percentage of Agree and Strongly Agree from the responses of the respondents.

TABLE 5-13
Analyses of Responses to Questions in Sub-Construct 7

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
D1.2	2.22%	2	3.33%	3	8.89%	8	25.56%	23	60.00%	54
D1.3	0.00%	0	5.56%	5	13.33%	12	25.56%	23	55.56%	50
D1.4	1.11%	1	2.22%	2	14.44%	13	22.22%	20	60.00%	54
D1.5	1.11%	1	4.44%	4	18.89%	17	30.00%	27	45.56%	41
D2.3	2.22%	2	6.67%	6	18.89%	17	26.67%	24	45.56%	41
D2.4	2.22%	2	4.44%	4	15.56%	14	22.22%	20	55.56%	50
D3.1	1.11%	1	4.44%	4	15.56%	14	23.33%	21	55.56%	50

D3.2	1.11%	1	10.00%	9	10.00%	9	24.44%	22	54.44%	49
D3.3	3.33%	3	5.56%	5	14.44%	13	20.00%	18	56.67%	51
D3.4	3.33%	3	7.78%	7	13.33%	12	21.11%	19	54.44%	49
D3.5	4.44%	4	4.44%	4	18.89%	17	34.44%	31	37.78%	34
D4.1	2.22%	2	5.56%	5	16.67%	15	13.33%	12	62.22%	56
D4.2	3.33%	3	6.67%	6	12.22%	11	31.11%	28	46.67%	42
D4.3	2.22%	2	8.89%	8	10.00%	9	27.78%	25	51.11%	46
D4.4	1.11%	1	7.78%	7	12.22%	11	23.33%	21	55.56%	50

Sub-Construct 7 summarises general data issues or employee_roles/responsibilities with 15 questions (D1.2; D1.3; D1.4; D1.5; D2.4; D3.1; D3.2; D3.4; D3.5; D4.1; D4.2; D4.3 and D4.4). The questions relate to the employee roles and employee responsibilities. The analyses of responses to questions in sub-construct 7 are shown in Table 5-13. Statements D1.2 and D1.4 are the most important ones because they have the highest percentage of Agree and Strongly Agree from the respondents.

TABLE 5-14
Analyses of Responses to Questions in Sub-Construct 8

	Strongly Disagree		**Disagree**		**Neutral**		**Agree**		**Strongly Agree**	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
D1.1	0.00%	0	2.22%	2	6.67%	6	28.89%	26	62.22%	56
D2.1	0.00%	0	2.22%	2	12.22%	11	28.89%	26	56.67%	51
D5.1	0.00%	0	1.11%	1	4.44%	4	24.44%	22	70.00%	63
D5.2	0.00%	0	0.00%	0	5.56%	5	35.56%	32	58.89%	53
D5.3	0.00%	0	1.11%	1	5.56%	5	28.89%	26	64.44%	58
D5.4	0.00%	0	2.22%	2	6.67%	6	35.56%	32	55.56%	50

Sub-Construct 8 relates to data security model or organisation_security models with six questions (D1.1; D2.1; D5.1; D5.3 and D5.4). The questions relate to the organisational security models. The analyses of responses to questions in sub-construct 8 are revealed in Table 5-14. Statement D5.1 is the most important one because it has the highest percentage of Agree and Strongly Agree from the responses of the respondents.

TABLE 5-15
Analyses of Responses to Questions in Sub-Construct 9

	Strongly Disagree		**Disagree**		**Neutral**		**Agree**		**Strongly Agree**	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	
E2.2	0.00%	0	0.00%	0	12.22%	11	21.11%	19	66.67%	60
E2.3	0.00%	0	2.22%	2	7.78%	7	23.33%	21	66.67%	60
E2.4	1.11%	1	3.33%	3	10.00%	9	23.33%	21	62.22%	56
E2.5	0.00%	0	5.56%	5	7.78%	7	21.11%	19	65.56%	59
E3.1	1.11%	1	4.44%	4	16.67%	15	12.22%	11	65.56%	59
E3.2	2.22%	2	6.67%	6	21.11%	19	28.89%	26	41.11%	37
E3.3	6.67%	6	4.44%	4	16.67%	15	16.67%	15	55.56%	50
E3.4	5.56%	5	5.56%	5	13.33%	12	12.22%	11	63.33%	57
E4.1	2.22%	2	4.44%	4	15.56%	14	25.56%	23	52.22%	47
E4.3	1.11%	1	1.11%	1	13.33%	12	27.78%	25	56.67%	51
E4.4	2.22%	2	4.44%	4	15.56%	14	21.11%	19	56.67%	51
E5.1	1.11%	1	7.78%	7	13.33%	12	24.44%	22	53.33%	48
E5.2	2.22%	2	1.11%	1	12.22%	11	31.11%	28	53.33%	48
E5.3	1.11%	1	1.11%	1	12.22%	11	35.56%	32	50.00%	45
E5.4	1.11%	1	2.22%	2	16.67%	15	24.44%	22	55.56%	50

Sub-Construct 9 relates to general control or monitor/control_tools/migration issues/risk assessment/migration duration/network bandwidth with fifteen questions (E2.2; E2.3; E2.4; E2.5; E3.1; E3.3; E3.4; E4.1; E4.3; E5.1; E5.2; E5.3 and E5.4). The questions relate to the monitoring/control of tools; monitoring/control of migration issues; monitoring/control of risk assessment; monitoring/control of migration duration and monitoring/control of network bandwidth. The analyses of responses to questions in sub-construct 9 are highlighted in Table 5-15. Statements E2.2 and E2.3 are the most important ones because they have the highest percentage of Agree and Strongly Agree from the responses of the respondents.

TABLE 5-16
Analyses of Responses to Questions in Sub-Construct 10

	Strongly Disagree		**Disagree**		**Neutral**		**Agree**		**Strongly Agree**	
Statements	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**	**% of Total**	**N**
E1.1	1.11%	1	1.11%	1	3.33%	3	22.22%	20	72.22%	65
E1.2	1.11%	1	2.22%	2	6.67%	6	17.78%	16	72.22%	65
E1.3	3.33%	3	5.56%	5	5.56%	5	14.44%	13	71.11%	64
E1.4	1.11%	1	2.22%	2	4.44%	4	16.67%	15	75.56%	68
E4.2	1.11%	1	0.00%	0	7.78%	7	23.33%	21	67.78%	61

Sub-Construct 10 refers to migration planning or migration processes_application identification/time management/servers de-staging/source data backup/data quality with five questions (E1.1; E1.2; E1.3; E1.4 and E4.2). The questions relate to the following processes: migration processes involve application identification; migration processes involve time management; migration processes involve server de-staging; migration processes involve source data backup and migration processes involve data quality. The analyses of responses to questions in sub-construct 10 are developed in Table 5-16. Statement E1.1 is the most important one because it has the highest percentage of Agree and Strongly Agree from the responses of the respondents.

5.6 Calculation and Comparison of the Sub-Construct Scores

In this section, the means and the standard deviation of these sub-constructs are calculated and used to perform the correlation analysis between the sub-constructs. To calculate a sub-construct score, the average of the reliable items that loaded onto the sub-construct (factor) were taken. For example, the Awareness accountability score was calculated as the average of items B1.1, B1.2 and B3.1. The following interpretation was used to interpret the mean score: a mean score towards 1 indicates strongly disagree while a mean score towards 5 indicates strongly agree.

The comparisons among the sub-constructs with respect to the means and the standard deviations of the sub-constructs are shown in Table 5-17. The sub-construct B Consequences of sensitive data or (Employee_Training/Info Non-Protection Consequences) is the most important considered sub-construct by the respondents with a mean of 4.66 (towards stongly agree).

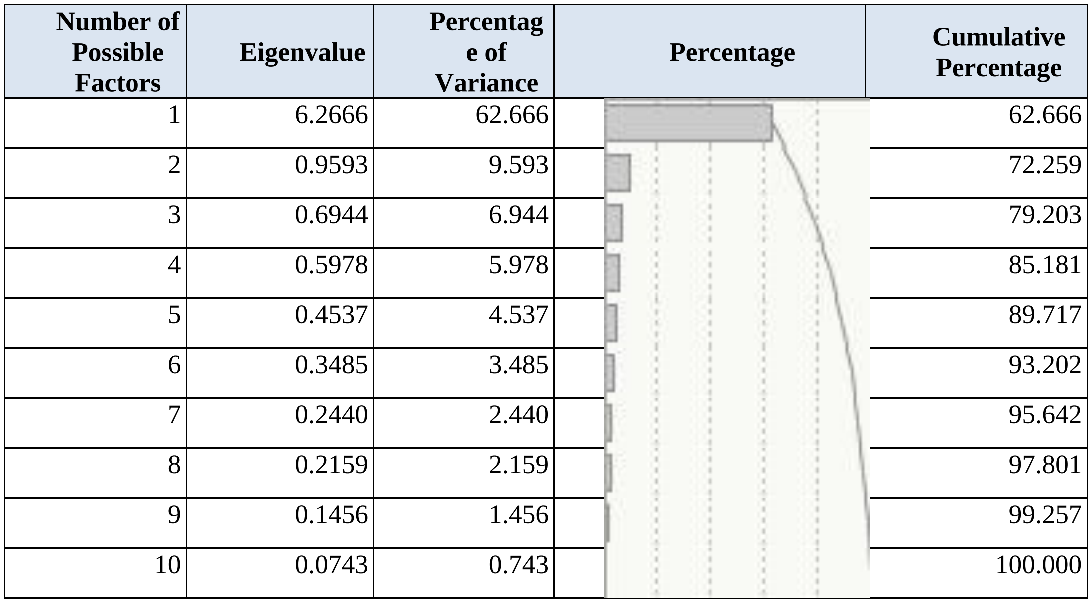
TABLE 5-17
Comparison of the Sub-Constructs

Sub-Constructs	Mean	Std Dev
B Awareness accountability score or (Employee_Awareness/Information Handling/Accountability)	4.49	0.64
B Training handling or (Employee_Course Type/Sensitivity Classification)	4.21	0.88

B Consequences of sensitive data or (Employee_Training/Info Non-Protection Consequences)	4.66	0.42
C General data policies or (Organisation_Strategy/Culture/Communication/Data)	4.50	0.65
C Specific sensitive data policy or (Organisation_Data Security Policy/Sensitive Info Identification)	4.28	0.77
C Access to sensitive data or (Data_Access/Controls/Standards Enforcement)	4.51	0.66
D General data issues or (Employee_Roles/Responsibilities)	4.21	0.84
D Data security model or (Organisation_Security Models)	4.51	0.53
E General control or (Monitor/Control_Tools/Migration Issues/Risk Assessment/Migration Duration/Network Bandwidth)	4.31	0.78
E Migration planning or (Migration Processes_Application Identification/Time Management/Servers De-Staging/Source Data Back Up/Data Quality)	4.57	0.69

5.7 Exploratory Factor Analysis of the Ten Sub-Constructs

In order to develop the management framework, EFA was performed on the ten sub-constructs to obtain the final main constructs. Figure 5-23 shows the output of the EFA for the ten sub-constructs in Table 5-17 (the eigenvalues and the % of variance declared by the factors). The first factor shows a cummulative percentage of 62.666%. The analysis shows only one factor explaining 62.67% of the data, which shows that all the sub-constructs are related to each other. However, the researcher wanted to explore the structure of the sub-constructs in order to develop the management framework. Therefore, the first four factors are considered because the factor loadings make logical and theoretical sense, since one factor cannot be considered only. Therefore, the four factors are extracted as shown in Figure 5-23.



Number of Possible Factors	Eigenvalue	Percentage of Variance	Percentage	Cumulative Percentage
1	6.2666	62.666		62.666
2	0.9593	9.593		72.259
3	0.6944	6.944		79.203
4	0.5978	5.978		85.181
5	0.4537	4.537		89.717
6	0.3485	3.485		93.202
7	0.2440	2.440		95.642
8	0.2159	2.159		97.801
9	0.1456	1.456		99.257
10	0.0743	0.743		100.000

Figure 5-23: Output of the EFA for the Ten Sub-Constructs: The Eigenvalues and % of Variance Declared by the Factors

Figure 5-24 illustrates the Scree Plot. In this figure, it can be seen that the first four factors decline most steeply. This means that the decision will be to keep the four factors.

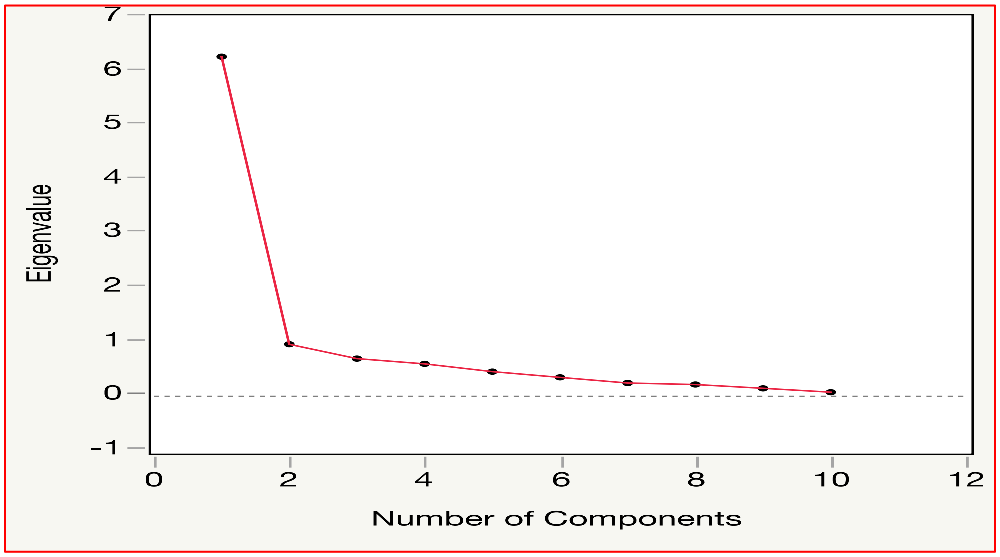


Figure 5-24: Output of the EFA on the Ten Sub-Constructs – Scree Plot

Figure 5-25 shows the output of the EFA on the ten sub-constructs with emphasis on the rotated matrix with factor loadings which is used to determine the composition of the factors.

Sub-Constructs	Factor 1	Factor 2	Factor 3	Factor 4
B1 Awareness accountability score	0.03	**0.67**	0.49	0.19
B2 Training handling	0.54	**0.59**	0.22	0.20
B3 Consequences of sensitive data	**0.43**	0.29	0.28	0.42
C4 General data policies	0.29	0.23	**0.92**	0.16
C5 Specific sensitive data policy	0.47	**0.71**	0.15	0.24
C6 Access to sensitive data	**0.77**	0.11	0.23	0.20
D7 General data issues	**0.71**	0.40	0.29	0.33
D8 Data security model	0.32	0.27	**0.47**	0.26
E9 General control	**0.75**	0.30	0.21	0.48
E10 Migration planning	0.37	0.24	0.22	**0.87**

Figure 5-25: Output of the EFA on the Ten Sub-Constructs: The Rotated Matrix with Factor Loadings

Table 5-18 shows how the ten sub-constructs are re-grouped within the four factors after performing the EFA to ensure validity of the sub-constructs. The following names for the factors are proposed that make logical and theoretical sense derived from the set of the sub-constructs that make up the Factor: Factor 1: Sensitive Data Management; Factor 2: Sensitive Data Awareness; Factor 3: Data Governance: Factor 4: Migration Planning. These four factors now constitute the main constructs that will be used to develop the structure of the preliminary management framework.

TABLE 5-18
Re-Grouping of the Ten Sub-Constructs within the Four Factors

Factor 1: Sensitive Data Management	Factor 2: Sensitive Data Awareness	Factor 3: Data Governance	Factor 4: Migration Planning
B3	B1	C4	E10
C6	B2	D8	
D7	C5		
E9			

In Table 5-18, Factor 1 comprises of sub-constructs that mostly pertain to sensitive data management, while Factor 2 comprises of sub-constructs that belong to sensitive data awareness. Factor 3 comprises of sub-constructs that pertain to data governance while Factor 4 can be said to belong to migration planning. Consequently, Factors 1, 2, 3 and 4 are labelled sensitive data management, sensitive data awareness, data governance and migration planning consecutively.

5.8 Reliability of the Four Main Constructs

The reliability analysis results of the four main constructs (Factors 1 to 4) are shown in Table 5-20 and the table indicates that all the four factors have higher Cronbach Alphas. This implies that all the four factors are highly reliable.

TABLE 5-20
Reliability Analysis Results of the Four Main Constructs

Factors	Sub-Constructs	Cronbach Alpha
Factor 1: Sensitive data management	B3; C6; D7; E9	0.8945
Factor 2: Sensitive data awareness	B1; B2; C5	0.8364
Factor 3: Data Governance	C4; D8	0.7623

Factor 4: Migration Planning	E10	0.7623

5.9 The Distributions of the Four Main Constructs

The distributions of the four main constructs (Factors 1 to 4) are illustrated in Table 5-19 and the Table shows the Mean, Standard Deviation, Standard Error Mean, Skewness and Kurtosis of the four factors. The average of the sub-constructs is used to calculate the means of the factors.

TABLE 5-19
Distributions of the Four Main Constructs

Factors	Mean	Std Dev	Std Err Mean	Skewness	Kurtosis
Factor 1: Sensitive data management	4.4243	0.6056	0.0638	1.3814	1.0641
Factor 2: Sensitive data awareness	4.3275	0.6685	0.0704	1.2282	1.1098
Factor 3: Data Governance	4.5058	0.5319	0.0561	1.1104	0.3567
Factor 4: Migration Planning	4.5711	0.6938	0.0731	2.6022	8.2783

It can be inferred from Table 5-19 that the data of the four factors are concentrated around the mean since they all have lower standard deviations. The standard deviation implies that the views are concentrated around the mean. Skewness is a measure of lack of symmetry and kurtosis is a measure of whether the data is flat or peaked near the mean. The skewness of a normal distribution is zero (0), which implies that any symmetric data will have a skewness near zero (0). The skewness of all the four factors are more than one (1), therefore this implies that their data are skewed right, that is, the right tail is longer relative to the left tail. It also means that the data are not normally distributed, therefore Pearson's correlation cannot be used. Consequently the Spearman's correlation is used to determine their correlations. Spearman's correlation does not assume normality, and it is a non-parametric technique.

The kurtosis for a standard normal distribution is 3 and for Factors 1, 2 and 3, their kurtosis values are less than 3 which implies that they do not have too much peak in their data distributions. However, Factor 4 has a kurtosis value of 8.2783, therefore it has a wider peak than the other factors. The standard error of the mean and the standard error of the estimate are the most commonly used standard error statistics. The standard error mean of all the factors have lower values and this indicates more precise estimates of their population mean.

5.10 Multivariate Correlations of the Four Main Constructs

The significance of a Spearman's correlation is determined by the p-value. If the p-value is lower than 0.05 the correlation is significant at a 95% level of confidence. The strength of the relationship between two constructs is measured by the correlation. The following criteria prevail for the strength of two correlated constructs: If the significant probability $p \leq 0.05$ and the spearman's correlation coefficient (r) is:

- 1, then the two constructs are perfectly correlated.
- 0.3, then there exists a weak correlation between the two constructs.
- 0.5, then there exists a medium correlation between the two constructs.
- 0.7, then there exists a strong correlation between the two constructs.
- 0, then there is no correlation between the two constructs.

5.10.1 Testing the Relationships between the Four Factors using Hypotheses

A number of hypotheses will be tested from the relationships among the four factors. The hypotheses are labelled as H1, H2, H3, H4, H5 and H6 respectively and for each test the usual null hypothesis and the negation thereof. Subsequently the researcher investigates:

$H1_0$ (Null Hypothesis): There is no significant correlation between sensitive data management (factor 1) and sensitive data awareness (factor 2).

$H1_1$ (Research Hypothesis): There is a significant correlation between sensitive data management (factor 1) and sensitive data awareness (factor 2).

$H2_0$: There is no significant correlation between sensitive data management (factor 1) and data governance (factor 3).

$H2_1$: There is a significant correlation between sensitive data management (factor 1) and data governance (factor 3).

$H3_0$: There is no significant correlation between data management (factor 1) and migration planning (factor 4).

$H3_1$: There is a significant correlation between data management (factor 1) and migration planning (factor 4).

$H4_0$: There is no significant correlation between sensitive data awareness (factor 2) and data governance (factor 3).

$H4_1$: There is a significant correlation between sensitive data awareness (factor 2) and data governance (factor 3).

$H5_0$: There is no significant correlation between sensitive data awareness (factor 2) and migration planning (factor 4).

$H5_1$: There is a significant correlation between sensitive data awareness (factor 2) and migration planning (factor 4).

$H6_0$: There is no significant correlation between data governance (factor 3) and migration planning (factor 4).

$H6_1$: There is a significant correlation between data governance (factor 3) and migration planning (factor 4).

The spearman's correlation analysis will be used to test these hypotheses indicated above. Table 5-21 depicts the non-parametric Spearman's correlation coefficient (r) of the four main constructs and Spearman's correlation is used to identify if two constructs (variables) are related in a monotonic function, that is, when one value increases, so does the other one, and vice versa. The strength of the relationships between the variables will be assessed by the Spearman's correlation coefficients given above. Correlation coefficients vary from -1 to +1, and a correlation coefficient near 0 indicates a weak or no linear correlation. A correlation towards 1 indicates a strong positive linear correlation. A correlation near -1 is indicative of a strong negative linear correlation. Table 5-21 provides the (strength of the) correlations as well as the p-values that provide the significance of the correlations.

TABLE 5-21
Non-Parametric: Spearman's Correlation Coefficients (r) and p-values

Variable	**By Variable**	**Spearman Correlation Coefficient (r)**	**Probability value (p-value)**	**Significance**
Factor 4: Migration Planning	Factor 3: Data Governance	0.4554	<0.0001	Highly significant

Factor 4: Migration Planning	Factor 1: Sensitive data management	0.6147	<0.0001	Highly significant
Factor 3: Data Governance	Factor 1: Sensitive data management	0.7350	<0.0001	Highly significant
Factor 2: Sensitive data awareness	Factor 1: Sensitive data management	0.7243	<0.0001	Highly significant
Factor 4: Migration Planning	Factor 2: Sensitive data awareness	0.5861	<0.0001	Highly significant
Factor 3: Data Governance	Factor 2: Sensitive data awareness	0.6641	<0.0001	Highly significant

$H1_1$

The p-value (p<0.0001) is smaller than 0.01 and this indicates a highly significant correlation between sensitive data management (factor 1) and sensitive data awareness (factor 2) at a 99% level of confidence. Since this correlation is positive and very strong (r = 0.7243), it means that should more employees be trained on the handling and protection of sensitive information, then more of them will be aware of how to handle and protect sensitive information; understand information classification; identify sensitive information; and adhere to data security policy in their organisations.

$H2_1$

The p-value (p<0.0001) is smaller than 0.01 and this indicates a highly significant correlation between sensitive data management (factor 1) and data governance (factor 3) at a 99% level of confidence. This correlation is positive and very strong (r = 0.7350). This means that a better organisational security model should lead to improved data access, controls and standards in the organisation. In addition, if the organisational security strategy is well communicated to the employees, it will facilitate better monitoring and controlling of sensitive information and improve employee roles and responsibilities.

$H3_1$

The p-value ($p<0.0001$) is smaller than 0.01 and this indicates a highly significant correlation between sensitive data management (factor 1) and migration planning (factor 4) at a 99% level of confidence. This correlation is positive and strong (r = 0.6147), meaning better monitoring and controlling of sensitive information during migration ought to enhance migration processes such as time management and data quality. Also, if the roles and responsibilities of employees during migration are made very clear, then it will enhance better migration processes.

$H4_1$

The p-value ($p<0.0001$) is smaller than 0.01 and this indicates a highly significant correlation between sensitive data awareness (factor 2) and data governance (factor 3) at a 99% level of confidence. This correlation is positive and strong (r = 0.6641). This means that a better organisational security model will improve the data security policy, employee awareness and handling of sensitive information, as well as the identification of sensitive information by the employees.

$H5_1$

The p-value ($p<0.0001$) is smaller than 0.01, indicating a highly significant correlation between sensitive data awareness (factor 2) and migration planning (factor 4) at a 99% confidence level. This correlation is positive and medium (r = 0.5861). This means that if more employees are aware of, and accountable to sensitive information handling, then this will lead to improved migration processes. Also a better data security policy will enhance migration processes.

$H6_1$

The p-value (p<0.0001) is smaller than 0.01 and this indicates a highly significant correlation between data governance (factor 3) and migration planning (factor 4) at a 99% level of confidence. This correlation is positive and medium (r = 0.4554). This means that the better the organisational security strategy and security model is, the better will be the migration processes.

Sensitive data management is strongly correlated with the rest. Since the four main constructs are highly correlated among themselves, it implies that all the factors must be taken into consideration during migration of sensitive information. None of these factors can be omitted during the sensitive migration of platforms.

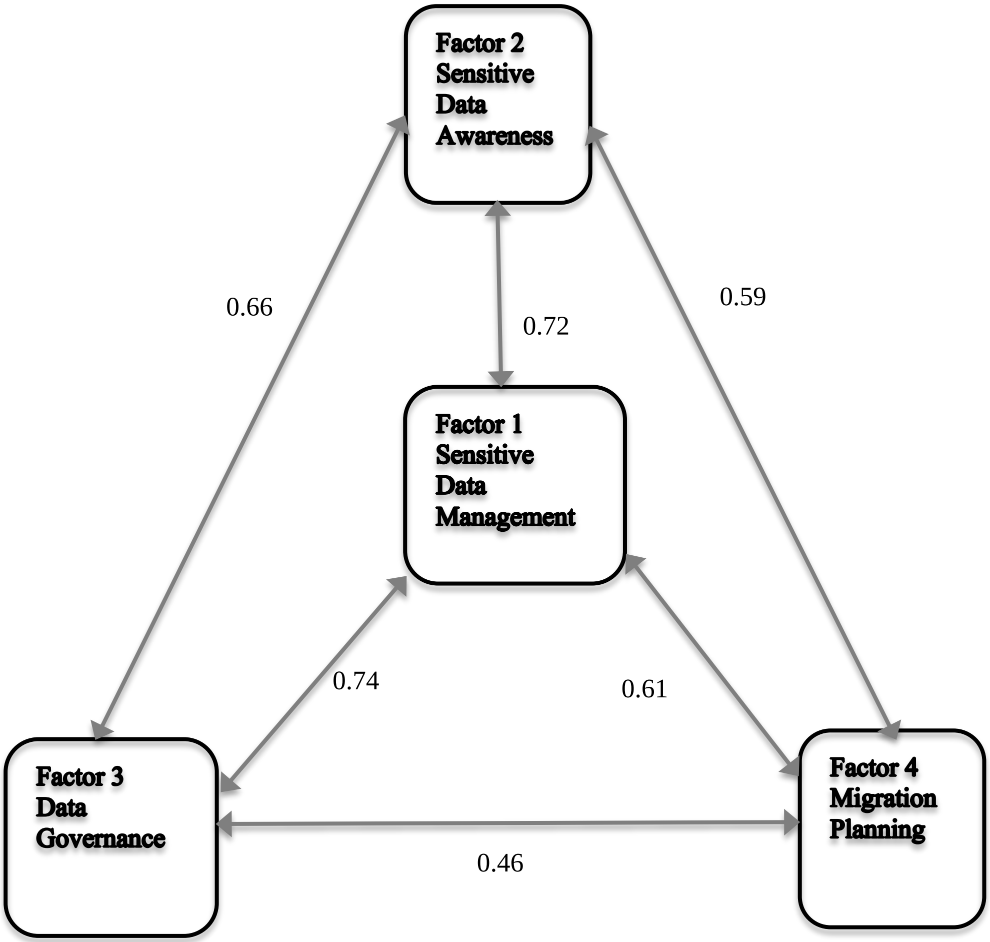


Figure 5-26: The Relationships among the Four Main Constructs with Spearman's Correlation Coefficients (r).

Figure 5-26 is now expanded to show all the variables within each factor and this forms the Preliminary Management Framework resulting from quantitative analysis (Figure 5-27).

5.11 Resulting Framework from the Quantitative Analysis Results

Table 5-22 illustrates the short names assigned to the sub-constructs that are used to expand Figure 5-26 in order to develop the Preliminary Management Framework.

TABLE 5-22
The New Nomenclature of the Sub-Constructs

Sub-Constructs	**Short Name**	**Associated Variables**
B Awareness accountability score or (Employee_Awareness/Information Handling/Accountability)	Awareness accountability score	Awareness/Information Handling/Accountability
B Training handling or (Employee_Course Type/Sensitivity Classification)	Training handling	Course Type/Sensitivity Classification
B Consequences of sensitive data or (Employee_Training/Info Non-Protection Consequences)	Consequences of sensitive data	Training/Info Non-Protection Consequences
C General data policies or (Organisation_Strategy/Culture/Communication/Data)	General data policies	Strategy/Culture/Communication/Data
C Specific sensitive data policy or (Organisation_Data Security Policy/Sensitive Info Identification)	Specific sensitive data policy	Data Security Policy/Sensitive Info Identification
C Access to sensitive data or (Data_Access/Controls/Standards Enforcement)	Access to sensitive data	Access/Controls/Standards Enforcement
D General data issues or (Employee_Roles/Responsibilities)	General data issues	Roles/Responsibilities
D Data security model or (Organisation_Security Models)	Data security model	Security Models

E General control or (Monitor/Control_Tools/Migration Issues/Risk Assessment/Migration Duration/Network Bandwidth)	General control	Tools/Migration Issues/Risk Assessment/Migration Duration/Network Bandwidth
E Migration planning or (Migration Processes_Application Identification/Time Management/Servers De-Staging/Source Data Back Up/Data Quality)	Migration planning	Application Identification/Time Management/Servers De-Staging/Source Data Back Up/Data Quality

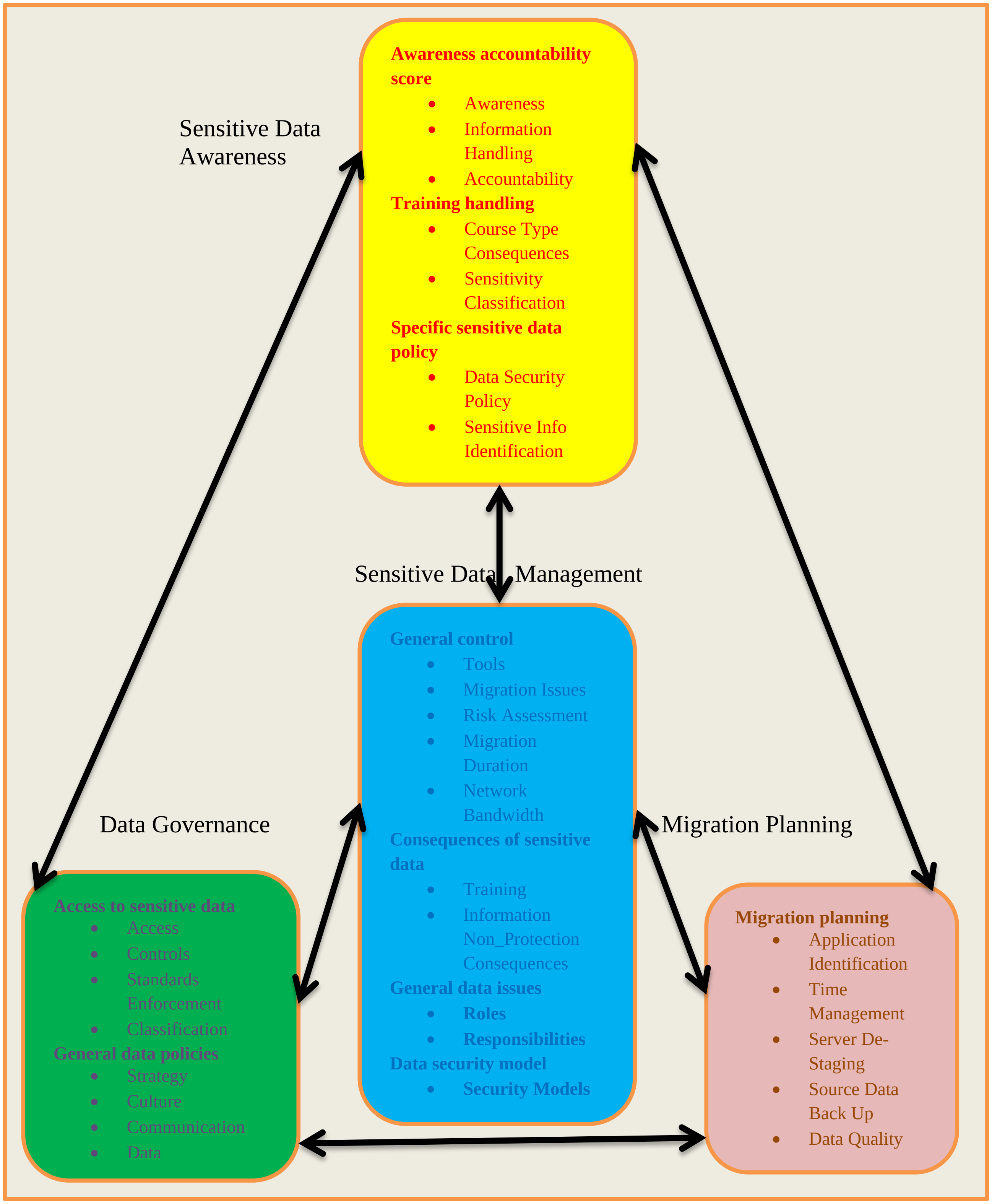


Figure 5-27: The Resulting Management Framework from Quantitative Analysis

Figure 5-27 represents the resulting Preliminary Management Framework from quantitative analysis developed from Figure 5-26. It also includes the new structure obtained after performing the EFA and Spearman's correlations on the sub-constructs on the quantitative data. Since all four main constructs are highly correlated among themselves, it implies that all the factors must be taken into consideration during migration of sensitive information. None of these factors can be left out during the sensitive migration of platforms. The Preliminary Management Framework has inputs from the Rudimentary Management Framework (Figure 2-2) in section 2.14 as well as the Security Model (Figure 3-1) in section 3.5. The double ended arrow indicates that the two factors that are joined together are correlated. More discussion follows in the paragraphs below.

5.11.1 Discussion of the Management Framework from the Quantitative Analysis

There exist very strong correlations among all the four factors namely sensitive data awareness, sensitive data management, data governance and migration planning. Therefore, all the variables indicated under all these factors are very important and must be considered during migration of sensitive information between software platforms. It is recommended that organisations must take cognizance of these variables when performing such software migrations.

The roles and responsibilities of the members of the migration team ought to be clearly defined before the commencement of the project. Organisations ought to provide training and awareness on sensitive information protection and handling. Induction courses ought to cover various aspects of the risks attached to the management of sensitive data. Training ought to spell out the consequences of the misuse of sensitive data and also the risk of not protecting sensitive data.

All employees ought to be educated about the different classification levels, their respective markings and when to apply them. Employees ought to value accountability when they handle sensitive data and handle sensitive information with utmost care as outlined in their data security policy. Employes need to be aware of what sensitive information is and how it should be protected within organisations having a process to identify sensitive information that is worth protecting. Employees working on sensitive data ought to undergo vetting in order to ascertain their confidential sensitivity levels.

Organisations should have a data security policy which lists data security methods and sensitive data management. These policies and procedures should be regularly communicated to and enforced for all staff. There should be a continual update of the data security policy and data integrity should be the hallmark of any organisation.

Organisational strategy ought to include the protection of sensitive information and should be aligned with clear objectives on how sensitive data should be handled. Protecting sensitive information should be part of any organisational corporate culture. Security models should be developed to support organisational strategy and such models should ensure confidentiality, integrity and reliability of data during protection of sensitive information.

The organisational data access by employees should be controlled and monitored and organisational data should be defined through data discovery and classification. Confidentiality, integrity, identifying authorised users and monitoring access should be undertaken by organisations to ensure sensitive data protection. Organisation networks should always be protected at all times. Organisations should provide for continual management of data sensitivity and risk management.

All the data created by users (information creators) should be classified or identified and proactively marked before the data is migrated. Data classification roles and responsibilities (e.g. data creators, data owners, data users, and data auditors) should be clearly defined within the organisation. Enough time should be planned for the data migration process. All the host servers, functions, applications and storage impacted by the data migration should be identified during the data migration. All the data in the servers, memory and buffers should be commited to non-volatile storage before performing migrations. It is important for organisations to know the timing of migration, the migration duration and system down-time (if necesssary).

The source data should be backed up prior to data migrations to the destination. The issues of data corruption, missed data or data loss should be considered during migration. Technical controls should be in place to ensure effective sensitive data protection during migrations. The necessary monitoring systems and risk assessment systems should be in place. The network bandwidth capacity utilisation needs to be measured before migration as well as when it will be available to ensure smooth migration. Verification or comparing migrated data versus source data should be performed, and, if problems persist, a data quality process should be performed.

The above discussion addresses aspects of Figure 5-27. Organisations must consider all the aspects indicated in all the four factors when performing migration of software platforms such as training and awareness on sensitive information, data security awareness, sensitivity clasification, inclusion of the protection of sensitive information in organisational strategy, controlling of organisational data access, protection of organisational networks, backing up of source data prior to data migration, ensuring technical controls, implementing network and risk assessment systems and sensitive information identification.

5.12 Validity and Reliability in Quantitative Research

Validation approaches in quantitative research include: internal validity, external validity and construct validity. Validity assessments in quantitative research are usually executed by statistical analysis (e.g. content, construct and predictive validity) or relate to questionnaire design (e.g. face validity and bias) (Symonds & Gorard 2009).

There is a high degree of acceptance on how quantitative analysis should be assessed, however, this consensus is lacking for qualitative analysis (Bamberger 2007). He further suggests seven categories for mixed methods validation: (a) confirmability/objectivity; (b) dependability/reliability; (c) credibility/internal validity; (d) statistical conclusion validity; (e) construct validity; (f) transferability/external validity; and (g) utilisation.

Construct validity is defined as the experimental demonstration that a test is measuring the construct it claims to be measuring and it should be determined by a buildup of evidence, e.g. correlation coefficients, factor analysis, ANOVA (Brown 2000). Construct validity refers to the complex question of whether test score analyses are consistent with a theoretical and observational terms (Cronbach & Meehl 1955). A researcher will be able to prove a test's construct validity with more evidence (McLeod 2013).

Construct validity is established by the exploratory factor analysis and the Spearman's correlations of the constructs (sections 5.3; 5.7 and 5.8.2). The reliability is established by the Cronbach's alpha values of the constructs (sections 5.4 and 5.8.1). Internal validity is established by the triangulation of results from the interviews.

Internal validity checks if the effects observed in a study are due to the alteration of the independent variable and not on some other factor. External validity is the degree to which the results of a study can be generalised to other settings (ecological validity), other people (population validity) and over time (historical validity) (McLeod 2013).

5.13 Conclusion

This chapter highlights the various statistical techniques applied to perform the quantitative data analysis. The data distribution, the Exploratory Factor Analysis (EFA) and the reliability analysis (using the Cronbach's Alpha value) are utilised to ascertain the reliability of newly formed constructs.

The EFA is used to regroup the sub-constructs in order to have regroupings that are valid. After the regroupings of the sub-constructs were done, the reliability of the newly grouped main constructs was undertaken using reliability analysis (also known as Item Analysis) to calculate their Cronbach Alpha coefficients. The four main constructs were subsequently compared to each other to determine how well they were correlated. It was found that the constructs were well correlated to each other with the significance ranging from medium to strong. The Preliminary Management Framework from Quantitative Analysis was developed on the strength of the various statistical analyses performed.

The Preliminary Management Framework is validated using a qualitative analysis in the next chapter.

Chapter 6

Final Management Framework: Qualitative Data Analysis

6.1 Introduction

The previous chapter focused on developing the Preliminary Management Framework using quantitative data analysis. In this chapter, the Final Management Framework is developed by using qualitative data analysis. The resulting Final Management Framework is validated and the outcome of the validation is presented.

Semi-structured interviews were held with ten participants selected from the seven South African government organisations mentioned in section 1.1.1. The same set of questions were asked of all the participants on different occasions and the set of questions is shown in Appendix C. The researcher took a realistic and flexible approach (i.e. pragmatic stance) by incorporating the content of the resulting Preliminary Management Framework from quantitative analysis, Figure 5-27 (section 5.11) in the questions put to the interviewees. The questions were designed to enhance and validate the Preliminary Management Framework from quantitative analysis. This is in accordance with the visual model of mixed methods design: sequential explanatory design procedures developed by Ivankova *et al.* (2006) as described in Figure 4-1.

The NVIVO Version 10 software is used to analyse the qualitative data resulting in the resulting management framework from Qualitative Analysis. The software is used to develop categories from the qualitative data which then form the themes that are used to develop the framework. The qualitative analysis is used to validate the result sets of the quantitative analysis in a mixed methods research setting. The resulting Preliminary Management Framework from the quantitative analysis is validated using these qualitative statistical tools in order to obtain the Final Management Framework on information sensitivity.

6.2 Validity and Reliability in Qualitative Research

Qualitative research practitioners have adopted criteria such as validity of their methods that are taken from the positivistic paradigm in order to reduce bias and subjectivity (Pozzebon 2003). The validity of a study is the degree of the trustworthiness of its results and the outcomes are not biased by the reseacher's own perceptions (Runeson & Host 2009). There are four types of validity (Runeson & Host 2009):

- Construct validity, which looks into what extent the operational actions are being studied, that is, what the researcher has in mind and what is investigated according to the research questions.
- Internal validity, which is important when causal relations are being examined. A researcher in investigating whether one factor affects another one might lead to the possibility that the investigated factor might be affected by a third factor. If the researcher is not aware of the existence of the third factor and/or knowing the extent of how the third factor affects the investigated factor, then, there is a threat to the internal validity.
- External validity, which relates to knowing the extent to which to generalise the research findings and also to know the limits to which the results will be of importance to other people outside the investigation, e.g. defining a theory. The researcher tries to examine the degree of the extent the findings are of importance for other cases.
- Reliability, which applies to the extent that data and analyses are reliant on the specific researchers; the same result must be obtained if another

> researcher conducts the same study at a later stage. It is concerned with the degree that the data and the analysis are reliant on the specific researchers.

Qualitative reliability is best performed by a basic redundancy test and reliability is satisfied if a group of interviewees as a whole give the same answers to the questions posed to them (Trotter II 2012). Gulati and Taneja (2013) indicate the importance of using the right kind of sample in order to arrive at valid and meaningful outcomes. All ten qualitative research respondents (respondents A to J) are well experienced in the practice of information security and information sensitivity. These respondents are identified using purposive sampling. Interviewees in qualitative research studies should be chosen purposefully, have a clear reasoning and satisfy a specific purpose to the research question (Cleary *et al.* 2014; Collingridge & Gantt 2008; Patton 1990). Multiple researchers have validated that their qualitative sample should not be too large in order to avoid bias (Gulati & Taneja 2013). Sample size in qualitative studies should be justified on the data quality and small sample size is justified since they can be studied intensively (Cleary *et al.* 2014). Therefore using ten interviewees in this research is justifiable because they were chosen purposefully with clear reasoning and also to avoid bias.

Johansson (2003) has pointed out that triangulation can be used to ensure the validity of research and this means having different data sources, or many data collection methods. Data in this study were collected from many different sources (data triangulation) and this contributed to the validity of the research.

Other common ways employed in qualitative validation includes credibility, confirmability, dependability and transferability. Credibility allows others to distinguish the experiences contained within the research through the interpretation of the participants' experiences. Credibility is equivalent to internal validity in quantitative research. The representativeness of the data as a whole must be checked in order to achieve credibility and this involves the researcher reviewing the individual transcripts, looking for similarities across and within the interviewees. Reflexivity, member checking, and peer debriefing or peer examination are used to establish credibility (Thomas & Magilvy 2011). Credibility of the study is confirmed in section 4.6 by the researcher's reflexivity due to the movements between the quantitative and qualitative analyses.

Confirmability refers to the degree in which the conclusions drawn are from the available evidence and if the research is relatively free from bias (Guba & Lincoln 1989). It also measures the extent that the research's methods and procedures are adequately described and if the methods used are enough to control bias. Confirmability is confirmed in sections 4.3; 4.4; 4.5 and 4.6 which adequately describe the research's methods and procedures.

Transferability is the ability to transfer research findings or methods from one group to another, or the determination of the extent to which the research findings of a study are applicable in other contexts or with other subjects/participants. It is equivalent to external validity in qualitative research (Lincoln & Guba 1985). Transferability could not be confirmed in this study due to the nature of the research setting.

Dependability refers to a situation when another researcher can follow the decision trail used by the researcher, and it is related to reliability in quantitative research. It involves the description of the study purpose; discussion around the reasons behind the selection of the participants for the study; the place and time lasted for data collection; discussion around data analysis; an explanation on how the data were reduced or transformed for analysis and explaining the specific techniques used to conclude the credibility of the data (Thomas & Magilvy 2011). Sections 4.3.9; 4.5

and 4.6 confirm the dependability of this study since the study purpose, discussion around the reasons behind the selection of the participants for the study, the place and time lasted for data collection, discussion around data analysis how the data were transformed for analysis are all explained in these sections of the thesis. The external validity is established by the validation of the management framework (section 6.6.2).

6.3 Qualitative Data Analysis Steps

Figure 6-1 indicates the data analysis steps for the qualitative case study that the researcher obtained from Stockdale and Standing (2002). The data analysis of the qualitative data is performed by using the NVIVO Version 10 Software. This is to simplify the coding process and to generate computerised reports.

Coding refers to the process of assigning categories, concepts, or ‘codes’ to segments of information that are of interest to the research objectives. Coded data implies that textual parts are allocated a code that represents for instance a certain theme, construct or area (Runeson & Host 2009).

A code can be allocated to many text pieces while one text piece can be allocated to more than one code. The resulting codes can create a hierarchy of codes (or categories), sub-codes and sub sub-codes. The resulting categories are then grouped together to form themes which, in this thesis result in the framework on information sensitivity during migration of platforms.

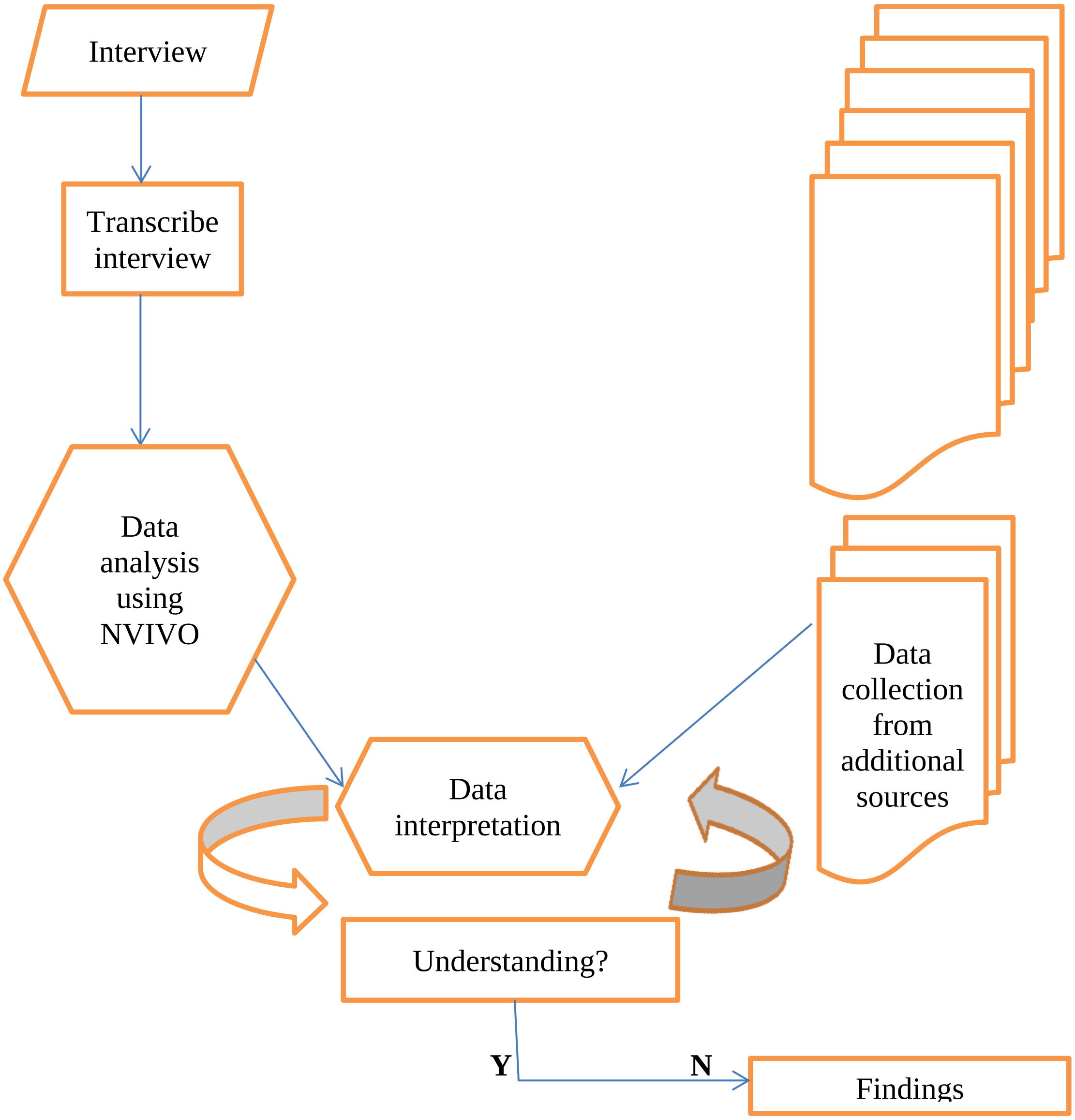


Figure 6-1: Qualitative Data Analysis Steps (from Stockdale & Standing 2002)

The semi-structured interviews were conducted with ten participants drawn from the seven South African government organisations (mentioned in section 1.1.1) using the same set of questions as shown in Appendix C. Each interview lasted between thirty and sixty minutes. All the interviews were transcribed by the researcher and the interview transcripts were uploaded as Microsoft Word documents in the NVIVO software for data analysis. The NVIVO software is used to perform the qualitative data analysis and the results are presented in the sections that follow.

6.4 Interview Narratives

The narratives of the interviews from the ten interviewees are presented below:

1. All ten of the interviewees said that they understand the difference between sensitive information and non-sensitive information and they all explained the difference between the two types of information. Most of them described sensitive information as *the information that is classified as information that should not be accessible or accessed by any other person except the one that it is intended for while non-sensitive information is that information that can be accessed by anyone without any repercussions*. One of them said '…sensitive information is the information that is restricted in terms of who can access it and it is also to some extent information that if accessed can compromise the security policies of that organisation'. Sensitive information identification is part of the preliminary management framework.
2. On the protection of sensitive information during software migration, most of the interviewees mentioned that *encryption techniques should be used as well as techniques such as hashing should be used*. Some also mentioned that employees handling sensitive information need to be vetted and obtain security clearance to know the type of information they can handle. Others also said that data must be classified first before migration so that they can know the kind of protection measures applicable to the various data sensitivity levels. Encryption is a new concept and it is added to the preliminary management framework to form the final management framework.
3. All ten interviewees agreed that *organisations must have security models to support their organisational strategy*. One of them said that '…a security

model can definitely help to guide what is included in any organisational strategy…by including security from the beginning'. The view of most of them is that security models will provide a common platform, common goal, a sense of responsibility, and understanding of the sensitivity of information. Security models are part of the preliminary management framework.

4. All the interviewees agreed that *the organisational strategy should incorporate how the organisational data will be protected and handled*. The reasons they gave include not allowing sensitive information to fall into the hands of malicious people with harmful intentions, and also allowing for data classification. Organisational strategy is part of the preliminary management framework.
5. There was a common agreement among all the interviewees on the concept that *employees handling organisational data should be trained on how to handle sensitive information in order to avoid data corruption, and that information is an asset that should be well taken care of*. One of them said that '…the weakest link in security is mostly the people, so if people are not aware of the risks or the vulnerabilities, cases of mishandling of sensitive data will rise'. Training is part of the preliminary management framework.
6. All the interviewees agreed that *employees should perform sensitivity assessment as part of the organisational strategy in protecting their organisational data depending on their environments*. One of the reasons mentioned is that it would stimulate their conscience to regard information as a valuable asset to their organisation. One of them said that '…Depending on the environment, you get environments that really don't have varying levels of sensitivity but in an environment like in government, the sensitivity can definitely vary and therefore the protection mechanisms will also vary based on the sensitivity. If you have more than one level of sensitivity, that assessment was assessed in segregating the data or at least labelling the data depending on how to segregate them to prescribe which

ones should have the strongest protection mechanisms applied. This is so often referred to as the top level security. If you have more sensitive data than the other same environment you will have multiple levels of security mechanisms that are applied based on the sensitivity. This is in essence crucial, since you don't know what data that is sensitive, and therefore, you cannot protect it accordingly to the prescripts. Therefore, you should definitely do assessment to ensure proper security in order to apply the appropriate mechanism of sensitivity'. This implies that government employees are subjected to security clearance based on their exposure to levels of information sensitivity. Therefore, government employees should only be allowed to work with information in line with the level of their security clearance. Also the protection mechanism should be based on the level of security clearance of government employees. Sensitive assessment is a new concept and is included into the final management framework.

7. All the interviewees agreed that *organisations should have policies and procedures on handling sensitive information and that it is a matter of compliance with governance*. Policies should inform employees on how data is used or handled and should be enforced so that people abide by the standards and procedures. One of them said that '…reason for this is to bring everybody on board so that they can understand what is regarded as sensitive and what is non-sensitive'. Policies can be enforced by providing awareness to the employees in the form of training. Some respondents mentioned that policies should be enforced by linking it to the performance review of the employees. Policies and procedures are part of the preliminary management framework.
8. There was a general consensus among all the interviewees that *it is important for organisations to control and monitor their data access by their employees to avoid data corruption by the employees*. Other reasons mentioned are to ensure accountability of data and also to limit data access. One of them said '…the access to data exposes your data firstly to leakage or to modification or whatever the intent anybody may have. Firstly by

limiting access to data implies that someone does not know what exists and this will not bother them. Therefore, we limit the access to allow them to use what they need to do their job and that will assist avoiding data corruption. If they know there are secret data somewhere in another database and they cannot have access to it then the possibility of their exploiting that access or using the access is just so much better. Having access to what they need to do their job definitely protects the inner security level that you can enforce on data.' Data access controls are part of the preliminary management framework.

9. All the interviewees agreed that *technical controls should be put in place by organisations during data migrations*. Some of the reasons given include: to ensure that there are no unwanted devices accessing the network during migration, to provide an audit trail or history of what has happened so that if something went wrong it is easier to trace what went wrong, and who was there and who violated it. Also it can be used to enforce the policies. Technical controls are part of the preliminary management framework.
10. All the interviewees agreed that *the organisational source data be backed up prior to migration* because if anything goes wrong during the migration process, it would be difficult to roll back to the previous state and afterwards the migration can proceed again. Some suggested keeping the backup copy of sensitive data off-site as a precautionary measure to protect sensitive information. One of the interviewees said that '…this is important so that if something goes wrong then you can fall back in terms of your operations and your business continuity'. Backing up of source data is part of the preliminary management framework.
11. All of the interviewees agreed that *it is important to determine the duration of the migration process before the migration of data* because it is part of project planning in order to ensure a successful migration project and it will assist in the business continuity planning. One of them said '…it is very much important…you can also factor in the cost and it assists in managing

the schedules in terms of the data migrations'. Migration duration is part of the preliminary management framework.

12. They all agreed that *organisations should ensure that the necessary monitoring and risk assessment systems are in place prior to migration of data* because they will assist in the determination of risks, guide against data loss, and keeping track of events for accountability. Some said that this should be in the organisational strategy and it would facilitate risk mitigation. Monitoring and risk assessment systems are part of the preliminary management framework.
13. All the interviewees agreed that *it is important for organisations to ensure the availability of adequate network bandwidth before commencing on a migration process*. This is to avoid slow response rate, network going down and improve network productivity. One of them said '…Yes, very important…we have a migration that after it was done, the system becomes slow because proper network bandwidth assessment was not done….so it is important because it would affect the usability and end throughput in the network. Availability of adequate network bandwidth is part of the preliminary management framework.
14. All of them agreed that *proper migration tools and strategies be provided prior to migration of data so that the planning and the execution process proceeds in a coherent manner*. It should have been spelt out in the user specifications requirements at the beginning of the migration project. A strategy is a roadmap and it included project monitoring tools in order to ensure a successful migration project. Migration tools are part of the preliminary management framework.
15. All the interviewees agreed that *database activities should always be monitored* since the database is the life of the organisation and therefore it must be secured. Examples of monitoring activities include: Who accesses the database? Are the database requests normal? What has happened in the database? What did they do with the data? Database activities of users like modifications, deletions and alterations can be selectively monitored.

Database activities monitoring is a new concept and has been incorporated into the final management framework.

16. There was a common agreement on the issue that *organisations should identify the functions, applications, hosts, host servers and storage*. This is to enhance the effort, and the requirements in these applications and also to know the interfaces between the applications, as well as the application formats and the data storage requirements prior to migration. Identification of applications is a new concept and has been incorporated in the final management framework.
17. All of the interviewees agreed that *organisational data should be classified prior to migration as part of the security strategy*. This will show who should access the data based on the data classifications and their security level clearance. It will also aid in the protection of sensitive information since only employees that have security clearance to handle such information will be allowed to do so. One of the interviewees said that '…Yes,…it is an indication of how that data should be handled. Now if it is classified, then it would be handled according to its classification.' This can also aid in knowing which data is more important than the other and can be used to prioritise the migration process. Data classification is part of the preliminary management framework.
18. All of them agreed that *it is important for organisations to clearly define the data classification roles and responsibilities of employees involved in the data migration (e.g. data creators, data owners, data users and data auditors)*. This will give a separation of duties for auditing purposes, so that the one who creates the data is not the one who audits it. One of them said that '…the data owner normally classifies the data and determines the level of the data sensitivity…So as to enforce the roles of the different role players'. Responsibilities of employees during migration are part of the preliminary management framework.
19. They all agreed that *the Extract, Transfer and Load (ETL) scripts used to perform the migration be reviewed for reliability and accuracy*, in order to

check if they are working well or not. Additionally they will ensure the correctness of the data that is transferred so that the original source data is the same as the target data. Review of ETL scripts is a new concept and has been included into the final management framework.

20. They all agreed that *the flow of sensitive data should be monitored during the migration process* so that sensitive data arrives at the right destination at the same level of quality. This will avoid sensitive information leakage and is one of the protection mechanisms of sensitive information migration. Monitoring of the flow of sensitive data is a new concept and has been incorporated into the final management framework.
21. All the interviewees concurred that *servers, memory and buffers should be commited to non-volatile storage before migration* so that the data is not corrupted by the uncleanliness of these devices. Also it will ensure the optimal performance of these devices. Commitment of memory and buffers to non-volatile storage is part of the preliminary management framework.
22. All agreed that *business rules should be examined in order to provide a basis for information categorisation with respect to sensitivity*. This will enable organisations to stay current with the standard of information since they guide the categorisation and classification of information. The business rules should prescribe the data modifications that have influence on the security of data. It should be determined at the beginning of the migration project and it is used to determine the sensitivity that the business needs the data for, e.g. top secret level for information like trade secrets. Examination of business rules is a new concept and has been included into the final management framework.
23. All the interviewees agreed that *security tools such as Continuous Data Protection (CDP) and Data Loss Prevention (DLP) and Cipher Text Encryption Tools be used to protect sensitive information during migration of data*. This is because it can prevent data loss and avoid data corruption during network transmission. Sensitive data must always be encrypted and should only be decrypted when it is accessed by the authorised users for

readability purposes. This will also ensure Confidentiality, Integrity and Availability (CIA) of data. One of them said that '…certain data stay in encrypted form wherever they are stored. Other data is encrypted only while they are transmitted.' Encryption is a new concept and has been incorporated into the final management framework.

24. All the interviewees agreed that *the migrated data should be tested and validated in order to ensure data accuracy and integrity and also be subjected to a data quality process after the migration process*. It is a requirement by the Auditor General that all data be backed up and tested after migration. It ought to enable successful transfer and avoid loss of data or data corruption and also ensure data integrity. One of the interviewees said that '…Yes, just to make sure that you've got everything and what you have is correct…a lot of things can go wrong during migration…it can be human errors'. The data quality process will ensure the correctness of the data and there is an auditable process that can assist if things go wrong during the migration process. Testing and validating of the migrated data are new concepts and have been added into the final management framework.
25. All the interviewees agreed that *the Total Cost of Ownership (TCO) of the new software should be taken into consideration during migration*, since the return on investments must be looked into as part of the budgetary process. One of the interviewees said that '…l remember a company that made a mistake, they migrated from their Point Of Sale (POS) to SAP, only to realise that they incurred a lot of cost, which means that they didn't plan and calculate the Total Cost Of Ownership (TCO). Cost is a new concept and has been incorporated into the final management framework.
26. All the interviewees agreed that *organisational structure and culture should be taken into consideration when planning migrations*. The organisational structure will come up with the resource capacity within the organisation to understand if the organisation has got the right resources to undertake the migration project. The culture of the organisation needs to

understand the process of the migration so that the migration project is not jeopardised since migrating data is a process that involves people and the way they do things. There is a change in the way the organisation does business which can cause some people not supporting the new system. Organisational structure and culture are part of the preliminary management framework.

27. There was a general consensus among the interviewees that *IT standards such as ISO/IEC 17799 should be adhered to during software migrations* because standards give the best practice baseline for IT governance since they are the basis of the foundations of information security. Organisation data security policies should be based on such standards to ensure protection of their data during migrations and ensure interoperability of information across organisations. The above consensus is significant for the frameworks developed in this thesis, since the treatment and classifications of sensitive information are based on the ISO/IEC 17799 standard. Enforcement of IT standards is a new concept and has been included into the final management framework.
28. All the interviewees agreed that *the employee's attitude and behaviour are some of the important human elements that should be considered during software migrations*. The migration team should be composed of dedicated and enthusiastic people that are committed to the success of the project. Members of the migration team must be certified at least up to a secret classification level. It is very important that members of the migration team have the right attitude and behaviour and that they also adhere to the organisational security policies and procedures. One of them said that '…people who should do information security duties should be people of higher integrity…employees' negative attitude can be changed by training them and awareness …and re-orienting them'. Employees's attitude and behaviour are new concepts and have been incorporated into the final management framework.

6.5 Categorisation of Information (Nodes Identification)

The ten transcribed qualitative interviews transcripts were imported into the NVIVO Version 10 Software to perform the coding process on the data. The NVIVO software refers to categories as nodes and the researcher was able to code the data using the categories identified having determined the words used most often. Categories, sub-categories and sub sub-categories were used to code the data.

Table 6-1 illustrates the word table of the most frequently used 100 words with a minimum word length of four. It shows the length and the count of the words as well as their weighted percentage. Possible variables of the management framework are shown in bold print in the table.

TABLE 6-1
Word Table of the Qualitative Data (most frequently used first 100 words)

Word	Length	Count	Weighted Percentage (%)
data	4	671	4.23
migration	9	431	2.71
information	11	318	2.00
important	9	221	1.39
think	5	196	1.23
sensitive	9	191	1.20
need	4	156	0.98
people	6	148	0.93
process	7	139	0.88
organisations	13	128	0.81
organisational	14	127	0.80
ensure	6	125	0.79
must	4	124	0.78
know	4	122	0.77
security	8	120	0.76
migrations	10	107	0.67
also	4	100	0.63

terms	5	85	0.54
access	6	84	0.53
policies	8	79	0.50
make	4	74	0.47
software	8	73	0.46
sure	4	73	0.46
going	5	72	0.45
migrated	8	69	0.43
strategy	8	69	0.43
actually	8	68	0.43
employees	9	64	0.40
organisation	12	62	0.39
just	4	61	0.38
like	4	61	0.38
place	5	61	0.38
sensitivity	11	60	0.38
platform	8	59	0.37
management	10	57	0.36
business	8	53	0.33
done	4	51	0.32
things	6	51	0.32
network	7	50	0.31
systems	7	50	0.31
prior	5	49	0.31
back	4	47	0.30
cost	4	47	0.30
servers	7	47	0.30
believe	7	46	0.29
framework	9	46	0.29
source	6	46	0.29
system	6	46	0.29
project	7	45	0.28
time	4	45	0.28
even	4	44	0.28
issues	6	44	0.28
might	5	43	0.27
tools	5	43	0.27
environment	11	42	0.26
want	4	42	0.26
bandwidth	9	41	0.26
change	6	41	0.26
integrity	9	41	0.26
protected	9	41	0.26
standards	9	41	0.26

assessment	10	40	0.25
monitored	9	40	0.25
able	4	39	0.25
example	7	39	0.25
understand	10	39	0.25
whatever	8	39	0.25
risk	4	38	0.24
take	4	36	0.23
used	4	36	0.23
well	4	36	0.23
classification	14	35	0.22
encrypted	9	35	0.22
part	4	35	0.22
planning	8	35	0.22
culture	7	34	0.21
encryption	10	34	0.21
open	4	34	0.21
protection	10	34	0.21
controls	8	32	0.20
handling	8	32	0.20
protect	7	32	0.20
plan	4	31	0.20
rules	5	31	0.20
something	9	31	0.20
transfer	8	31	0.20
whole	5	31	0.20
classified	10	30	0.19
critical	8	30	0.19
first	5	30	0.19
migrate	7	30	0.19
procedures	10	30	0.19
accuracy	8	29	0.18
order	5	29	0.18
quality	7	29	0.18
structure	9	29	0.18
work	4	29	0.18
working	7	29	0.18
difference	10	28	0.18
mechanisms	10	28	0.18

Figure 6-2 illustrates the word cloud of the most frequently used 100 words with a minimum word length of four.

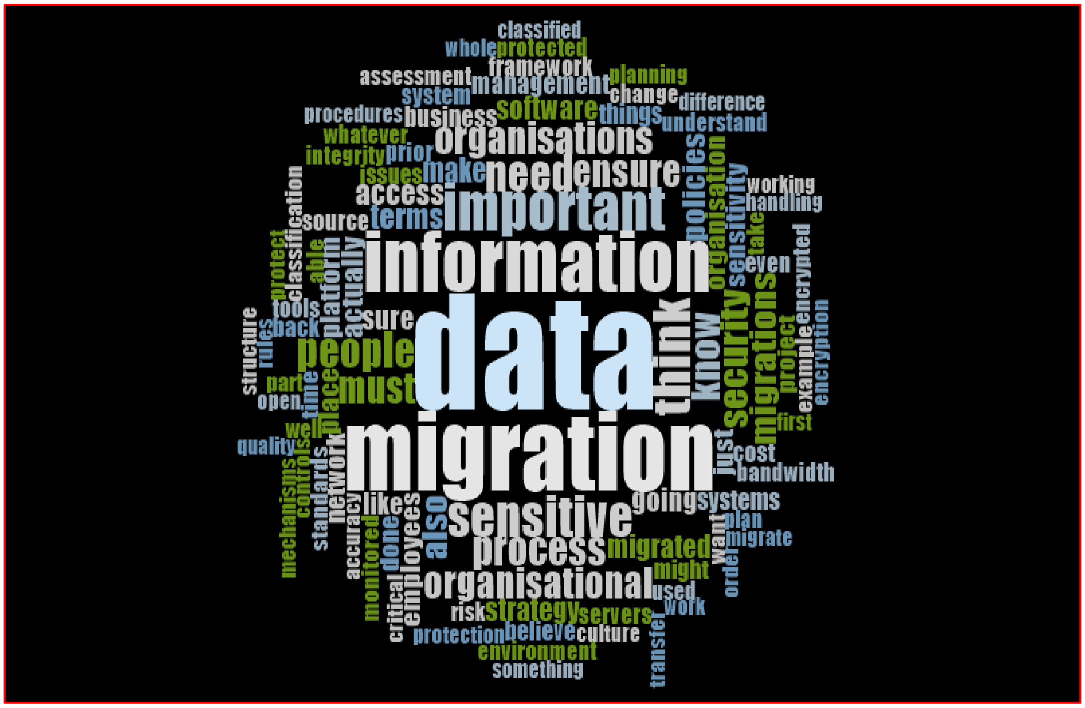


Figure 6-2: Word Cloud Diagram for the Qualitative Data Analysis

The larger the size of the word, the more frequently it is used in the data e.g. data is the largest word followed by migration, information and sensitive, therefore these are the most frequently used words in the data in Figure 6-2. The tree map of the word frequency query is shown in Figure 6-3. The larger the box in the tree map the more frequently is the word inside the box is used in the data. That implies that data, migration and information are the most frequently used words in the data.

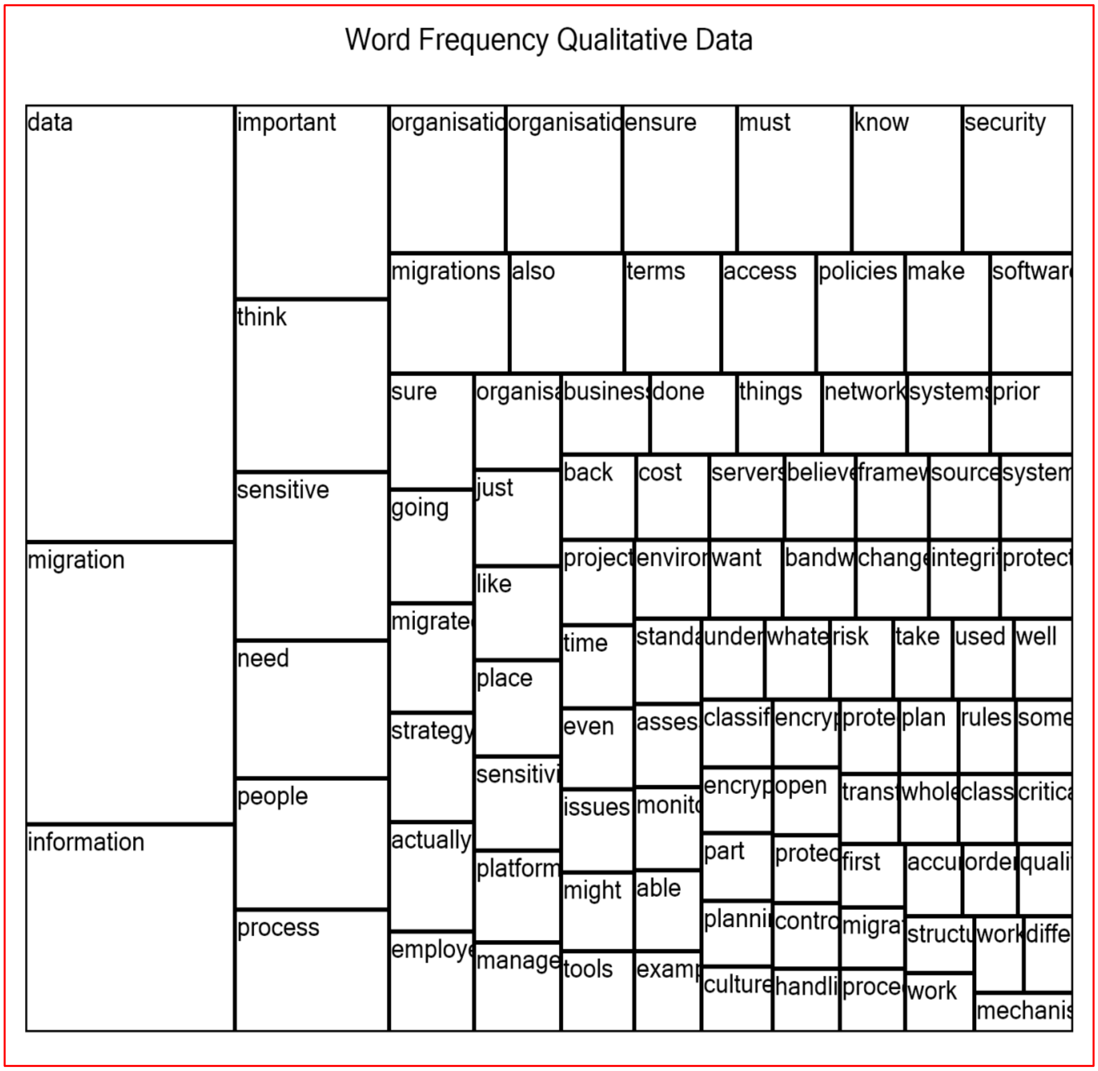


Figure 6-3: Tree Map for the Qualitative Data Analysis

Figure 6-4 shows the cluster analysis for the query – the most frequently occurring 100 words with a minimum word length of four.

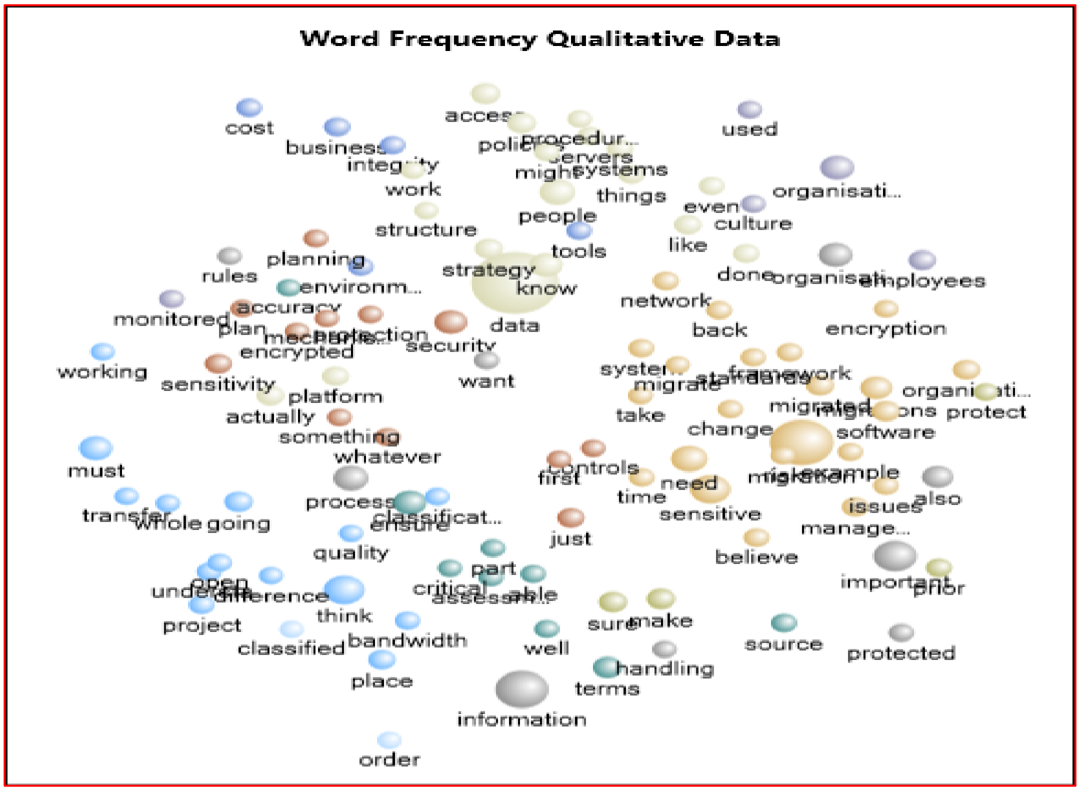


Figure 6-4: Cluster Analysis for the Qualitative Data Analysis

Figure 6-4 highlights how the first 100 words in the data are clustered together. The Figure 6-4 allows the viewing of closer words.

Table 6-2 elucidates the distribution of the nodes and the references of the nodes in the interview transcripts across the ten interviews.

TABLE 6-2
The Distribution of the Nodes and References in the Qualitative Data

Name	Nodes	References
Interview with Person F	33	65
Interview with Person A	38	122
Interview with Person B	37	79
Interview with Person C	38	85
Interview with Person D	37	71

Interview with Person E	35	74
Interview with Person G	37	74
Interview with Person J	22	34
Interview with Person H	30	45
Interview with Person I	27	39

In Table 6-2, 33 nodes were coded with 65 references in the interview transcript of Person F while for the interview with Person A, 38 nodes were coded with 122 references. The categories, sub-categories and sub sub-categories as well as the data narratives in section 6.4 were all used to develop the final framework explained in section 6.6.

6.6 Management Framework on Information Sensitivity during Software Migrations

Sections 6.4 and 6.5 confirm most of the variables in the Preliminary Management Framework (Figure 5-27). Moreover, new variables were obtained from these two sections and they were added to the Preliminary Management Framework (Figure 5-27) to obtain the Final Management Framework (Figure 6-7). These new variables are: sensitivity assessment; attitude and behaviour; test/validate; cost; IT standards enforcement; database activities; sensitive data; business rules; encryption; applications; servers and ETL scripts. Figure 6-5 illustrates the resulting final management framework on information sensitivity during migration of software platforms.

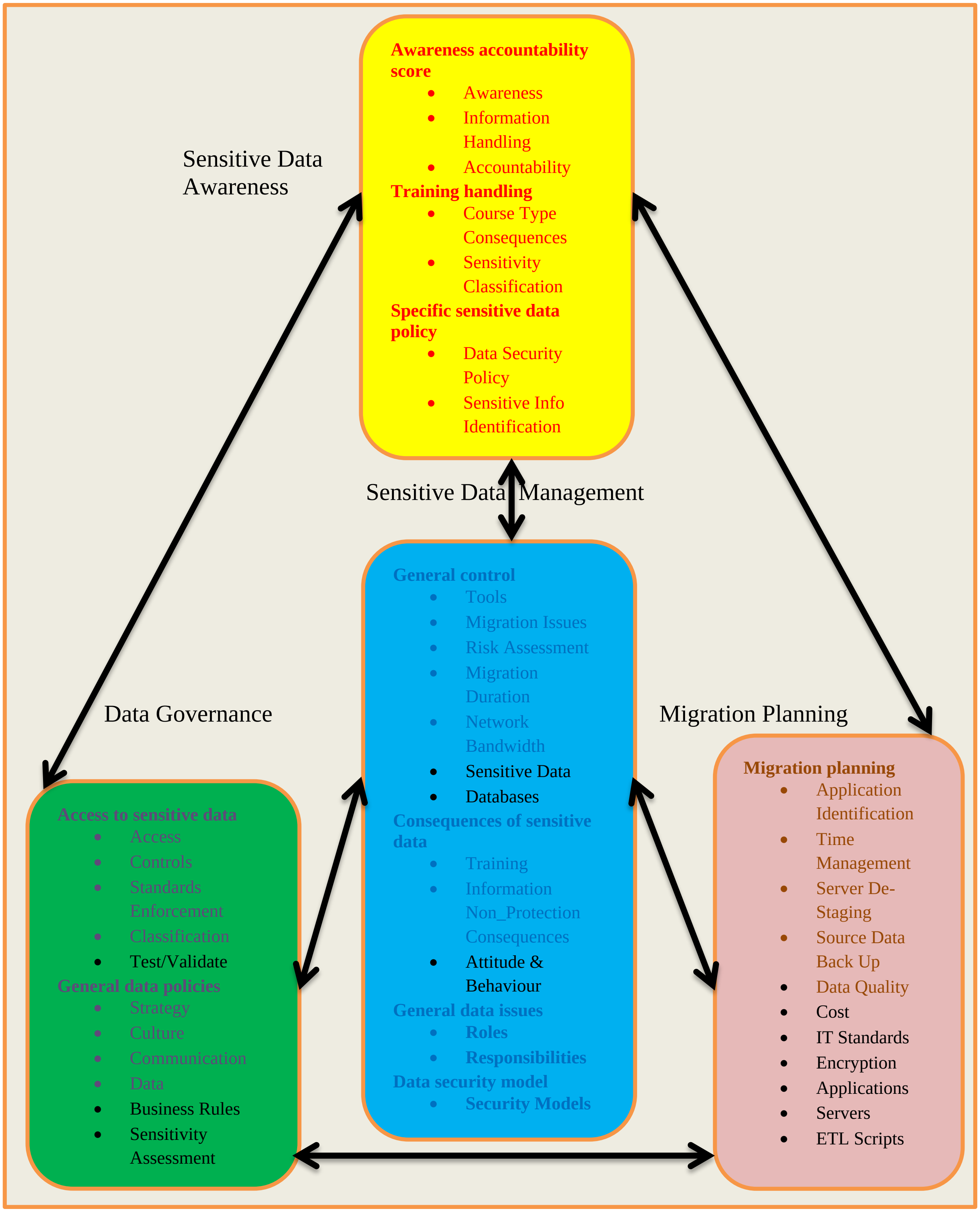


Figure 6-5: Management Framework on Information Sensitivity during Software Migrations

Figure 6-5 is an enhancement of Figure 5-27 after the qualitative analysis have been performed. The new variables added to Figure 5-27 are shown in bold/black letters in Figure 6-5. Sensitive data management is the core of the whole migration process and it is linked to all the variables. All the variables are linked together and no one item can be left out, all the items must be considered and this also comes out (emerges) from the statistical correlations. The discussion on the Final Management Framework on information sensitivity is presented in section 6.6.1 and the validation of the Final Management Framework is presented in section 6.6.2.

The Final Management Framework is also an expansion of both the Rudimentary Management Framework (Figure 2-2) and the security challenges during OSS migrations model (Figure 3-1). More variables that are not previously indicated in Figure 2-1 and Figure 3-1 are now included in the Final Management Framework.

6.6.1 Discussion of the Management Framework on Information Sensitivity

The Final Management Framework indicates that the roles and responsibilities of the people (migration team members) are part of the management framework and should be clearly defined before the commencement of the migration project. Training and awareness need to be provided to all employees handling sensitive information as a protection measure to safeguard the organizational sensitive information. Employees should be accountable for the sensitive information that they handle within their organisations. Employees should be made aware of the consequences of sensitive information non-protection. Employees should be educated on the different classification of information within their organisation in order to serve as a protection mechanism for the organizational sensitive data. All employees should be educated in the different classification levels, their respective markings, and when to apply them. Employees should value accountability when they handle sensitive data, and handle sensitive information with care – as outlined in their data security policy. Employees need to be aware of what is sensitive information and how it should be protected, with organisations having a process to identify sensitive information that is worth protecting.

The Final Management Framework also points out that organisations should have a security strategy in place which should incorporate the culture of the organisation with respect to information security. Organisations should have data security policy which should be regularly communicated to and enforced among all the employees. Organisations should communicate all the various information security guidelines to their staff to ensure that employees adhere to these security guidelines. Security models should be developed to support organisational strategy, and such models should ensure confidentiality, integrity and reliability of data, in order to protect sensitive information. Organisations should access the cost of their migration projects before embarking on such projects. The total cost of ownership of the migration projects should be computed during the migration planning stage, to facilitate the completion of the migration project within its initial budget allocation. The benefits, value, and return on investment must be explored before embarking on the migration project, in order to ensure that the migration project is beneficial to the organisation

The Final Management Framework indicates the use of tools such as encryption technology (Continuous Data Protection and Data Loss Prevention) during the migration of software platforms. Organisations should have the required tools, applications, databases, servers and data migration strategies in place, in order for them to have a successful migration. Organisational networks should be protected at all times. The organisational data access by employees should be controlled and monitored, and organisational data should be defined through data discovery and classification. Confidentiality, integrity, identifying authorised users, and monitoring access, should be undertaken by organisations, to ensure sensitive data protection. Organisations should enforce hardware and software standards in order to eliminate unknown factors that might access their sensitive information. The attitude and behaviour of the migration team members should be taken into consideration before the composition of the team. The migration team should be composed of dedicated and enthusiastic people who are committed to the success of the project.

The Final Management Framework shows that all the data created by users (information creators) should be classified or identified, and proactively marked before they are migrated. Data classification roles and responsibilities (e.g. data creators, data owners, data users, and data auditors) should be clearly defined within the organisation. Business rules should be examined, in order to provide a basis for data classification. The flow of sensitive data communication monitoring, as well as database activity monitoring, should be in place. Enough time should be planned for the data migration process, and all the functions, applications, host servers, and storage impacted by the data migration, should be identified during the data migration. All the data in the servers, memory and buffers, should be de-staged to disc before performing migrations. It is important for organisations to know the timing of migration, the migration duration period, and the system's downtime (if necesssary). Scripts (if used) during the migration should be reviewed for reliability and accuracy.

The Final Management Framework indicates that migrated sensitive data should always be encrypted during and after migration, and should only be decrypted when the data is accessed by the authorised user, for readability. The necessary monitoring and risk assessment systems should be in place. The issues of data corruption, missed data or data loss, should be considered during migration. Migrated data should be tested and validated after migration, in order to ensure data accuracy and integrity. The network bandwidth capacity utilisation needs to be measured before migration and there is a need to know the network availability in order to ensure smooth migration. Verification or comparing migrated data with source data should be performed, and if problems persist, then a data quality process should be performed. Standards such as ISO/IEC 17799 should be adhered to when compiling security policies and procedures, in order to ensure protection of information during migration.

Organisations should enforce hardware and software standards in order to eliminate unknown factors that might access their sensitive information. Organisations should have the required tools, applications, databases, servers and data migration strategies in order for them to have a successful migration.

Business rules should be examined in order to provide a basis for data classification. The flow of sensitive data communication monitoring as well as database activity monitoring should be in place.

Organisations should use Continuous Data Protection (CDP) technology and Data Loss Prevention (DLP) tools to protect sensitive information during data migrations. Scripts (if used) during the migration should be reviewed for reliability and accuracy.

Migrated data should be tested and validated after migration in order to ensure data accuracy and integrity. Migrated sensitive data should always be encrypted during and after migration and should only be decrypted when the data is accessed by the authorised user for readability. The empirical results are integrated with insights from the literature in the sections below.

The roles and responsibilities of the migration team members should be clearly defined before the commencement of the project. Dhillon and Backhouse (2000) have stressed the importance of the integrity, roles and responsibilities of users as good values of information security management. Users are seen as the weakest connection in the information security chain (Schneier 2000), so, the information security function of each user should be seen as part of information security (Albrechtsen 2007). Albrechtsen (2007) further reiterates that users should be made to know their role in the total information security process.

Organisations should provide training and awareness of sensitive information protection and handling. Training of employees in detecting manipulative attempts is one of the methods proffered by CPNI (2009) to protect organisations against manipulation and sabotage risks. Security topics and requirements should be part of the normal business behaviour by having a clear policy and educating employees (Colwill 2009). Induction courses should cover various aspects of the risks attached to the management of sensitive data. Training should spell out the consequences of the misuse of sensitive data and also the risk of not protecting sensitive data. User awareness of the risks of their organisation's information systems has been identified by Humphreys (2008) to be part of good business practice. This might be in the form of regular awareness briefings, newsletters and circulars and the organisational awareness programme should be re-examined and continuous brought up to date when necessary.

All employees should be educated about the different classification levels, their respective markings and when to apply them. Employees should value accountability when they handle sensitive data and handle sensitive information with care as outlined in their data security policy. Employees need to be aware of what is sensitive information and how it should be protected within organisations having a process to identify sensitive information that is worth protecting. Employees working on sensitive data should undergo vetting in order to ascertain their confidential sensitivity levels. Colwill (2009) states that it is essential for organisations to perform effective employee background checks and vetting before they commence work and that the vetting process should apply to all staff levels, most especially to management and employees allocated to roles with powerful privileges, e.g. those with access to sensitive information. Members of the migration team must be certified at least up to a secret level.

Organisational strategy should include the protection of sensitive information and should be aligned with clear objectives on how sensitive data should be handled. Protecting sensitive information should be part of any organisational corporate culture. Some authors have recognised that an organisation's security culture is an important factor when maintaining an adequate information systems security level in their organisations (Ruighaver *et al.* 2007; Nosworthy 2000; Borck 2000; Von Solms 2000; Beynon 2001). According to Borck (2000), organisations willing to have effective security must also involve the corporate culture when they deploy the latest technology. Cultural change needs to be managed as Colwill (2009) indicates since it can lead to fear, uncertainty and doubt in employees, and these can have an effect on employees' atitudes towards security.

Organisations should have a data security policy which lists data security methods and sensitive data management. These procedures and the policy should be regularly communicated and enforced to all staff. There should be a continual update of the data security policy, and data integrity should be the hallmark of any organisation. This is also the view of Ross (2008) and Kavanagh (2006) that organisations should have a policy in place and the policy as well as the standards need to be enforced by the level of management that does the enforcing. Security models should be developed to support organisational strategy and such models should ensure confidentiality, integrity and reliability of data to protect sensitive information. Security is related to change management and the change management should be properly communicated to end users to ensure that they receive it well in their organisation (Ashenden 2008). There should be sufficient communication on information security with end users by management.

The organisational data access by employees should be controlled and monitored and organisational data should be defined through data discovery and classification. Employees should be given access based on their job's role and the information they are required to perform in their duties (Humphreys 2008). He points out that there should be separation of duties in order to enhance access protection against the insider threat. Confidentiality, integrity, identifying authorised users and monitoring access should be undertaken by organisations to ensure sensitive data protection. McCue (2008) has pointed out that research shows that 70% of computer fraud is perpetrated by insiders but 90% of security controls and monitoring are concentrated on external threats. Technical controls must be used to prevent unauthourised data access and they should not be used in an isolated manner (Jones & Colwill 2008).

Organisations should enforce hardware and software standards in other to eliminate unknown factors that might access their sensitive information. Organisations should have the required tools, applications, databases, servers and data migration strategies in order for them to have a successful migration. Organisational networks should always be protected at all times. Proper integration of people, process and technology should be undertaken in order to facilitate successful information security management (Eminagaoglu *et al.* 2009). Organisations should provide for continual management of data sensitivity and risk management. Eminagaoglu *et al.* (2009) indicate that organisations must always audit, check and measure their tasks within any information security programme.

All the data created by users (information creators) should be classified or identified and proactively marked before they are migrated. Data classification roles and responsibilities (e.g. data creators, data owners, data users, and data auditors) should be clearly defined within the organisation. Business rules should be examined in order to provide a basis for data classification. The flow of sensitive data communication monitoring as well as database activity monitoring should be in place.

Sufficient time should be planned for the data migration process and all the functions, applications, host servers, and storage impacted by the data migration should be identified during the data migration. All the data in the servers, memory, and buffers should be de-staged to disc before performing migrations. It is important for organisations to know the timing of migration, the migration duration period, and the systems down time (if necesssary). Scripts (if used) during the migration should be reviewed for reliability and accuracy.

Organisations should use Continuous Data Protection (CDP) technology and Data Loss Prevention (DLP) tools to protect sensitive information during data migrations (Nawafleh *et al.* 2013). The source data should be backed up prior to data migrations to the destination. Backups should be managed properly since they can cause critical points of weakness (Humphreys 2008). He suggests the encryption of backup tapes and using e-vaulting of data to protect sensitive information. The issues of data corruption, missed data or data loss should be considered during migration. Migrated data should be tested and validated after migration in order to ensure data accuracy and integrity. Technical controls should be in place to ensure effective sensitive data protection during migrations. In addition, the view of Colwill (2009) is that encryption, access control, monitoring, auditing and reporting should be part of the technical controls against insider attacks.

Migrated sensitive data should always be encrypted during and after migration and should only be decrypted when the data is accessed by the authorised user for readability. The necessary monitoring systems and risk assessment systems should be in place. Colwill (2009) has argued that a holistic approach that includes human factors, technical controls and implementing focused risk assessments are necessary to protect the organisation from the malicious insider attacker. The network bandwidth capacity utilisation needs to be measured before migration and when the network bandwidth will be available to ensure smooth migration. Verification or comparing migrated data versus source data should be performed, and if problems persist, then a data quality process should be performed.

The attitude and behaviour of the migration team members should be taken into consideration before the composition of the team. The migration team should be composed of dedicated and enthusiastic staff who are committed to the success of the project. It is very important that members of the migration team have the right attitude and behaviour and that they also adhere to the organisational security policies and procedures. Albrechtsen and Hovden (2010) highlight that there is a need for user awareness and good behaviour as part of the important aspects of the information security performance. Employee awareness and training are important, but equally changing the behaviour of employees through targeted training should be employed by educating employees on unacceptable, and non-malicious behaviour (Sasse *et al.* 2007). Organisations should reward and reinforce good security behaviour (Kavanagh 2006).

The Total Cost of Ownership of the migration projects should be computed during the migration planning stage to facilitate the completion of the migration project within its initial budget allocation. The benefits, value and the return on investment must be explored before embarking on the migration project in order to ensure that the migration project is beneficial for the organisation.

Standards such as ISO/IEC 17799 should be adhered to when compiling security policies and procedures in order to ensure protection of information during migration. Organisations have applied best information security practice for decades and many of them are incorporated into the international standards such as ISO/IEC 27001 and ISO/IEC 27002 (Humphreys 2008). Such standards can be used to monitor and control the migration processes. The standards would give the best practice baseline for IT governance since they are the basis of the foundations of information security. Humphreys (2008) emphasises that due diligence should be performed to reveal risks and manage them in terms of information security of organisational assets and their protection. This should be done by implementing effective systems of control and undertaking regular monitoring and reviews. He maintains that organisations should embark on information security governance in order for them to protect their information assets.

6.6.2. Validation of the Management Framework

All ten of the interviewees were asked whether they believe the researcher addressed all the issues pertaining to the properties of the resulting management framework on information sensitivity during software migrations. They all unanimously agreed that the resulting management framework is comprehensive. One of them said that '…in my mind it looks good'. Another person said that '…actually you've covered a lot.' Another person said that '…there is nothing to add…it is a full-fledged framework as long as those things are followed. It is fine.' Another person said that '…l think you really covered it…l think this is very much comprehensive to be honest.' Another person said that '…l think you've covered even more than what l was expecting…l think they are all covered…So for me l think you've covered quite in detail.' Another person said that '…Yah, for me l think we've done everything in terms of migration or management framework of sensitive information.' Another person said that '…l think everything is covered, when l am looking at management framework on information sensitivity, everything is covered and according to our interview it makes it a good framework.' In conclusion, the management framework on information sensitivity during software migrations is valid and reliable as verified by the qualitative analysis conducted in this research.

6.7 Ethical Considerations

Ethical issues that the researcher avoided include:

- *Bias* – which is a deliberate attempt to hide the research findings in the study or showing something disproportionately that it's very true. The researcher did not hide any research findings in this study. The raw data is provided on CD and can be checked by auditing researchers.
- *Inappropriate research methodology* – either by selecting a sample that is highly biased or using questionnaires that are not valid or using wrong conclusions. The questionnaires were proven to be valid by using item analysis indicating that no wrong conclusions were made in the study.
- *Invalid reporting* – reporting findings that serve the researcher's or someone else's interest. The researcher sent out questionnaires for the quantitative part of the study and also interviewed some IT practitioners to ensure that there was no invalid reporting during the research.
- *Using information inappropriately* – using information that directly or indirectly has adverse effects on the respondents. The researcher did not disclose the identities of all the respondents and the interviewees to ensure that information was not used inappropriately that can cause adverse effects on the respondents in this study.

Some of the issues raised above are in line with what Bryman (2004) states on ethical issues during research since he identifies four main concerns:

- Harm to participants
- Lack of informed consent
- Invasion of privacy
- Deception

There was no harm to participants, consent was obtained from all the respondents, no invasion of privacy and participants were not subjected to deception (Refer to Appendix D – Research participation form and ethics committee letter).

6.8 Conclusion

The final management framework is developed and validated in this chapter using qualitative data analysis. The qualitative data analysis steps followed in this research are presented. The qualitative interviews were transcribed and imported into the NVIVO software and data analysis was performed.

The narratives from the transcripts of the ten interviews conducted were presented and used in conjunction with the outcomes of the coding process from the NVIVO software to develop the management framework on information sensitivity during software migrations. The management framework was validated by interviewing experienced IT specialists in the South African government departments and agencies and the outcomes of the validation presented. The next chapter presents the conclusions and recommendations for future work.

Chapter 7

Conclusion and Recommendations for Future Work

7.1 Introduction

The previous chapter is an analysis of the qualitative data and conceptualisation of discussion of the management framework on information sensitivity during the migration of software platforms. The NVIVO software was used to carry out the qualitative data analysis in the previous chapter in order to generate categories and their hierarchy to gather the themes that are generalised in the resulting management framework. This chapter concludes the thesis on the development of a management framework on information sensitivity during migrations of software platforms. A synopsis of all the previous chapters and a general conclusion on the research are presented in this chapter. This chapter concludes with the recommendations for future work. The content of this thesis was synthesised into a research publication (Ajigini *et al.* 2016).

7.2 Limitations of the Research

The researcher got ninety respondents and could not get a high number of respondents (say at least 150) to complete the questionnaires. Some of the government departments/agencies turned down the researcher's request to have their employees participate in the study due to the nature of the research which focuses on information security. These organisations fear that their sensitive information might be compromised if they allow their employees to participate in the research.

The importance of estimating the sample size required for a quantitative study (Exploratory Factor Analysis) has been highlighted by some authors (Beavers *et al*. 2013; Guadagnoli & Velicer 1988; Devane *et al*. 2004). Beavers *et al*. (2013) suggest a minimum sample size of 150 because the family of factor analysis procedures involves multivariate tools and these methods require larger sample sizes than univariate methods. Hutcheson and Sofronion (1999) suggest at least 150 to 300 sample size while Guadagnoli and Velicer (1988) predict a sample size of 150. Other authors recommend a minimum sample size of 100 to 200 (Comrey 1973, 1978; Loo 1983; Gorsuch 1983; Lindeman *et al*. 1980; Hair *et al*. 1979; Guilford 1954). Therefore from the literature, a sample size of 150 should be adequate for Exploratory Factor Analysis. The Exploratory Factor Analysis (EFA) was performed twice (first on the items and then later on the sub-constructs) in this research in order to overcome this limitation.

7.3 Synopsis of the Research Questions

This thesis develops a management framework to manage sensitive information during the migration of software platforms. It examines the following research questions:

RQ: How should organisations manage sensitive information during its migration between software platforms?

SQ1: What are the differences between sensitive information and other information capital in an organisation?

SQ2: How can protection mechanisms be implemented during the migration of information from one platform to another, e.g. from a proprietary platform to an open source platform?

SQ3: What would be the properties of a management framework for the migration of sensitive information during platform migrations?

SQ4: Why is the management framework necessary to protect sensitive data during software migrations?

RQ is addressed in section 2.5 – information sensitivity and information classification and finally answered in section 6.6 where a management framework on information sensitivity between software migrations was developed and is illustrated in Figure 6-5. The researcher's view is that organisations should use the BS 17799 classification scheme to classify their sensitive information during migrations of software platforms since it is an international standard that has been tested and widely approved. The BS 17799 Classification scheme proposes a five-layered classification level starting from: public documents (Level 1); internal use only (Level 2); proprietary (Level 3); highly confidential (Level 4) and finally top secret (Level 5). Highly sensitive information should be classified as top secret (Level 5) and less highly sensitive information can be classified as highly confidential (Level 4). Organisations can use the management framework developed in this thesis to manage their sensitive information during its migration between software platforms.

SQ1 is answered in sections 2.2 and 6.4. In section 2.2: What are the differences between sensitive information and other information capital in an organisation? the researcher explored the literature on sensitive information definitions as listed in Table 2-1: Definitions of sensitive information. By analysing and integrating concepts from the various existing definitions the researcher formulated a synthesised definition of sensitive information (definition 2.1). Any other information that does not fall within this definition is termed to be non-sensitive information. Additionally, in section 6.4 which summarises ten interview narratives, the researcher collated and synthesised/amalgamated the responses of the ten interviewees on the question that they must state the difference between sensitive information and non-sensitive information. The interviewees were able to point out and confirm the main difference between sensitive information and non-sensitive information. Using the definition 2.1, sensitive information is defined as:

Definition 2.1:

Sensitive information is protected information that the owner does not want to reveal to others and which is not to be divulged outside the organisation, as well as information concerning an individual’s ethnic origin or race, criminal record, sexual preferences or practices, and other information that include political beliefs, political association membership, trade union membership, religious associations or philosophical opinions; efforts should be made to conceal such information not being revealed to other people.

Non-sensitive information is any information that does not conform to the definition 2.1.

SQ2 is best resolved by the management framework on information sensitivity during software migrations that was developed in this thesis. A sequence of steps was followed to obtain the final management framework on information sensitivity during software migrations. To begin with, in section 2.14: ‘Towards a rudimentary management framework’, the properties of the management framework obtained from the literature were conceptualised to form the rudimentary management framework on information sensitivity during software migrations as illustrated in Figure 2-2 and Table 2-5.

Secondly, in section 3.5: ‘A model for addressing the security challenges during migration to OSS’ was conceptualised with inputs from the rudimentary management framework of section 2.14 and other concepts from the literature. This led to the conceptualisation of Figure 3-1: ‘Modelling security challenges during OSS migrations' which was used to conceptualise the later stage in the conceptualisation of the management framework.

Thirdly, in section 5.11: 'Resulting framework from the quantitative analysis results', the Preliminary Management Framework was developed using inputs from the former Rudimentary Management Framework and the model for addressing security challenges during migration to OSS. A Preliminary Management Framework was conceptualised as illustrated by Figure 5-27. Finally, a management framework on information sensitivity between software migrations was developed in section 6.6 and is illustrated in Figure 6-5. The protection mechanisms, ought to be implemented during the migration of information from one platform to another, can be effectively performed by utilising the final management framework in Figure 6-5.

SQ3 is addressed in section 2.13: 'Properties of a management framework' and sub-section 6.6.1: 'Discussion of the management framework on information sensitivity.' These are the building blocks of the management framework of Figure 6-5, and an explanation of how they have contributed to the management framework is given in sub-section 6.6.1.

SQ4 is addressed in section 1.1.2, which focuses on the statement of the problem as well as section 2.3, which focuses on the management of sensitive information. A management framework is necessary to protect sensitive data during software migrations because information in organisations has to be protected according to its sensitive levels, how critical it is and its value, irrespective of the storage media, the processing systems (manual or automated), or the information distribution methods. The protection of information - in accordance with its sensitivity - is substantiated by section 5 of the ISO17799 standard which stipulates that it is essential to perform information classification according to its actual value and sensitivity levels in order to implement the appropriate security level. Corporations have been motivated to invest in information security by safeguarding their confidential data and their customers' personal information (Kalyvas *et al.* 2013; Acquisto *et al*. 2006). The non-protection of sensitive information can damage the reputation of an organisation (Kalyvas *et al.* 2013; Rasmussen 2008). Therefore, a framework is necessary to protect sensitive information during migration of such data.

7.4 Synopsis of the Thesis Chapters

Chapter one begins with an explanation of the purpose of the study and a clarification of the statement of the problem. In this study, the researcher develops and validates a management framework for the migration of sensitive information during the migration of platforms (only software platforms) by using a case study methodology and a mixed method approach. The study concentrates on South African government departments and parastatals that have performed software migrations. The main focus is the development and validation of a management framework for the migration of sensitive information between platform migrations. This is necessary because, as far as the researcher could determine, there is currently no management framework to manage information sensitivity between migration of platforms in academic research.

The migration problems encountered by some South African departments and agencies are highlighted. These include unauthorised access; information theft; information leakage; phishing; and stealing of sensitive information. All the government departments and agencies that participated in the research are chosen because they have performed software migrations most especially from proprietary to open source. The goals and objectives of the research are explained. Some of these are (a) the conceptualisation of information sensitivity, (b) defining the protection measures that should be undertaken during the migration of software platforms and (c) the development of a management framework that can be used to protect sensitive information during platform migrations.

The contribution of the research to the IS world and body of knowledge is explained and this is the enrichment of the theory of information systems with respect to information sensitivity management. A management framework that can be used to protect and handle sensitive information during migration of software platforms was developed and the main work of this thesis had been published by the African Journal of Information Systems (AJIS) (Ajigini *et al.* 2016). A brief explanation of research design and methodology is discussed motivating the use of a case study methodology and a sequential explanatory mixed methods approach. The data gathering and the data analysis processes are explained with the data analysis consisting of both quantitative and qualitative methods in a mixed methods research setting.

Chapter two involves the critical reviewing of the literature on the main aspects of the research. The definition of sensitive information is synthesised from the definitions given by some authors in the literature. A mathematical description of sensitive information is developed and explained. The importance of information sensitivity and classification is also emphasised. Information classification relates to tagging the organisational data so that the necessary protection mechanisms can be applied to various levels of classified data. The BS17799 standard is the recommended information classification scheme that organisations can use to protect their sensitive information. On the strength of this standard, a formal description of sensitive information is developed and discussed. The migration from closed source software to open source software is examined in terms of the security challenges and how they can be overcome by using the rudimentary framework. The properties of the rudimentary management framework are highlighted in Table 2-3 and the rudimentary framework is conceptualised from these properties of the framework.

In Chapter three, the literature is critically reviewed focusing on migrations from a proprietary platform to an OSS platform. The security challenges during OSS migrations are discussed, notably phishing, stealing sensitive information e.g. account details and cookies and getting hacked during the process (Mtsweni & Bierman 2008). The OSS and CSS security are compared and the security of OSS is found to be roughly of the same quality as that of CSS. The various challenges that are encountered during migrations from proprietary to OSS are identified and discussed. Some technical challenges include security; data migration; and OSS code maintainance (Sarrab *et al.* 2013; ElHag & Abushama 2009). A model for addressing the security challenges during migration to OSS is presented in section 3.5 (Figure 3-1). This model is conceptualised from the Anner and Cid (2010) open source assessment framework and the rudimentary management framework in Chapter 2.

Chapter four details the research design and the research methodology in which various philosophical perspectives are explained. Positivism is the testing of theory in order to increase the predictive understanding of a phenomenon (Myers 1997) using data collection methods such as sample surveys, controlled experiments and inferential statistics (Johari 2009). Positivism follows the universalist methods of science principle which disagrees that there is a fundamental difference between the social and the natural sciences (Myers & Klein 2011). Interpretivism involves understanding the phenomenon within cultural and contextual situations (Orlikowski & Baroudi 1991). IS research can be regarded as being interpretive if based on the assumption that our knowledge of reality is obtained via social constructions like tools, documents, language and other artifacts (Klein & Myers 1999).

Pragmatism is the philosophical approach for mixed methods research (Johnson & Onwuegbuzie 2004) since pragmatism calls for research that combines theory and practice as suggested by Ormerod (1996). Pragmatism is a philosophy of science based on actions (Taatila & Raij 2012; Dewey 1929; Peirce 1992; Blosch 2001). Pragmatism is positioned toward solving practical problems in the real world (Feilzer 2010) and it is not based on the postulations about the nature of knowledge. Pragmatism is the underlying philosophical paradigm used in this research.

The research methods are explored in Chapter four. These include (a) induction, deduction and abduction reasoning; (b) qualitative research method; (c) quantitative research method; (d) mixed methods research. The author used sequential explanatory mixed methods research to perform the research under a pragmatist stance, specifically the sequential exploratory mixed methods design used as described in Table 4-1. The case study research methodology is highlighted. This research is a sequential explanatory mixed method case study. Data collection was done for both quantitative and qualitative methods. In quantitative research, item analysis was performed on the 25 responses received for data analysis. This was carried out in order to identify the items that form internally consistent constructs and to eliminate those that do not in order to reflect the extent of inter-correlation among the items by using the Cronbach's Alpha Coefficient (α).

Chapter five discusses the development of the preliminary management framework using statistical techiques. This involves descriptive statistics and correlation statistics. The JMP version 11 software was used to perform the data analysis. Ninety responses were received out of 250 questionnaires that were distributed to IT specialists in the visited government organisations, thus giving a response rate of 36 percent. Firstly, the biographical data distributions were analysed and the outcomes are illustrated in Figures 5-1 to 5-10. Exploratory Factor Analysis (EFA) is used to identify new constructs that are valid. Ten (10) sub-constructs are found to be valid and reliable. Reliability of the sub-constructs was undertaken using item analysis (Cronbach's Alpha Coefficient). Estimates of internal consistency as measured by Cronbach’s Alpha all exceeded 0.80 with the exception of three sub-constructs that score less than 0.70 and are reported on in Table 5-6. This indicates good reliability for the seven sub-constructs that have Cronbach’s alpha exceeding 0.80.

EFA was performed on the ten sub-constructs to obtain the four main constructs as illustrated in Table 5-18. The four main constructs are described and compared using means and standard deviation as shown in Table 5-19. The reliability of the four main constructs is illustrated in Table 5-20 and the table shows that all the four factors are highly reliable. The Spearman’s correlation between the main constructs were performed and it was found that the significance of the displayed paired main constructs are mostly medium and strong as depicted in Table 5-21. The relationships among the four main constructs with Spearman’s correlation coefficients are shown in Figure 5-26 and this was expanded to obtain the preliminary management framework (Figure 5-27). The resulting framework from the quantitative analysis is explained in section 5.11 by using Figure 5-27.

Chapter six entails the outcomes of the qualitative data analysis and the development of the final management framework on information sensitivity during software migrations. Ten semi-structured interviews were performed across all the participating government organisations in Pretoria, South Africa. The NVIVO Version 10 software was used to perform the qualitative data analysis. The software was used to develop categories and sub-categories in the qualitative data and these categories were used to form the themes of the management framework. The various steps undertaken during the qualitative data analysis are explained.

This involves the transcription of the ten interviews and then importing them into the NVIVO software for further data analysis. An explanation on the interview narratives is provided. The categories and sub-categories developed were then used to code the transcripts of the ten interviewees. The first 100 most frequently occuring words in the qualitative data are displayed as a word cloud diagram (Table 6-2); tree map (Figure 6-3); and cluster analysis (Figure 6-4). New variables were identified after the qualitative analysis and these variables were then included in the preliminary management framework (Figure 5-27) to obtain the final management framework (Figure 6-5). The final management framework on information senstivity during software migrations is presented in section 6.6. This management framework was then further validated and the result is presented in sub-section 6.6.2.

7.3.1 Concluding Remarks

In conclusion, it was important to do this research because it developed a management framework that can be used to handle and protect sensitive information during software migrations. The resulting final management framework shown in Figure 6-5 is a fully-fledged, conscise, valid and reliable management framework that organisations may utilise to assist them to protect their classified sensitive information during migrations of software platforms.

Recommendations for Future Work

It is the hope of the researcher that this research on information sensitivity has suggested avenues for future scholars to more deeply investigate the research questions outlined in this thesis. The question that will define the next challenge for research and practice in the area of information sensitivity will be how to deploy this management framework in organisations. This leads to new challenges for the conceptualisation and development of what areas of the management framework truly need to be redesigned to realise the full potential of the management framework.

Research to understand human conceptualisations of sensitive information and also to understand the difference between sensitive information and non-sensitive information needs to be further pursued. Research into the design of data classification systems in organisations - with a view to enhance the quality of the data sets to be classified, the information audit and how the information will be physically stored and categorised over its lifetime should be further investigated. The formal description of sensitive information developed in Chapter 2 needs to be validated and enhanced through deeper investigations into the sectors identified.

Lastly, research to confirm the final management framework by using Confirmatory Factor Analysis (CFA) with more emphasis on Structural Equation Modelling (SEM) needs to be further explored. This may require the development of new research questions and hypotheses. Additionally, the questionnaire will be refined according to the Structural Equation Modelling.

References

Abraham, S., 2011. ‘Information security behavior: factors and research directions.’ *Proceedings of the 17th Americas Conference on Information Systems (AMCIS),* Detroit, Michigan.

Acello, R., 2009. ‘Feds ready to tackle cybercrime.’ *ABA Journal*, 95(2), 37.

Acquisto, A., Friedman, A. & Telang, R., 2006. ‘Is there a cost to privacy breaches?: An event study.’ *27th International Conference on Information Systems*, Milwaukee.

Agerfalk, P. J., 2010. ‘Getting pragmatic.’ *European Journal of Information Systems*, 19, 251 – 256.

Ahmad, A., Bosua, R. & Scheepers, R., 2014. ‘Protecting organizational competitive advantage: A knowledge leakage perspective.’ *Computers & Security*, 42, 27 – 39.

Ahmad, A., Maynard, S. B. & Park, S., 2014. ‘Information security strategies:

towards an organisational multi-strategy perspective.’ *Journal Intelligent Manufacturing*, 25, 357 – 370.

Ajigini, O. A., Van der Poll, J. A. & Kroeze, J. H., 2012. ‘Towards a Management Framework to protect sensitive information during migrations.’ The 2nd International Conference on Design and Modeling in Science,

Education and Technology (DeMSet), Orlando, Florida, USA, (ISBN-13: 978-1-936338-76-4) ISBN-13: 978-1-936338-76-4 CD / ISBN-13, 6-13.

Ajigini, O. A., Van Der Poll, J. A. & Kroeze, J. H., 2014. ‘Towards a Model on

Security Challenges during Closed Source Software to OSS Migrations.’ *The 9th International Conference for Internet Technology and Secured Transactions (ICITST Proceedings),* London, UK, 8 – 10 Dec., 275-284. ISBN: 978-1-908320-31-5.

Ajigini, O. A., Van der Poll, J. A. & Kroeze, J. H., 2016. ‘A Framework to Manage Sensitive Information during its Migration between Software Platforms.’ *The African Journal of Information Systems*, Vol. 8, Issue 2, Article 2, 21 – 44. ISSN 1936-0282.

Albrechtsen, E., 2007, ‘A qualitative study of users’ view on information security.’ *Computers & Security*. 26(4), 276 – 289.

Albrechtsen, E. & Hovden, J., 2010. ‘Improving information security awareness and behaviour through dialogue, participation and collective reflection – An intervention study.’ *Computers & Security*, 29(4), 432 – 445.

Alhazmi, O., Malaiya, Y., & Ray, I., 2007. ‘Measuring, analyzing and predicting security vulnerabilities in software systems.’ *Computers & Security*, 26(3), 219 – 228.

Allen, J. P., 2012. ‘Democratizing business software: Small business ecosystems for open source applications.’ *Communications of the Association for Information Systems*, 130 (28), 483-496.

Alomari, Z., Halimi, O. E. & Sivaprasad, K., 2015. ‘Comparative studies of six programming languages,’ *arXiv*:1504.00693. Available from: arxiv.org/labs/1504.00693. Accessed on 2015/10/05.

ALRC, 2000. ‘ALRC report 108.’ Available on www.alrc.gov.au.

Aner, Y. & Cid, C., 2010. ‘Open-source security assessment.’ Royal Holloway Series, University of London. Accessed on 2012/02/17 and available on:

http://media.techtarget.com/searchSecurityUK/downloads/RHUL_Yoav_v2.pdf

Arai, M. & Tanaka, H., 2009. ‘A proposal for an effective information flow control model for sharing and protecting sensitive information.’ Australasian Information Security Conference (AISC), Wellington, New Zealand. *Conferences in research and practice in Information Technology (CRPIT)*, Ljiljana Brankovic and Willy Susilo, Eds.

Ashenden, D., 2008. ‘Information security management: A human challenge?’ *Information Security Technical Report*, 195 – 201.

Atkinson, D. 2015.’Approaches and strategies of social research.’ Available from: http://minyes.its.rmit.edu.au/~dwa/Methods.html. Accessed on 2015/12/02.

Augustinos, T., 2009. ‘Preventing and reacting to a data breach.’ *Risk Management*, 56(10), 45.

Azorin, J. M. & Cameron, R., 2010. ‘The application of mixed methods in organisational research: A literature review,’ *The Electronic Journal of Business Research Methods*, 8(2), 95 – 105. Available online at www.ejbrm.com.

Bamberger, M., 2007. ‘A framework for assessing the quality, conclusion validity and utility of evaluations: Experience from International development and lessons for developed countries.’ AEA.

Baskerville, R. & Myers, M., 2004. ‘Special issue on action research in information systems: making IS research relevant to practice – foreword.’ *MIS Quarterly*, 28(3), 329 – 335.

Bataller, E., 2012. ‘Data classification tips and technologies.’ *InformationWeek Report*, available from: www.networkcomputing.com/unified-communication/data-classification-tips-and-technologies/d/d- id/1233510? Accessed 2012/02/25.

Bayuk, J., 2009. ‘Data-centric security.’ *Computer Fraud & Security*, 7 – 11.

Beavers, A. s., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J. & Esquivel, S. L., 2013. ‘Practical considerations for using exploratory factor analysis in educational research.’ *Practical Assessment, Research and Evaluation Journal*, 18(6).

Becker, J. & Niehaves, B., 2007. ‘Epistemological perspectives on IS research: a framework for analyzing and systematizing epistemological assumptions.’ *Info Systems Journal*, 17, 197 – 214.

Benbasat, I., Goldstein, D. K. & Mead, M., 1987. ‘The case study research strategy in studies of Information Systems.’ *MIS Quarterly*, 11(3), 369 – 386.

Benzies, K. M. & Allen, M. N., 2001. ‘Symbolic interactionism as a theoretical perspective for multiple method research.’ Blackwell Science Ltd., *Journal of Advanced Nursing*, 33(4), 541 – 547.

Beynon, D. 2001. ‘Talking heads.’ *Computerworld*, 24(33), 19 – 21.

Bhatt, K. & Dongre, A., 2014. ‘A survey of sensitive information hiding techniques.’ *International Journal of Emerging Technology and Advanced Engineering*, 4(1), 402 – 406.

Bhattacherjee, A., 2012. *Social science research: Principles, methods and practices.* University of South Florida, USA.

Biot-Paquerot, G. & Hasnaoui, A., 2009. ‘Stakeholders’ perspective and ethics in financial information systems.’ *Journal of Electronic Commerce in Organisations*, 7(1), 59 – 70.

Bleek, W., & Finck, M., 2011. ‘Migrating a development project to open source software development.’ Available from: www.flosshub.org/system/files/bleek10-14.pdf.

Bloch, C., Sorensen, M. P., Graversen, E. K., Schneider, J. W., Schmidt, E. K., Aagaard, K. & Mejlgaard, N., 2014. ‘Developing a methodology to assess the impact of research grant funding – A mixed methods approach.’ *Journal of Evaluation and Program Planning*, 43, 105 – 117.

Blosch, M., 2001. ‘Pragmatism and organisational Knowledge Management.’ *Knowledge and Process Management*, 8(1), 39 – 47.

Boland, R., 1979. ‘Control, causality and information system requirements.’

Accounting, Organisations and Society, 4(4), 259 – 272.

Borck, J., 2000, 'Advice for a secure enterprise: implement the basics and see that everyone uses them.' *InfoWorld*, 22(46), 90.

Bradley T., 2007. 'Securing sensitive information: Protecting your network against information leakage.' *CISSP-ISSAP*, Microsoft MVP - Windows Security.

Brin, S., Davis, J. & Garcia-Molina, H., 1995. 'Copy detection mechanisms for digital documents.' Proceedings of the *ACM SIGMOD* International Conference on Management of Data, San Jose, CA, USA, May 22 – 25, *ACM*, 398 – 409.

British Standards Institute, 2000. 'Information Technology code of practice for information security management (Standard 0-580-36958-7).' London: British Standards Institute.

Brown, C. V., DeHayes, D. W., Hoffer, J. A., Martin, E. W. & Perkins, W. C., 2014. '*Managing information technology.'* 7th edn. Upper Saddle River, NJ: Prentice Hall.

Brown, J. D., 2000. 'What is construct validity.' *JALT testing & evaluation, SIG Newsletter*, 4 (21), 8 – 12.

Bruce, L. S., 2003. 'Information Security – Key issues and developments.' Retrieved on November 2, 2009 from:

www.pwcglobal.com/jm/images/pdf/information%20Security%20Risk.pdf.

Bryman, A., 2004. *Social research methods*. 2nd edn. Oxford University Press, Oxford.

Bryman, A., 2006a. 'Integrating quantitative and qualitative research: how is it done?' *Qualitative Research*, 6(1), 97 – 113.

Bryman, A., 2006b.'Paradigm peace and the implications for quality.' *International Journal of Social Research Methodology*, 9(2), 111 – 126.

Bryman, A., 2007. 'Barriers to integrating quantitative and qualitative research.'

Journal of Mixed Methods Research, 1(8), 8 – 22.

Buchanan, J, & Jones, M. L., 2010. 'The efficacy of utilising NVIVO for interview data from the electronic gaming industry in two jurisdictions.' *Review of Management Innovation & Creativity,* 3(5), 1 – 15.

Cala, J., Czekierda, L. & Zielinski, K., 2004. 'Migration aspects of Telemedical software architectures.' Studies in Health Technology and Informatics, 105, 80 – 91.

Cameron, R., 2011.'Mixed methods research: The five Ps framework.' *The Electronic Journal of Business Research Methods,* 9(2), 96 – 108.

Carroll, J. & Swatman, P., 2000. 'Structured-case: a methodological framework for building theory in information systems research.' *Proceedings of the 8th European Conference on Information Systems,* Vienna University, Vienna, 116 – 123.

Casey, E., 2006. 'Investigating sophisticated security breaches.' *Communications of the ACM,* 49(2).

Cate, F. H., 2006. 'The privacy and security policy vacuum in higher education.' *EDUCAUSE Review,* 41(5).

Cavaye, A., 1996. 'Case study research: A multi-faceted approach for IS.' *Information Systems Journal,* 6, 227 – 242.

Cecez-Kecmanovic, D., Klein, H. K., & Brooke, C., 2008. 'Exploring the critical agenda in information systems research.' *Information Systems Journal,* 18(2), 123 – 135.

CENATIC, 2008. 'Open source software for the development of the Spanish public administration: An overview.' Available from: www.cenatic.es

CENATIC, 2009. 'Study on the situation of open source software in universities and R&D centers.' Report, Almendralejo. Available from: www.cenatic.es

Cepeda, G. & Martin, D., 2005. 'A review of case studies publishing in management decision 2003 – 2004, Guides and criteria for achieving quality in qualitative research.' *Management Decision*, 43(6), 851 – 876.

Chang, M., Hung, Y., Yen, D. C. & Tseng, P. T. Y., 2009. 'The research on the critical success factors of knowledge management and classification framework project in the executive Yuan of Taiwan government.' *Expert Systems with Applications*, 36, 5376 – 5386.

Chavhan, N. B., Wankhade, P. J. & Tagalpallewar, S. K., 2013. 'Implementation of data leakage detection & protection using allocation strategies.' *International Journal for Engineering Applications and Technology*, 174 – 181.

Chen, W. & Hirschheim, R., 2004. 'A paradigmatic and methodological examination of information systems research from 1991 to 2001.' *Information Systems Journal*, 14, 197 – 235.

Chengalur-Smith, I., Nevo, S. & Demertzoglou, P., 2010. 'An empirical analysis of the business value of open source infrastructure technologies.' *Journal of the Association for Information Systems*, 11, Special issue, 708 - 729.

Choi, Y. B., Hayward, A., Forkey, S. J. & Griffin, R., 2014. 'Information systems management in government: Ongoing issues and approaches.' *International Journal of Computer and Information Technology*, 3(5), 993 – 998.

Chua, W. F., 1986. 'Radical developments in accounting thought.' *The Accounting Review*, 61, 601 – 632.

Cicoria, S., Sherlock, J., Clarke, L. & Muniswamaiah, M., 2012. *Open source software opportunities and risks*.

Clake, R., Dorwin, D. & Nash, R., 2009. 'Is open source software more secure?' Homeland Security/Cyber Security, viewed 2 February 2003, from www.cs.washington.edu/education/courses/csep590/05au/whitepaper_turnin/oss%2810%29.pdf.

Clark, A. M., 1998. ‘The qualitative-quantitative debate: moving from positivism and confrontation to post-positivism and reconciliation.’ *Journal of Advanced Nursing*, 27, 1242 – 1249.

Clark, V. L. P., 2005. ‘Cross-disciplinary analysis of the use of mixed methods in physics education research, Counseling Psychology, and Primary Care.’ Doctoral Dissertation, University of Nebraska-Lincoln.

Clark, V. L. P., Creswell, J. W. & Shope, R. J., 2010. ‘Mixing quantitative and qualitative approaches, An introduction to emergent mixed methods research.’ In Handbook of Emergent Methods, Edited by Hesser-Biber, S. N. and Leavy, P., The Guilford Press, New York.

Cleary, M., Horsfall, J. & Hayter, M., 2014. ‘Data collection and sampling in qualitative research: does size matter?’ *Informing Practice and Policy Worldwide through Research and Scholarship*, 473 – 475.

Coaley, K., 2010. *An introduction to psychological assessment and psychometrics*. London, England: Sage.

Cole, R., Purao, S., Rossi, M. & Sein, M. K., 2005. ‘Being proactive: Where action research meets design research.’ *Proceedings of the 26th International Conference on information systems, Association for Information Systems*, Atlanta, 325 – 336

Collingridge, D. S. & Gantt, E. E., 2008. ‘The quality of qualitative research.’ *American Journal of Medical Quality*, 23(5), 389 – 395.

Collins, K. M. T., Onwuegbuzie, A. J. & Jiao, Q. G., 2007. ‘A mixed methods investigation of mixed methods sampling designs in social and health science research.' *Journal of Mixed Methods Research*, 1(3), 267 – 294.

Colwill, C., 2009. 'Human factors in information security: The insider threat – Who can you trust these days?' *Information Security Technical Report*, 14, 186 – 196.

Comrey, A. L., 1973. *A first course in factor analysis*. New York: Academic Press.

Comrey, A. L., 1978. 'Common methodological problems in factor analytic studies.' *Journal of Consulting and Clinical Pyschology*, 46, 648 – 659.

CPNI, 2009. 'Insider attacks.' www.cpni.gov.uk/MethodsOfAttack/insider.aspx.

Cresswell, J., 1994. *Research design: qualitative and quantitative approaches*. Sage: Thousand Oaks.

Cresswell, J., 2002. *Qualitative inquiry and research design*. Sage: Thousand Oaks.

Creswell, J., 2005. *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Upper Saddle River, NJ: Merrill Prentice Hall.

Creswell, J., 2007. *Qualitative inquiry and research design: choosing among five approaches.* (2nd ed.), Thousand Oaks: Sage

Cresswell, J. W., 2009. *Research design, qualitative, quantitative and mixed methods approaches.* 3rd edn. Los Angeles: Sage.

Creswell, J. W. & Clark, V. L. P., 2007. *Designing and conducting mixed methods research.* Thousands Oaks, CA: Sage Publications.

Creswell, J. W., Tashakkori, A., Jensen, K. D. & Shapley, K. L., 2003. 'Teaching mixed methods research: Practices, dilemmas, and challenges', In: *Handbook* *of mixed methods in social and behavioural research,* A. Tshakkori and C. Teddlie (eds.). Thousand Oaks, CA: Sage Publications, 91 – 110.

Cronbach, L., 1951. ‘Coefficient alpha and the internal structure of tests.’ *Psychometrika*, 16, 297 – 334.

Cronbach, L. J., & Meehl, P. E., 1955. ’Construct validity in psychological tests.’ *Psychological Bulletin*, 52, 281 - 302.

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M. & Baskerville, R., 2013. ‘Future directions for behavioural information security.’ *Computers & Security*, 32, 90 – 101.

Crowe, S. Creswell, K, Robertson, A., Huby, G., Avery, A. & Sheikh, A., 2011. ‘The case study approach.’ *BMC Medical Research Methodology*, 11, 100.

Crowston, K., Wei, K. & Howison, J., 2012.’Free/Libre Open-source software development: What we know and what we do not know.’ *ACM Computing Surveys*, 44(2), Article 7.

Cunningham, D. J., 1998. ’Cognition as a Semiosis: The role of inference.’ *Theory and Psychology*, 8(6), 827 – 840.

Danermark, B., 2001. *Explaining Society: An Introduction to Critical Realism in the Social Sciences.* Routledge, Florence, KY.

Daniel, J., 2009, ‘Open Source vs. Closed Source Software: The Great Debate.’ viewed 1 February 2013, from: http://www.articlesbase.com/internet-articles/open-source-vs-closed-source-software-the-great-debate-1040071.html

Danish Board of Technology Working Group, 2002. ‘Open Source Software in e-government, a report on the analysis and recommendations drawn up by a working group under the Danish Board of Technology.’ Viewed 2012/04/07 from: www.teckno.dk/pdf/projekter/p03_opensource_paper_english.pdf.

Datta, L., 1994. ‘Paradigm wars: A basis for peaceful coexistence and beyond.’ In *The Qualitative-Quantitative debate: New perspectives*, C. S. Reichardt and S. F. Rallis (eds.), San Francisco: Jossey-Bass Publishers, 53 – 70.

Da Veiga, A. & Eloff, J., 2010. ‘A framework and assessment instrument for information Security Culture.’ *Computers & Security*, 29(2), 196.

Denscombe, M. 2008. ‘Communities of practice.’ *Journal of Mixed Methods Research*, 2, 270 – 283.

Denzin, N. K. & Lincoln, Y. S., 2000. *Handbook of Qualitative Research,* 2nd edn. Thousand Oaks: Sage

Department of Public Service and Administration, 2006. ‘Policy on free and open source software use for South African Government.’ available online at www.sita.co.za/FOSS/FOSS.html, accessed on 6/11/2008 at 9:48am.

Deshmukh, P. S. & Pande, P., 2014. ‘A study of electronic document security.’ *International Journal of Computer Science and Mobile Computing (IJCSMC)*, 3(1), 111 – 117.

Devani, D., Begley, C. M. & Clarke, M., 2004. ‘How many do l need? Basic principles of sample size estimation.’ *Journal of Advanced Nursing*, 47(3), 297 – 302.

Devers, K,. J., 1999. ‘How will we know good qualitative research when we see it? Beginning the dialogue in Health Services Research.’ *Health Services Research*, 34(5), 1153 – 1188.

Dewey, J., 1929. *The Quest for Certainty: A study of the relation of knowledge in action.* New York, Minton, Balch and Company.

Dewey, J., 1960. *The Quest for Certainty: A study of the relation of knowledge in action*. Capricorn Books: New York.

Dhillon, G. & Backhouse, J., 2000. ‘Information system security management in the new millennium.’ *Communications of the ACM*, 43(7), 125 – 8.

Di Bella, E, Sillitti, A. & Succi, G., 2013. ‘A multivariate classification of open source developers.’ *Information Sciences*, 221, 72 – 83.

Diffie, W., 2008. ‘Information security: 50 years behind, 50 years ahead.’ *Communications of the ACM,* 51(1).

Doinea, M., 2010. ‘Vulnerability assessment in open source distributed applications.’ *Open Source Journal,* 2(3), 107 – 114.

Doyle, L., Brady, A. & Bryne, G., 2009. ‘An overview of mixed methods research.’ *Journal of Research in Nursing,* 14(2), 175 – 185.

Drake, P., Shanks, G. & Broadbent, M., 1998. ‘Successfully completing case study research: combining rigour, relevance and pragmatism.’ *Information Systems Journal,* 8(4), 273 – 289.

Drozdik, S. & Kovacs, G. L., 2005, ‘Risk Assessment of an open source migration project.’ *Proceedings of the First International Conference on Open Source Systems*, Geneva, M. Scotto & G. Succi (eds.) 246 – 249.

Duri, S., Elliott, J., Gruteser, M., Liu, X., Moskowitz, P., Perez, R., Singh, M. & Tang, J., 2004. ‘Data protection and data sharing in Telematics.’ *Mobile Networks and Applications*, 9, 693 – 701, Kluwer Academic Publishers, Netherlands.

Dwan B., 2004. ‘Open source vs. closed.’ *Network Security,* 5, 11-13.

Edhlund, B. M. & McDougall, A. G., 2013. ‘NVIVO 10 Essentials, Your guide to the World’s most powerful qualitative data analysis software.’ Form & Kunskap AB, Sweden.

ElHag, H. M. A. & Abushama, H. M., 2009. ‘Migration to FOSS: Readiness and challenges.’

EMC Corporation, 2011. ‘Planning a data center migration: Five Key Success Factors.’ Published in the USA, *09/11 EMC Perspective H6151.1*

Eminagaoglu, M., Ucar, E. & Eren, S., 2009. 'The positive outcomes of information security awareness training in companies – A case study.' *Information Security Technical Report*, 223 – 229.

Engstrom, C., 2013. 'NSA asked Linus Torvalds to install backdoors into GNU/Linux.' InfoPolicy. Available from: http://falkvinge.net/2013/11/17/nsa-asked-linus-torvalds-to-install-backdoors-into-gnulinux/. Accesed on: 2015/23/10.

Falconer, D. J. & Mackay, D. R., 1999. 'The key to the mixed method dilemma.' *Proceedings of the 10th Australasian Conference on Information Systems*, Victoria University, Wellington, New Zealand.

Falconer, D. & Mackay, D. R., 2000. 'The myth of multiple methods.' *American Conference on Information Systems (AMCIS) Proceedings*, 1467 – 1473.

Fakhri, B., Fahimah, N. & Ibrahim J., 2015. 'Information security aligned to enterprise management.' *Middle East Journal of Business*, 10(1), 62 – 66.

Farrell, G., 2002. 'Former Anderson executive to testify.' *USA Today*, p.B1, April 10.

FATF/OECD, 2009. 'Anti-money laudering and combating the financing of terrorism in South Africa, Mutual Evaluation Report.' available from: www.fic.gov.za/DownloadContent/NEWS/General/2008/FATF_ME_2009_FINAL. Accessed on 2012/03/02.

Federal Trade Commission (FTC), 2009. 'Protection of personal information: A guide for business.' Available from: http://www.ftc.gov/infosecurity, Accessed on 2012/02/23.

Feilzer, M. Y., 2010. 'Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm.' *Journal of Mixed Methods Research*, 4, 6 – 16.

Fidel, R., 2008. ‘Are we there yet? Mixed methods research in Library and Information Science.’ *Library & Information Science Research*, 30, 265 – 272.

Fitzgerald, B., 2006. ‘The transformation of open source software.’ *MIS Quarterly*, 30(3), 587-598.

Fitzgerald, B. & Bassett, G., 2003. ‘Legal issues relating to free and open source software.’ *Essays in Technology Policy and Law*, Queensland University of Technology, School of Law.

Fitzgerald, B. & Howcroft, D., 1998.‘Towards resolution of the IS research debate: from polarization to polarity.’ *Journal of Information Technology*, 13, 313 – 326.

Fowler, S., 2003. ‘Information Classification – Who, Why and How?’, SANS Institute. Available from:

www.sans.org/reading-room/whitepapers/auditing/information-classification-who-846. Accessed on 2011/05/25.

Frej, M. B. H., Bach, C., Shock, R. & Desplaines, E., 2015.’Open-source:

Adoption and challenges.’ ASEE Northeast Section Conference, *American Society for Engineering Education*.

Friedman, L. W. & Friedman, H. H., 2015. ‘Connectivity and convergence: A

whimsical history of Internet culture.‘ *SSRN Electronic Journal.*

Fung, P. & Jordan, E., 2002. ‘Implementation of information security: A knowledge-based aproach.’

Gable, G. G., 1994. ‘Integrating case study and survey research methods: an example in information systems.’ *European Journal of Information Systems*, 3(2), 112 – 126.

Gallegoa, M. D., Lunab, P. & Bueno, S., 2008. ‘User Acceptance Model of open source software.’ *Computers in Human Behaviour*, 24(5), 2199 - 2216.

Gallivan, M. J., 1997. ‘Value in triangulation: A comparison of two approaches for combining qualitative and quantitative methods.’ presented at Proceedings of the IFIP 8.2 Working Group, Philadelphia, PA, 417 – 444.

Gallopino, R., 2009. ‘Open Source TCO: Total Cost of Ownership and the Fermat’s Theorem.’ Available at:

http://robertogaloppininet/2009/01/08/open-source-TCO-total-cost-of-ownership-and-the-Fermats-theorem/(ed.).

Garfinkel, S. L., 2014. ‘Leaking sensitive information in complex document files and how to prevent it.’ *IEEE Computer and Reliability Societies*, 20 – 27.

Gartner, 2008. ‘Gartner highlights: Key predictions for IT organisations and users in 2008 and beyond’. Available from: www.gartner.com/newsroom/id/593207. Accessed on 2012/03/02.

GAO (US Government Accounting Office), 2000. ‘Federal information security: Actions needed to address widespread weaknesses.’ Retrieved on December 6, 2009, from http://www.gao.gov.

GCIS, 2007. ‘Cabinet Statement.’ Feb 22, available at:

www.gcis.gov.za/media/cabinet/2007/070222.htm. Accessed: 2 February 2012.

Geetha, S., 2012. ‘Possible challenges of developing migration projects.’

International Journal of Computers &Technology, 3(3), 463 - 465.

Gennotte, G. & Trueman, B., 1996. ‘The strategic timing of corporate disclosures.’ *Review of Financial Studies*, 9 (2), 665 – 690.

Gephart, R. P., Jr., 2004. ‘From the editors: Qualitative research and the academy of management journal.’ *Academy of Management Journal*, 47(4), 454 – 462.

Ghiani, G., Polet, J., Antila, V. & Mantyjarvi, J., 2015. ‘Evaluating context-aware user interface migration in multi-device environments.’ *Journal Ambient Intelligence Human Computing*, 6, 259 – 277.

GITOC, 2003. ‘Using open source software in the South African Government.’ A proposed strategy compiled by the Government Information Technology Officers’ Council, Version 3.3.

Goldkuhl, G., 2004. ‘Meanings of Pragmatism: Ways to conduct information systems research.’ In *Proceedings of the 2nd International Conference on Action in Language, Organisations and Information Systems (ALOIS)*, Linkoping University, Linkoping.

Goldkuhl, G., 2008. ‘What kind of pragmatism in information systems research?’ *AIS SIG Prag Inaugural Meeting*, Paris.

Goldkuhl, G., 2012. ‘Pragmatism vs. interpretivism in qualitative information systems research.’ *European Journal of Information Systems*, 21, 135 – 146.

Goles, T. & Hirschhein, R., 2000. ‘The paradigm is dead, the paradigm is dead…long live the paradigm: the legacy of Burell and Morgan.’ *Omega*, 28, 249 – 268.

Gorsuch, R. L., 1983. ‘Factor analysis (2nd ed.).’ Hillsdale, NJ: Erl-baum.

Goyal, A., Yadav, R. & Rawal, S., 2015. ‘Exploring secure Unix system with other OS.’ *International Journal of Computer Science and Mobile Computing*, 4(4), 49 – 53.

Graham, B., 2005, ‘Hackers attack via Chinese web sites: US agencies networks are among targets.’ *Washington Post*. Thursday, August 25.

Gratchoff, J. & Kroon, G., 2015. ‘Project Spartan Forensics.’ Available from: www.ipv4.os3.nl/-media/2014-2015/courses/ccf/guido_and_james.pdf. Accessed on 2015/10/05.

Gray, D. E., 2004. *Doing research in the real world*. Sage Publications, London.

Green, J. C., 2007. *Mixed methods in social inquiry*. San Francisco, CA: John Wiley & Sons.

Green, J. & Caracelli, V., 2003. ‘Making paradigmatic sense of mixed methods inquiry.’ In *Handbook of Mixed methods in Social & Behavioural Research*, Tashakkori, A. & Teddlie, C. (Eds.), Sage, California.

Guadagnoli, E. & Velicer, W. F., 1988. ‘Relation of sample size to the stability of component patterns.’ *Psychological Bulletin*, 103(2), 265 – 275.

Guba, E. G., 1990. ‘The Alternative paradigm dialog.’ In E. G. Guba (ed.), *The Paradigm Dialog.* Newbury Park, CA: Sage Publications, 17 – 30.

Guba, E. G. & Lincoln, Y. S., 1994. ‘Competing paradigms in qualitative research.’ In N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of qualitative research.* London: Sage, 105 – 117.

Guilford, J. P., 1954. *Psychometric methods*. New York: Mc Graw-Hill.

Gulati, S. & Taneja, U., 2013. ‘Specialty healthcare in India: A research design review of mixed methods approach.’ *Journal of Business Management & Social Sciences Research*, 2(10), 49 – 56.

Gupta, M., 2010. ‘A new strategy for the protection of intellectual property.’ *Computer Fraud & Security*, 8 – 10.

Gwebu, K. L. & Wang, J., 2010. ‘An exploratory study of free open source software users’ perceptions.’ *The Journal of Systems and Software*, 83(11), 2287-2296.

Gwebu, K. L. &Wang, J., 2011, ‘Adoption of open source software: The role of social identification.’ *Decision Support Systems*, 51, 220 - 229.

Haack, S., 1976. ‘The pragmatist theory of truth.’ British Journal or Philosophy of Science, 27, 231 – 249.

Hair, J. F., Anderson, R. E., Tatham, R. L. & Grablowsky, B. J., 1979. *Multivariate data analysis*. Tulsa, OK: Petroleum.

Hall, R., 2012, ‘Mixed Methods: In search of a Paradigm.’ Available at:

http://www.auamii.com/proceedings_Phuket_2012/Hall.pdf. Accessed 2014/03/15.

Hannes, K. & Lockwood, C., 2011. 'Pragmatism as the philosophical foundation for the Joanna Briggs meta-aggregattive approach to qualitative evidence synthesis.' *Journal of Advanced Nursing*, 1632 – 1642.

Hanson, B., 2008a. *Questioning qualitative inquiry*. Critical essay, London: SAGE.

Hanson, B., 2008b, 'Wither qualitative inquiry? Grounds for methodological convergence.' *Quality & Quantity*, 42, 97 – 111.

Hansen, J. G. & Jul, E., 2004. 'Self-migration of operating systems.' In *Proceedings of the 11th ACM SIGOPS European Workshop*, 126 – 130.

Hansen, M., Kohntopp, K. & Pfitzmann, A., 2002. 'The open source approach – opportunities and limitations with respect to security and privacy.' *Computers & Security*, 21(5), 461-471.

Harrison III, R. L., 2013. 'Using mixed methods designs in the Journal of Business Research, 1990 – 2010.' *Journal of Business Research*, 66, 2153 – 2162.

Harvey, C., 2015. 'Open source software list:2015 ultimate list.' Available from: www.datamation.com/open-source-software-list-ultimate-list-1.html. Accessed on 2015/10/06.

Hassan, N. K., 2007. 'Making information systems research more valuable.' *13th Americas Conference on Information Systems (AMCIS)*, Keystone, Colorado, USA, August 9 – 12, 9 – 12.

Hauge, O., Ayala, C. & Conradi, R., 2010. 'Adoption of open source software in software-intensive organisations – A systematic literature review.' *Journal of Information and Software Technology*, 52, 1133 – 1154.

Hedgebeth, D., 2007. ‘Gaining competitive advantage in a knowledge-based economy through the utilization of open source software.’ *VINE: The Journal of Information and Knowledge Management Systems*, 37(3), 284 – 294.

Hendrick, C. & Hendrick, S., 1986. ‘A theory and method of love.’ *Journal of Personality and Social Psychology*, 50, 392 – 402.

Henson, R. K., & Roberts, J. K., 2006, ‘Use of Exploratory Factor Analysis in published research: Common errors and some comment on improved practice.’ *Educational and Psychological Measurement*, 66(3), 393 – 416.

Heredero, P., De, C., Berzosa, D. L. & Santos, R. S., 2010. ‘The implementation of free software in firms, an empirical analysis.’ *The International Journal of Digital Accounting research*, 10(6).

Hevner, A. R., March, S. T., Park, J. & Ram, S., 2004. ‘Design Science in information systems research.’ *MIS Quarterly*, 28(1), 75 – 115.

Hislop, R., 2004. ‘Mossel Bay adopts Linux on the desktop’, *Electronic Government Africa*, 1(1), 14.

Hoepman, J. & Jacobs, B., 2007. ‘Increased security through open source.’ *Communications of the ACM*, 50(1).

Hogan, T. P., Benjamin, A. & Brezinksi, K. L., 2000. ‘Reliability methods: A note on the frequency of use of various types.’ *Educational and Psychological Measurement*, 60(4), 523 – 531.

Howe, K. R., 1988. ‘Against the quantitative-qualitative incompatibility thesis or dogmas die hard.’ *Educational Researcher*, 17(1), 10 – 16.

Hida, M. & Hisham, M, 2009.’Migration to FOSS: Readiness and challenges.’ Free Open Source Software (FOSS) Wokrshop, Sudan.

Hudson, J. R., 2015. ‘History of FOSS.’ Available from: www.johnrhudson.me.uk/computing/FOSS_in_academia.pdf. Accessed on

2015/10/02.

Humphreys, E. 2008. 'Information security management standards: Compliance, governance and risk management.' *Information Security Technical Report*, 247 – 255.

Hussain, K., Addulla, N., Rajan, S. & Moussa, G., 2005. 'Preventing the capture of sensitive information.' *43rd ACM Southeast Conference*, March 18 – 20, Kennesaw, GA, USA.

Hutcheson, G. & Sofronion, N., 1999. *The multivariate social scientist: introducing statistics using generalised linear models*. Thousand Oaks: Sage Publications.

Huth, C. L., Chadwick, D. W., Claycomb, W. R. & You, I., 2013. 'Guest editorial: A brief overview of data leakage and insider threats., *Information Systems Front*, 15, 1 – 4.

Hutton, E., 2009. 'An examination of post-positivism and postmodernism.' Available from: http://ericahutton.blogspot.co.za/2009/03/examination-of-postpositivism.html. Accessed on: 2015/11/21.

Iroju, O. & Ikono, R., 2013. 'A security based framework for interoperability of healthcare systems.' *International Journal of Applied Information Systems (IJAIS)*, 23 – 31.

ISO17799 News – Issue 9, 2007. 'Establishing information classification criteria.' The ISO17799 Newsletter – News & views on the ISO/IEC security standard.

IT Web, 2007. 'Two nabbed for eNaTIS fraud, IT Web News.' Available online at: http://www.itweb.co.za, accessed on 22/09/2008.

Ivankova, V., Creswell, J. W. & Stick, S. L., 2006. 'Using mixed-methods sequential explanatory design: From theory to practice.' *Field methods*, 18(1), 3 – 20.

James, S. & Van Belle, J., 2008. 'Ensuring the long-term success of OSS migration: a South African exploratory study.' 6th Conference on Information Science Technology and Management, New Delhi, India.

Jarvensivu, T. & Tornroos, J., 2010. 'Case study research with moderate constructionism: Conceptualisation and practical illustration.' *Industrial Marketing Management*, 39, 100 – 108.

Jericho Forum, 2009.'Information Classification.' Available at: www.opengroup.org/jericho/publications.htm. Accessed on 2013/04/15.

Jick, T. D., 1979. 'Mixing qualitative and quantitative methods: triangulation in action. *Administrative Science Quarterly*, 24, 602 – 611.

Joas, H., 1993. *Pragmatism and Social Theory*. University of Chicago Press, Chicago, IL.

Johansson, R., 2003. 'Case Study Methodology.' International Conference on Methodologies in Housing Research. *International Association of People-Environment Studies,* Stockholm, 22 – 24.

Johari, J., 2009. 'Interpretivism in Information Systems (IS) Research.' *Integration and Dissemination Research Bulletin*, 4, 25 – 27.

Johnson, R. B. & Onwuegbuzie, A. J., 2004. 'Mixed methods research: A research paradigm whose time has come.' *Educational Researcher*, 3(7), 14 – 26.

Johnson, R. B., Onwuegbuzie, A. J. & Turner, L. A., 2007. 'Toward a definition of mixed methods research.' *Journal of Mixed Methods Research*, 1(2), 112 – 133.

Johnston, K. A. & Seymour L. F., 2005. 'Why South Africans don't floss?' Proceedings of the International Business Information Management *Conference (IBIMA)*, 438 – 446, July, Lisbon, Portugal.

Jones, A. K., 2002. 'Network Security.' Paper presented at the Critical

Infrastructure and Information Assurance Symposium, Syracuse, NY.

Jones, A. & Colwill, C., 2008. 'Dealing with the malicious insider.' In: 9th *Australian information and Warfare security Conference*.

Juneja, G. K., 2013. 'Ethical hacking: A technique to enhance information security.' *International Journal of Innovative Research in Science, Engineering and Technology*, 2(12), 7575 – 7580.

Jurisch, M. C.,Wolf, P. & Krcmar, H., 2013. 'Using the Case survey method for synthesizing Case Study evidence in information systems research.' *Proceedings of the 19th Americas Conference on Information Systems,* Chicago, Illinois.

Kalyvas, J. R., Overly, M. R. & Karlyn, M. A., 2013. 'Cloud computing: A practical framework for managing cloud computing risk – Part II.' *Intellectual Property & Technology Law Journal*, 255(4), 19 – 27.

Kamthan, P., 2007. 'A perspective on software engineering education with open source', IGI Global, In K. St. Amant & B. Still (eds.), *Handbook of research on open source software: Technological, economic and social perspectives,* 690 – 702.

Kapitan, T., 1997.'Peirce and the structure of abductive inference.' In studies in the Logic of Charles Sanders Peirce, ed. Nathan Houser, Don D. Roberts, and James Van Evra. Bloomington: Indiana University Press.

Kaplan, B. & Duchon, D., 1988. 'Combining qualitative and quantitative methods in information systems research: A Case Study.' *MIS Quarterly*, 12, 570 – 586.

Kale, A. V., Dubey, P. S. & Bajpayee, V., 2015. 'A review on data leakage prevention.' *International Journal of Computer Science and Mobile Computing*, 4(4), 513 – 518.

Kaushal A., Khan, A. & Kumar, V., 2015. 'Big data: A brief investigation on different privacy issues.' *International Journal of Innovations & Advancement in Computer Science (IJIACS)*, 4(1), 122 – 129.

Kavanagh, J., 2006. 'Security special report: the internal threat.' *Computer Weekly*, www.computerweekly.com/Articles/2006/04/25/215621/security-special-report-the-internal-threat.htm

Kazimir, P., 2012. 'IT migration – a way to business sustainability.' *American International Journal of Research*, 2(4), 101 – 110.

Kelemen, M. & Rumens, N., 2012. 'Pragmatism and heterodoxy in organisation research: Going beyond the quantitative/qualitative divide.' *International Journal of Organisational Analysis*, 20(1), 5 – 12.

Kelle, U., 2005. 'Sociological explanations between micro and macro and the integration of qualitative and quantitative methods.' *Historical Social Research*, 30(1), 95 – 117.

Kemp, R., 2009. 'Current developments in open source software.' *Computer Law & Security Review*, 25, 569-582.

Kerlinger, F. N. & Lee, H. B., 2000. 'Foundations of behavioural research' (4th edn.) Holt New York: Harcourt College Publishers.

Khanjani, A. & Sulaiman, R., 2011. 'The aspects of choosing open source versus closed source.' *IEEE Symposium on Computers & Informatics*, 646 - 649.

Khazanchi, D. & Munkvold, B. E., 2002. 'On the rhetoric and relevance of IS research paradigms: a conceptual framework and some propositions.' *System Sciences*, Proceedings of the 36th Annual Hawaii International Conference on IEEE.

Kirda, E & Kruegel, C., 2005. 'Protecting users against phishing attacks with antiPhish.' *Proceedings of the 29th Annual Conference Computer Software and Applications Conference, IEEE.*

Klein, H. K. & Meyers, M. D., 1999. ‘A set of principles for conducting and evaluating interpretive studies in information systems.’ *MIS Quarterly*, 23(1), 67 – 93.

Klingner, J. K. & Boardman, A. G., 2011. ‘Addressing the “Research Gap” in special education through mixed methods.’ *Learning Disability Quarterly*, 34, 208.

Knapp, K. J., Marshall, T. E., Rainer, R. K. & Ford, N., 2007. ‘Information security: management’s effect on culture and policy.’ *Information Management & Computer Security*, 14(1), 24 – 36.

Kolkowska, E., 2011. ‘Security subcultures in an organisation - exploring value conflicts.’ *ECIS conference,* Helsinki, Finland 9-11 June 2011. *The 19th European Conference on Information Systems ITC and Sustainable Services Development.*

Kovacs, G. L., Drozdik, S., Zuliani, P. & Succi, G., 2004. ‘Open source software for the public administration.’ *Proceedings of the 6th International Workshop on Computer Science and Information Technologies*, Budapest, Hungary.

Kovacs, G. & Spens, K. M., 2005. ‘Abductive reasoning in logistics research.’ *International Journal of Physical Distribution & Logistics Management*, 35(2), 132 – 144.

Ku, W. & Chi, C. H., 2004. ‘Survey on the technological aspects of digital rights management.’ *Proceedings of the 7th International Conference, ISC,* Palo Alto, CA, USA, Sept 27 – 29, Springer Berlin / Heidelberg, 391 – 403.

Kumar, R., 2005. ‘Research Methodology – A step-by-step guide for beginners.’ Second edition, Pearson Education.

Kurita, H., Shioya, R., Irie, H., Goshima, M. & Sakai, S., 2007. ‘Dynamic information flow control for preventing information leakage.’

Proceedings of the IPSJ SIG Technical Report, HOKKE, Hokkaido, Japan, March 1 – 2, ARC- 172, IPSJ Press, 227 – 232.

Kvasny, L., & Richardson, H., 2006. 'Critical research in information systems: Looking forward, Looking back.' *Information Technology & People*, 19(3), 196 – 202.

Lacey, D., 2010. 'Understanding and transforming organisational security culture.' *Information Management & Computer Security*, 18(1), 4-13.

Lafuente, G., 2015. 'The big data security challenge.' *Network Security*, 12 – 14.

Lee, A., 1989. 'Integrating positivist and interpretive approaches to organisational research.' *Organisation Science*, 2(4), 342 – 365.

Lee, A. S., 1999. 'Rigour and relevance in MIS research: Beyond the approach of positivism alone.' *MIS Quarterly*, 23(1), 29 – 34.

Lee, A. & Nickerson, J., 2010. 'Theory as a case of design: lessons for design from philosophy of science.' *Proceedings of the 43rd Hawaii International Conference on System Sciences*.

Lee, A. S. & Hubona, G. S., 2009. 'A scientific basis for rigour in information systems research.' *MIS Quarterly* 33(2), 237 – 262.

Lee, S., Noh, S. & Kim, H., 2013.'A mixed methods approach to electronic word-of-mouth in the open-market context.' *International Journal of Information Management*, 33, 687 – 696.

Leedy, P. & Ormrod, J., 2012. *Practical research: Planning and design* (10th ed.) Upper Saddle River, NJ: Merrill Prentice Hall, Thousand Oaks: SAGE Publications.

Leonardi, P. M. & Barley, S. R., 2008. 'Materiality and change: Challenges to building better theory about technology and organizing.' *Information and Organisation*, 18, 159 – 176.

Lewis, J. A., 2007. ‘Government open source policies.’ Available from: http://www.csis.org/media/csis/pubs/070820_open_source_policies.pdf, Accessed on 2012/01/29.

Liddy, E. D., 2001. ‘Information security and sharing.’ *Online*, 25(3), 28 – 30.

Lincoln, Y., & Guba, E. G., 1985. *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lincoln, Y., & Guba, E. G., 1989. *Fourth Generation Evaluation*. Beverly Hills, CA: Sage.

Lindeman, R. H., Merenda, P. F. & Gold, R. Z., 1980. *Introduction to bivariate and multivariate analysis*. Glenview, IL: Scott, Foresman.

Linn, R. L. & Grondlund, N. E., 2000. *Measurement and Assessment in Teaching.* (8th ed.), New Jersey: Merrill (an imprint of Prentice Hall).

Little, G. & Stergiades, E., 2009. ‘Worldwide open source services 2009-2013 forecast.’ International Data Corporation.

Long, T. & Johnson, M., 2000. ‘Rigour, reliability and validity in qualitative research.’ *Clinical Effectiveness in Nursing*, 4, 30 – 37.

Loo, R., 1983. ‘Caveat on sample sizes in factor analysis.’ *Perceptual and Motor Skills*, 56, 371 – 374.

Lu, L., Arpaci-Dusseau, A. C., Arpaci-Dusseau, R. H. & Lu, S., 2014. ‘A study of Linux file system evolution.’ *ACM Transactions on Storage (TOS)*, 10(1), 3.

Luyt, R., 2012. ‘A framework for mixing methods in quantitative measurement development, validation and revision: A case study.’ *Journal of Mixed Methods Research*, 6(4), 294 – 316.

Manfield-Devine, S., 2008. ‘Open Source: does transparency lead to security?’ *Computer Fraud & Security*, 11-13.

Ma, Q. & Pearson, J. M., 2005a. 'ISO 17799: Best practices in information security management.' *Communications of the Association for Information Systems*, 15, 577 – 591.

Ma, Q. & Pearson, J. M. 2005b. 'The Inter-relationship between objectives and practices in information security management.' *Proceedings of the 11th Americas Conference on Information Systems*, Omaha, NE, USA, August 11 – 14.

Ma, Q., Schmidt, M. & Pearson, J., 2009. 'An integrated framework for information security management.' *Review of Business*, 30(1), 58 – 69.

Malterud, K., 2001. 'Qualitative research: standards, challenges, and guidelines.' *The Lancet*, 358, 483 – 488.

Manova Report, 2010. 'Multivariate Analysis.' available from: www.slideshare.net/Lordnikhil/manova-report-3593954. Accessed on 2014/02/12.

Manfield-Devine, S., 2008. 'Open Source: does transparency lead to security?' *Computer Fraud & Security*, 11-13.

Manthena, M., 2011. 'Adoption of open source software: The challenges and opportunities.' Available from: http://bit.ly/EFY_Times. Accessed on 2015/08/20

Maples, S. & Chen, W., 2015. 'An examination of recent network security failures.' *International Conference on Security and Management*, 266 – 271.

Markus, M. L., 1994. 'Electronic mail as the medium of managerial choice.' *Organisation Science*, 5, 342 – 365.

Markus, M. L., Manville, B. & Agres, C. E., 2014. 'What makes a virtual organisation work : Lessons from the open-source world.' Image. Available from: www.sloanreview.mit.edu/article/what-makes-a-virtual-organisation-work-lessons-from-the-opensource-world/. Accessed on 2015/10/05.

Marsan, J., Pare, G. & Wybo, M. D., 2012. 'Has open source software been institutionalised in organisations or not?' *Information and Software Technology*, 54, 1308 – 1316.

Marshall, P., Kelder, J. A. & Perry, A., 2005. 'Social constructivism with a twist of pragmatism: a suitable cocktail for information systems research.' 16th Australasian Conference on Information Systems, Sydney.

Martinez, V. G., Encinas, L. H., Dios, A. Q., Encinas, A. H. & Vaquero, J. M., 2013. 'Avoiding sensitive information leakage in moodle.' *Literacy Information and Computer Education Journal (LICEJ)*, 2(2), 1331 – 1341.

Marton, F., 1981. 'Phenomenography – describing conceptions of the world around us.' *Instructional Science*, 10, 177 – 200.

Mastin, L., 2008, 'The basics of philosophy.' from: www.philosophybasics.com/branch_solipsism.html. Accessed on 2014/03/24.

Maxcy, S. J., 2003. 'Pragmatic threads in mixed methods research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism.' In: A. Tashakkori, & C. Teddlie (eds.) *Handbook of mixed methods in social and behavior research*. Thousand Oaks, CA: Sage, 51 – 90.

McCormick, M., 2008. 'Data theft: a prototypical insider threat.' In: S. Stolfo, S. Bellovin, S. Hershkop, A. Keromytis, S. Sinclair & S. Smith (eds.) Inside attack and cyber security: beyond the hacker. New York: Springer, 52 – 67.

McCue, A., 2008, 'Beware the insider security threat.' CIO Jury, www.silicon.com/management/cio-insights/2008/04/17/beware-the-insider-security-threat-39188671/. Accessed on 2012/04/08.

McCullagh, K., 2007. 'Data sensitivity: resolving the conundrum.' *British & Irish Law, Education and Technology Association Annual Conference*. Hertfordshire 16 - 17 April.

McHugh, M. L., 2008. 'Standard error: meaning and interpretation, Lessons in biostatistics.' *Biochemia Medica*, 1, 7 – 13.

McLeod, S. A., 2013. 'What is validity.' Retrieved from: www.simplypsychology.org/validity.html. Accessed on 2015/08/30.

Metcalfe, R., 2012. 'Toptips for selecting open source software.' Open Source Software Advisory Service. Available at: www.oss-watch.ac.uk/resource/tips.xml

Mi2g Report, 2002, 'Press report' at: http://www.zdnet.com.au/newstech/os/story/0,2000024997,20266696,00.htm. Accessed on 2012/04/02.

Michell, J., 2011. 'Qualitative research meets the ghost of Pythagoras.' *Theory & Psychology*, 21, 241 – 259.

Miles, M. B. & Huberman, A. M., 1994. *Qualitative data analysis*. 2nd edn., Sage Publications, Thousand Oaks, CA.

Mingers, J., 2001. 'Combining IS research methods: Towards a pluralist Methodology.' *Information Systems Research*, 12(3), 240 – 259.

Mingers, J., 2003. 'The paucity of multi-method research: A review of the information systems journal.' *Information Systems Journal*, 13(3), 233 – 249.

Miscione, G. & Johnston, K., 2010. 'Free and open source software in developing contexts, from open in principle to open in the consequences.' *Journal of Information & Ethics in Society*, 8(1), 42-56.

Moody, G., 2015.'Open source has won: But it isn't finished.' ComputerWorld UK.

Morgan, D. L., 2007. 'Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods.' *Journal of Mixed Methods Research*, 1(1), 48 – 76.

Morgan, D. L., 2014. *Integrating qualitative and quantitative methods, A pragmatic approach*. Thousand Oaks, CA: Sage.

Morris, J., McNaughton, D., Mullins, R. & Osmond, J., 2009. 'Post-positivist epistemology.' Available from: http://admn502awiki.pbworks.com/f/Post%2Bpositivist.doc. Accessed on 2015/11/24.

Morse, J. M., 1994. 'Designing funded qualitative research.' In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 220 – 235). Thousand Oaks, CA: Sage.

Morse, J. M., Niehaus, L., Wolfe, R. R. & Wilkins, S., 2006. 'The role of the theoretical drive in maintaining validity in mixed-methods research.' *Qualitative Research in Psychology*, 3(4), 279 – 291.

Mtsweni, J. & Biermann, E., 2008. 'An investigation into the implementation of open source software within the SA government: An emerging expansion model.' *SAICSIT*, 6 – 8 October, Wilderness Beach Hotel, Wilderness, SA.

Mtsweni, J. & Biermann, E., 2010. 'A roadmap to proliferate open source software usage within SA Government servers.' *Third International Conference on Broadband Communications, Information Technology & Biomedical Applications, IEEE Computer Society*, 430 - 436.

Mutula, S. & Kalaote, T., 2010. 'Open source software deployment in the public sector: a review of Botswana and South Africa.' *Emerald*, 28(1), 63 – 80.

Myers, M. D., 1997. ‘Qualitative research in information systems.’ *MIS Quarterly*, 21(2), 241 – 242.

Myers, M. D., 2005. ‘Qualitative research in information systems.’ *Association of Information Systems*. Available on: www.qual.auckland.ac.nz. Accessed on 2014/03/23.

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

Myers, M. D. & Newman, M., 2006. ‘The qualitative interview in IS research: Examining the craft.’ *Information and Organisatio*n, 2 – 26.

Nagler, M., 2005. ‘Open Source adoption of the German Federal Office for information security.’ Available from: http://ec.europa.eu/idabc/servelets/Doc?id=21394. Accessed on 2012/01/25.

Naik, D. P. & Ghule, A., 2013. ‘An advanced data transformation algorithm for categorical data protection.‘ *International Journal of Computer Science and Information Technologies*, 4(6), 899 – 902.

Nawafleh, S. A., Hasan, M. Y. F. & Nawafleh, Y., 2013. ‘Protection and defense against sensitive data leakage problem within organisations.’ *European Journal of Business and Management*, 5(23).

Nazareth, D. L. & Choi, J., 2015. ‘A system dynamics model for information security management.’ *Information & Management*, 52, 123 – 134.

Nazeer, S., Bahadur, F., Iqbal, A., Ashraf, G. & Hussain, S., 2015. ‘A comparison of Windows 8 and Linux operating system (Android) security for mobile applications.’ *International Journal of Computer (IJC)*, 21 – 29.

Neuhaus, S., Zimmermann, T., Holler, C. & Zeller, A., 2007. ‘Predicting vulnerable software components.’ In *Proceedings of the 14th ACM*

Conference on Computer and Communications Security, Alexandria VA, Oct., 529 – 540.

Nguyen, A. T., Nguyen, H. A., Nguyen, T. T. & Nguyen, T. N., 2014. 'Statistical learning of API mappings for language migration.' *ICSE Companion*, May 31 – June 7, Hyderabad, India, 618 – 619.

Nielsen, J., 1993. *Usability Engineering*. Academic Press, Boston.

Niglas, K., 2004. 'The combined use of qualitative and quantitative methods in educational research.' Tallinn Pedagogical University Press, Tallinn.

NIH, 2008. 'Guide for identifying sensitive information.' Available on http://irm.cit.nih.gov/security/NIH_sensitive_info_Guide.doc. Accessed on: 201/07/13

NIST, 2008. 'An Introduction to Computer Security, The NIST Handbook.' Text from: http://csrc.nist.gov/publications/nistpubs/800-12/800-12-html/chapter1-printable.html. Accessed on: 2012/09/03.

Noor, K. B. M., 2008. 'Case study research methodology.' *American Journal of Applied Sciences*, 5(11), 1602 – 1604.

Nosworthy, J., 2000. 'Implementing information security in the 21st Century – do you have the balancing factors?' *Computers & Security*, 19(4), 337 – 347.

Novell Connection, 2009. 'About the National Library of South Africa.' from: https://www.novell.com/connectionmagazine/2009/04/national-library-of-south-africa.html. Accessed on: 2013/09/12

NSA (National Security Agency), 2001. 'Response #19' at: www.nsa.gov/selinux/faq.html

Nunnally, J. C. & Bernstein, I. H., 1994. *Psychometric Theory*. New York: McGraw-Hill.

Nurse, J. R. C., Erola, A., Goldsmith, M. & Creese, S., 2015. 'Investigating the leakage of sensitive personal and organisational information in email headers.' *Journal of Internet Services and Information Security (JISIS)*, 5(1), 70 – 84.

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

Olivier, M. S., 2002. 'Database privacy – Balancing confidentiality, integrity and availability.' *SIGKDD Explorations*, 4(2), 20.

O'Neill, A., 2012. 'Open source software.' University of Central Lancashire, Preston, UK, Elsevier Inc.

Onwuegbuzie, A. J. & Collins, K. M. T., 2007. 'A typolology of mixed methods sampling designs in social science research.' *The Qualitative Report*, 12(2), 281 – 316.

Onwuegbuzie, A. J., Johnson, R. B. & Collins, K. M. T., 2009. ' A call for mixed analysis: A philosophical framework for combining qualitative and quantitative.' *International Journal of Multiple Research Approaches*, 3, 114 – 139.

Onwuegbuzie, A. J. & Leech, N. L., 2005.'On becoming a pragmatic researcher: the imporatance of combining quantitative and qualitative methodologies.' *International Journal of Social Research Methodology*, 8(5), 375 – 387.

Open Source Report, 2008. 'Coverity report.' viewed 1 February 2013, from: http://scan.coverity.com/report/

Open Source Security Study, 2008. 'Fortify report', viewed 1 February 2013, from: www.fortify.com/ossreport.html

Oram, A., 2011. 'Promoting open source software in Government: The challenges of motivation and follow-through.' *Journal of Information Technology & Politics*, 8(3), 240 – 252.

Orlikowski, W. J. & Baroudi, J. J., 1991. 'Studying information technology in organisations: Research approaches and assumptions.' *Information Systems Research*, 2(1), 1 – 8.

Ormerod, R. J., 1996. 'Combining management consultancy and research.' *Omega*, 24(1), 1 – 12.

Ostlund, U., Kidd, L., Wengstrom, Y. & Rowar-Dewar, N., 2011. 'Combining qualitative and quantitative research within mixed methods research designs: A methodological review.' *International Journal of Nursing Studies*, 48, 369 – 383.

OSI (Open Source Initiative), 2014. 'Open source definition', www.opensource.org/osd.

Otter, A., 2007. 'SA government gets serious about ODF, IOIL Technology.' Available from: http://www.ioltechnology.co.za/article_page.php?iArticleId=4126700&iSectionId=2888, accessed on 2012/02/21.

Oyelami, J. O. & Ithnin, N. B., 2015. 'Establishing a sustainable information security management policies in organisation: A guide to information security management practice (ISMP).' *International Journal of Computer and Information Technology*, 4(1), 44 – 49.

Pan, S. L. & Tan, B., 2011. 'Demystifying case research: A structured-pragmatic-situational (SPS) approach to conducting case studies.' *Information and Organisation*, 21, 161 – 176.

Park, Y., Gates, S. C., Teiken, W. & Cheng, P., 2011. 'An experimental study on the measurement of data sensitivity.' *Badgers*, April 10, Salzburg, Austria, 70 – 77.

Pather, S. & Remenyi, D., 2004. 'Some of the philosophical issues underpinning research: From positivism to critical realism.' *Proceedings of the SAICSIT*, 141 – 146.

Patton, M. Q., 1988. 'Paradigms and pragmatism.' In D. M. Fetterman (ed.), *Qualitative Approaches to Evaluation in Education: The Silent Scientific Revolution*. New York: Praeger.

Patton, M. Q., 1990. *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.

Patton, M. Q., 2002. *Qualitative evaluation and research methods*. 3rd ed. Sage: Thousand Oaks, CA.

Pearson, H. E., 2000. 'Open source licenses, Open source – The death of closed source systems?' *Computer Law & Security Report*, 16(3), 151-156.

Pearson, S., 2009. 'Taking account of privacy when designing cloud computing services.' CLOUD '09, IEEE, May 23, Vancouver, Canada.

Pearson, S., Casassa-Mont, M. & Novoa, M., 2007. 'Securing information transfer in distributed computing environments.' *IEEE Computer Society*, 34 – 42.

Peirce, C. S., 1931-58. ' Collected papers of Charles Sanders Peirce.' vols. 1 – 6, ed. C. Hartshone and P. Weiss, vols. 7 – 8, ed. A. W. Burks, Cambridge, MA: Havard University Press.

Peirce, C. S., 1955. Philosophical Writings of Peirce. Selected and edited by J. Buchler, New York: Dover Publications, Inc.

Peirce, C. S., 1992. *The essential Peirce*. Selected philosophical writings, 1, Indiana University Press, Indianapolis, IN.

Peng, G. C., Nunes, J. M. B. & Annansingh, F., 2011. 'Investigating information systems with mixed methods research.' In: *Proceedings of the IADIS*

International Workshop on Information Systems Research Trends, Approaches and Methodologies (ISRTAM), Rome, Italy.

Performanta, L. A., 2015. 'Data loss prevention: the business case.' *Computer Fraud & Security*, 13 – 16.

Petter, S. C. & Gallivan, M. J., 2004. 'Toward a framework for classifying and guiding mixed method research in Information Systems.' *IEEE, Proceedings of the 37th Hawaii International Conference on System Sciences*.

Petty, N. J., Thomson, O. P. & Stew, G., 2012. 'Ready for a paradigm shift? Part 2: Introducing qualitative research methodologies and methods.' *Manual Therapy*, 17, 378 – 384.

Phillips, D. C. & Burbules, N. C., 2000. *Post-positivism and educational research*. New York: Rownan & Littlefield.

Pieta, S., 2010. 'IT systems security management in migration process.' *Foundations of Management*, 2(2), 63 – 80.

Plutynski, A., 2011. 'Four problems of abduction: A brief history.' *Journal of the International Society for the History of Philosophy of Science*, 1, 1 – 22.

PNC, 2007. 'Proposal to provide an open source software support, Change management and training services to Presidential National Commission' – by Impi Linux (Pty) Ltd.

Pokarna, K., Ingole, S. G. & Diwate, R. B., 2015. 'Comparative study for software selection of OSS vs CSS.' *International Journal for Research in Emerging Science and Technology*, 2(1), 376 – 379.

Polivec, 2002. *Security policy development process*. Colorado Springs, CO: PoliVec Inc.

Poulter, A., 2010. 'Open source in libraries: an introduction and overview.'

Emerald, 59(9), 655 – 661.

Pozzebon, M., 2003. 'Criteria for conducting and evaluating critical interpretive research in the IS field.' *Cahier du GreSI*, 13(14).

Putnam, H., 1987. *The many faces of Realism*. Open court press: La Salle.

PWC, 2008. 'Security breaches survey, enterprise and regulatory reform (BERR).' PricewaterhouseCoopers on behalf of the UK Department of Business, www.pwc.co.uk.

Raghunathan, S., Prasad, A., Mishra, B. K. & Chang, H., 2005. 'Open source versus closed source: Software quality in monopoly and competitive markets.' *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans*, 35(6), 903-918.

Rakers, J., 2010. 'Managing professional and personal sensitive information.' *SIGUCCS*, October 24 – 27, Norfolk, Virginia, USA.

Rafiq, M. & Ameen, K., 2009. 'Issues and lessons learned in open source software adoption in Pakistani libraries.' *The Electronic Library*, 2794, 601 – 610.

Rasmussen, G. T., 2008. 'Safeguarding sensitive information – An ounce of prevention.' Available from: www.gideonrasmussen.com/article-09.html. Accessed on 2011/08/21.

Ratajczak, D., 2015. 'Is Linux a better desktop operating system than Microsoft Windows?' Available from: www.grin.com/en/e-book/292829/is-linux-a-beter-desktop-operating-system-than-microsoft-windows.

Razavian, M. & Lago, P., 2014. 'A lean and mean strategy: a data migration industrial study.' *Journal of Software: Evolution and Process*, 26, 141 - 171

Renken, J. & Moswetsi, W., 2006. 'Experiences in interpretive information systems research: Investigating E-commerce adoption in the Botswana Defence Force in quality and impact of qualitative research.' In: Ruth, A. (ed.) Quality and Impact of Qualitative Research, 3rd Annual QualIT Conference, Brisbane: Institute for Integrated and Intelligent Systems, Griffith University, 104 – 115.

Richardson, B. T. & Michalski, J., 2007. 'Security framework for control, system data classification and protection.' SANDIA report. Available from: http://energy.sandia.gov/wp/wp-content/gallary/uploads/Richardson-2007-3888P.pdf. Accessed on: 2011/09/08.

Richardson, R. & Kramer, E. H., 2006. 'Abduction as the type of inference that characterises the development of a grounded theory.' *Journal of Qualitative Research*, 6(4), 497 – 513.

Ridenour, C. S. & Newman, I., 2008. *Mixed method research: Exploring the interactive continuum*. Carbondale, IL: Southern Illinois University Press.

Ritter, N. L., 2010. 'Understanding a widely misunderstood statistic: Cronbach's alpha.' Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans.

Rob, M. A., 2015. 'Software size estimation: Practical models and their applications in various phases of the SDLC.' *International Journal of Research in Business and Technology*, 6(3), 856 – 862.

Robey, D., 1996. 'Diversity in information systems research: Threat, promise, and responsibility.' *Information Systems Research*, 7(4), 400 – 408.

Robson, C., 2011. 'Real world research.' 3rd edition, Chichester: Wiley.

Rocco, T. S., Bliss, L. A., Gallagher, S. & Perez-Prado, A., 2003. 'Taking the next step: mixed methods research in organisational systems.' *Information Technology, Learning, and Performance Journal*, 21(1), 19 – 29.

Rodgers, C., 2012. 'Data Classification: Why is it important for Information Security?' *Secure State*, Available: http://blog.securestate.com/data-classification-why-is-it-important-for-information-security. Accessed on: 28-August-2015.

Rohm, A. J. & Milne, G. R., 2004. ‘Just what the doctor ordered – The role of information sensitivity and trust in reducing medical information privacy concern.’ *Journal of Business Research*, 57, 1000 – 1011.

Roode, D., 2003. ‘Information Systems research: a matter of choice?: editorial.’ *South African Computer Journal,* 30(1).

Rorty, R., 1991. *Objectivity, relativism and truth*. Philosophical papers, Cambridge, UK: Cambridge University Press.

Ross, S. J., 2008. ‘Enforcing information security: architecture and Responsibilities.’ *Network Security,* 7 – 10.

Rowlands, B., 2003. ‘Employing interpretive research to build theory of information systems practice.’ *AJIS,* 10(2), 3 – 22.

Rowley, J., 2012. ‘Conducting research interviews.’ *Management Research Review,* 35, Issue 3/4, 260 – 271.

Ruighaver, A. B., Maynard, S. B. & Chang, S., 2007. ‘Organisational security culture: Extending the end-user perspective., *Computers & Security,* 26, 56 – 62.

Runeson, P. & Host, M., 2009. ‘Guidelines for conducting and reporting case study research in software engineering.’ *Empirical Software Engineering*, 14, 131 – 164.

Ryan, A. B., 2006.’Post-positivist approaches to research. In: Researching and writing your thesis: A guide for postgraduate students.’ MACE: Maynooth Adults and Community Education, pp. 12 – 26.

Sale, J. E. M., Lohfeld, L. H. & Brazil, K., 2002. ‘Revisiting the quantitative-qualitative debate: Implications for mixed-methods research.’ *Quality & Quantity*, 36, 43 – 53.

Sampemane, G., 2015. ‘Internal access controls.’ *Communications of the ACM*, 58(1), 62 – 65.

Sarrab, M. & Bourdoucen, H., 2013. ‘Runtime monitoring using policy based approach to control information flow for mobile Apps.’ *International Journal of Electrical, Electronic, Electrical Science and Engineering*, 7(11), 526 – 533.

Sarrab, M., Elbasir, M. & Elgamel, L., 2013. ‘The Technical, Non-Technical Issues and the Challenges of Migration to Free and Open Source Software.’ *International Journal of Computer Science Issues*, 10(3), 464 – 469.

SAS, 2014. ‘JMP 11 Online documentation.’ Available from: www.jmp.com/academic/Learning_Library.shtml. Accessed on: 2014/04/21

Sasse, M. A., Ashenden, D., Lawrence, D., Coles-Kemp, L., Flechais, I. & Kearney, P., 2007. ‘Human vulnerabilities in security systems, human factors working group.’ Cyber security KTN human factors white paper.

Saunders, M., Lewis, P. & Thornhill, A., 2003. *Research methods for business students*. 3rd edn., Pearson Education, Essex.

Sawyer, S., 2000. ‘Studying organisational computing infrastructures: Multi-method approaches.’ presented at IFIP 8.2 Working Conference, Aalborg, Denmark, 213 – 231.

Sawyer, S., 2001. ‘Analysis by long walk: Some approaches to the synthesis of multiple sources of evidence.’ In: *Qualitative research in IS: Issues and trends*, E. Trauth, ed. Hershey, PA: Idea Group Publishing.

Schneier, B., 2000. *Secrets & lies, Digital security in a networked world*. New York, John Wiley.

Scholz, C., 1990. ‘The symbolic value of computerized information systems’, In P. Gagliardi (ed.) *Secrecy*, 161 – 177, New York: Human Sciences.

Schryen, G., 2009. ‘Security of open source and closed source software: An empirical comparison of published vulnerabilities.’ *AMCIS Proceedings*, Paper 387.

Schryen G., 2011. ‘Is open source security a myth? What does vulnerability and patch data say?’ *Communications of the ACM*, May, 54(5).

Schulenberg, J. L., 2007. ‘Analysing police decision-making: Assessing the application of a mixed-method/mixed-model research design.’ *International Journal of Social Research Methodology*, 10(2), 99 – 119.

Schulze, S., 2003. ‘Views on the combination of quantitative and qualitative research approaches.’ *Progressio*, 25(2), 8 – 20.

Schwandt, T. A., 2001. *Qualitative inquiry, Theory, method and practice*. London: Sage.

Scola, N., 2009. ‘Why the White House’s embrace of Drupal matters.’ *Personal Democracy Forum techPresident*.

SERPRO, 2005. ‘Fast move to free software in Brazil.’ available from: http://ec.europa.eu/idabc/en/documents/5131/528, accessed on 2012/01/20.

Shaikh, M. & Cornford, T., 2012. ‘Strategic drivers of open source software adoption in the public sector: Challenges and opportunities.’ European Conference on Information Systems AIS, Paper 237.

Sharma, A. & Adkins, R., 2006. ‘OSS in India’, in DiBona, C., Cooper, D. and Stone, M. (eds.), *Open Sources 2.0*. O’Reilly Media, Sebastopol, CA, 189 - 196.

Shaw, A., 2011, ‘Insurgent expertise: The politics of free/live and open source software in Brazil.’ *Journal of Information Technology & Politics*, 8, 253 – 272.

Shelly, G. B. & Rosenblatt, H. J., 2013. *Systems analysis & design*. Course

Technology.

Shostack, A., 2008. 'Experiences threat modelling at Microsoft,' In Modelling Security Workshop, Dept. of Computing, Lancaster University, UK, viewed 2 February 2013, from: http://blogs.msdn.com/sdl/attachment/8991806.ashx.

Shostack, A., 2014. *Threat modelling: Designing for security*. John Wiley & Sons Inc.

Sidi, F., Selamat, M. H., Ghani, A. A. A. & Ibrahim, H., 2009. 'An investigation into methods and concepts of qualitative research in information system research.' *Computer and Information Science*, 2(4), 47 – 54.

Sidorova, A., Evangelopoulos, N. Valacich, J. S. & Ramakrishnan, T., 2008. 'Uncovering the intellectual core of the Information Systems discipline.' *MIS Quarterly*, 32(3), 467 – 482.

Silic, M. & Back, A., 2015. 'Identification and importance of the technological risks of open source software in the enterprise adoption context.' *12th International Conference on Wirtschaftsinformatik*, Osnabruck, Germany, 1163 – 1176.

Singh, A., Bansal, R. K. & Jha, N., 2015. 'Open source software vs proprietary software.' *International Journal of Computer Applications*, 114(18), 26 – 31.

SITA FOSSFocus 2009. 'Report on FOSS implementation in National Government Departments.' from http://www.sita.co.za/FOSS/Docs/FOSSFOCUS%204th%20Issue%20Oct%2009.pdf

Small, M. L., 2011. 'How to conduct a mixed methods study: Recent trends in a rapidly growing literature.' *Annual Review of Sociology*, 37, 57 – 86.

Smit, N. C., 1990. 'The case study: A useful research method for information management.' *Journal of Information Technology*, Chapman and Hall, London, 5, 123 – 133.

Sodhi, G. K., 2015. ‘Database security and confidentiality.’ *International Journal of Advanced Research in Computer and Communication Engineering*, 4(2), 309 – 310.

Stake, R. E., 1995. *The art of case study research*. Sage.

Statista, 2015. ‘Projected revenue of open-source software from 2008 to 2020.’ Available from: www.statista.com/statistics/270805/projected-revenue-of-open-source-software-since-2008/. Accessed on: 2015/10/23.

STATISTICS SA, 2013. ‘Mid-year population estimates.’ Statistical Release P0302, available from: http://beta2.statssa.gov.za/publication/P0302/P03022013.pdf.

Stevenson, M., Helmond, A. & Driscoll, K., 2015. ‘Systems, syntax and snippy: Accounting for software in web history.’ Two day conference at Aarhus University, Denmark.

Stewart, C. D., 2009. ‘A multidimensional measure of professional learning communities: The development and validation of the Learning Community Culture Indicator (LCCI)’, Dissertation, Brigham Young University.

Stockdale, R. & Standing, C., 2002. ‘Case Studies in context: an examination of research influences.’ *ACIS Proceedings*, paper 14. Available from: http://aisel.aisnet.org/acis2002/14. Accessed on: 2014/03/21.

Stockdale, R. & Standing, C., 2006. ‘An interpretive approach to evaluating information systems: A content, context, process framework.’ *European Journal of Operational Research*, 173, 1090 – 1102.

Stol, K., Babar, M. A., Russo, B. & Fitzgerald, B., 2009. ‘The use of empirical methods in open source software research: Facts, trends, and future Directions.’ *FLOSS*, Vancouver, Canada, 19 - 24.

Stoyanov, B. & Kordov, K., 2014.‘Open source software alternatives in higher education following computer science curricula 2013.' *Research Journal of Applied Sciences, Engineering and Technology*, 8(9), 1160 – 1163.

Strauss, A. L. & Corbin, J., 1998. *Basics of qualitative research*. (2nd edn.), Thousand Oaks, CA: Sage.

Sullivan, G. M., 2011. 'A primer on the validity of assessment instruments.' *Journal of graduate medical education*, 119 – 120.

Svennevig, J., 2001. 'Abduction as a methodological approach to the study of spoken interaction.' *Norskrift*, 103, 1 – 22.

Svensson, D. & Magnusson, B., 2004. 'An architecture for migrating user interfaces.' In *NWPER*, Turku, Finland, 31 – 44.

Swartz, J., 2005. '2005 worst year for breaches of computer security.' In: *USA Today*.

Swiderski, F. & Snyder, W., 2004.*Threat modelling*. Redmond (WA): Microsoft Press.

Symonds, J. E., & Gorard, S., 2009. 'The death of mixed methods: Research labels and their casualties.' The British educational research asociation annual conference, Herwt Watt University, Edinburgh, Sept. 3 – 6.

Taatila, V. & Raij, K., 2012. 'Philosophical review of pragmatism as a basis for learning by developing pedagogy.' *Educational Philosophy and Theory*, 44(8), 831 – 844.

Tabachnick, B. G. & Fidell, L. S., 2001. *Using Mulivariate Statistics*. 4th edn. New York: Harper & Row.

Tankard, C. & Pathways, D., 2015. 'Data classification – the foundation of information security.' *Network Security*, 8 – 11.

Tartari, V. & Breschi, S., 2012. 'Set them free: Scientists' evaluations of the benefits and costs of university-industry research collaboration.' *Industrial and Corporate Change*, 21(5), 1117 – 1147.

Tashakkori, A., 2007. 'Editorial: the new era of mixed methods.' *Journal of Mixed Methods Research*, 1(1), 3 – 7.

Tashakkori, A. & Creswell, J. W., 2007. 'Editorial: the new era of mixed methods.' *Journal of Mixed Methods Research*, 1(1), 1 – 8.

Tashakkori, A. & Teddlie, C., 1998. *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks: Sage Publications.

Tashakkori, A. & Teddlie, C., 2003a. 'Issues and dilemmas in teaching research methods courses in social and behavioural sciences: A US perspective.' *International Journal of Social Research Methodology*, 6(1), 61 – 77.

Tashakkori, A. & Teddlie, C., 2003b., 'The Past and the future of mixed methods research: From methodological triangulation to mixed methods designs.' *In Handbook of Mixed Methods in Social and Behavioural Research*, A.

Tashakkori and C. Teddlie (eds.), Thousand Oaks, CA: Sage Publications, 671 – 701.

Tate, R. A., 2003. 'A comparison of selected empirical methods for assessing the structure of responses to test items.' *Applied Psychological Measurement*, 27, 159 – 203.

Tavakol, M. & Dennick, R., 2011. 'Making sense of Cronbach's alpha.' *International Journal of Medical Education,* 2, 53 – 55.

Taylor, R. G., 2006. 'Management perception of unintentional information security risks.' 27th International Conference on Information Systems, Milwaukee.

Taylor, S. S., Fisher, D. & Dufresne, R. L., 2002. 'The aesthetics of management storytelling: a key to organisational learning.' *Management Learning,* 33(3), 313 – 330.

Teddlie, C. & Tashakkori, A., 2003. 'Major issues and controversies in the use of mixed methods in the social and behavioural Sciences.' In *Handbook of Mixed Methods in Social and Behavioural Research*. A Tashakkori & C. Teddlie (eds.), Thousand Oaks, CA: Sage Publications, 3 – 50.

Teddlie, C. & Tashakkori, A., 2007. *Foundations of mixed methods research*. Thousand Oaks, CA: Sage Publications.

Teddlie, C. & Tashakkori, A., 2009. *Foundations of mixed methods research*. Thousand Oaks, CA: Sage Publications.

The Guardian, 2004. 'Open invitation taken up at last', Viewed on 2 February 2013, from: http://society.guardian.co.uk/e-public/story/0,,1362744,00.html.

The Open Group, 2009. 'COA paper – Information classification.' from:

www.opengroup.org/jericho/COA_informationClassification_v1.0.pdf.

Thomas, D., 2005. 'Going open source software in IT opportunities and

challenges.' *Journal of Object Technology*, 4(2).

Thomas, E. & Magilvy, J. K., 2011. 'Qualitative rigour or research validity in

qualitative research.' *Journal for specialists in pediatric nursing*, 151 – 155.

Thomas, J., 2007. 'Malaysian public sector OSS program Phase II: Accelerated adoption', available from:

http://www.oscc.org.my/documentation/phase2_launching/OSS-Phase-2Strategy-Plan-Launch.pdf.

Thompson, B., 1992. 'A partial test distribution for cosines among factors across samples.' In B. Thompson (ed.) *Advances in Social Science methodology*, 2, 81 – 97.

Thompson, B. & Daniel, L. G., 1996. 'Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines.' *Educational and Psychological Measurement*, 56, 197 – 208.

Thompson, E. D. & Kaarst-Brown, M. L., 2005. ‘Sensitive information: A review and research agenda.’ *Journal of the American Society for Information Science and Technology*, 56(3), 245 – 257.

Thomson, K. L. R., von Solms, R. & Louw, L., 2006. ‘Cultivating an organizational information security culture’ *Computer Fraud & Security*, 10, 7 – 11.

Tinsley, H. E. A. & Tinsley D. J., 1987. ‘Uses of factor analysis in counseling psychology research.’ *Journal of Counseling Psychology*, 34(4), 414 – 424.

TJNAF, 2007. ‘Security Plan for protection of sensitive information.’ Thomas Jefferson National Accelerator Facility.

TMPSOSSSMP, 2008. ‘The Malaysian public sector open source software master plan: Phase II – Accelerated adoption.’ Available from:

http://www.mampu.gov.my/seminar%20ict/kk2-   OSS.pdf. Accessed:

2012/02/19.

Torchiano, M., Di Penta, M., Ricca, F., De Lucia, A. & Lanubile, F., 2011. ‘Migration of information systems in the Italian industry: A state of the practice survey.’ *Information and Software Technology*, 53, 71 – 86.

Trauth, E. M. & Jessup, L. M., 2000. ‘Understanding computer-mediated discussions: Positivist and interpretive analyses of Group Support System use.’ *MIS Quarterly*, 24, 43 – 79.

Treiblmaier, H. & Filzmoser, P., 2009. ‘Exploratory factor analysis revisited: How robust methods support the detection of hidden multivariate data structures in IS research.’ Vienna University of Technology. Available from:

http://www.statistik.tuwien.ac.at.

Trochim, W., 2006. ’Positivism vs. post-positivism.’ Web center for social research methods. Available from:

http://www.socialresearchmethods.net/kb/positivism.htm. Accessed on

2015/11/24.

Trotter II, R. T., 2012. 'Qualitative research sample design and sample size: Resolving and unresolved issues and inferential imperatives.' *Journal of Preventative Medicine*, 55, 398 – 400.

US Congress, 2003. 'Joint enquiry into intelligence community activities before and after the terrorist attacks on September 11, 2001.' Retrieved on December 4, 2009, from: http://www.gpoaccess.gov/serialset/creports/911.html.

Vadalasetty, S. R., 2009. 'Security concerns in using open source software for enterprise requirements.' SANS Institute InfoSec Reading Room.

Van Belle, J., Brink, D., Roos, L. & Weller, J., 2006. 'Migrating to OSS-on-the-desktop: Lesson learnt and a proposed model. Proceedings of the 38th Southern Africa Computer Lecturers Association Conference, Somerset West, South Africa, 94 – 107.

Vargas, R. J. G., Anaya, E. A., Huerta, R. G. & Hernandez, A. F. M., 2012. 'Security controls for android.' *Fourth International Conference on Computational Aspects of Social Networks (CASoN),* 212 – 216.

Vasa, M., Jadatharan, A. & Srivasta, B., 2015. 'Towards risk-aware planning of service delivery operations.' 12th International Conference on Services Computing, *IEEE*, New York, USA.

Venkatesh, V., Brown, S. A. & Bala, H., 2013. 'Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems.' *MIS Quaterly*, 37(1), 21 – 54.

Vintila, B., 2010. 'Citizen oriented open source security.' *Open Source Science Journal* 2(3), 57 - 64.

Vital W., 2006. 'The South African adoption of open source.' White paper created by Vital Wave Consulting available online from www.vitalwaveconsulting.com/insights/South-African-Adoption-of-Open-Source.pdf, accessed on 20 February 2012.

Von Solms, B., 2000. ‘Information security – the third wave?’ *Computers & Security*, 19(7), 615 – 620.

Von Solms, R. & van Niekerk, J., 2013. ‘From information security to cyber security.’ *Computers & Security*, 38, 97 – 102.

Walia, N., Rajagopalan, B. & Jain, H., 2006. ‘Comparative investigation of vulnerabilities in open source and proprietary software: An exploratory study.’ *Americas Conference on Information Systems (AMCIS)*, 108, 848- 857.

Walker, T., 2004. ‘The future of open source software in government.’ Available from: http://www.oss-institute.org/newspdf/walker_oss_white_paper_2292004.pdf. Accessed on 2012/01/24.

Walsham, G., 1993. *Interpreting information system in organisations.* John Wiley, Chichester.

Walsham, G., 1995. ‘Interpretive case studies in IS research: nature and method.’ *European Journal of Information Systems*, 4(2), 74 – 81.

Walsham, G., 2006. ‘Doing interpretive research.’ *European Journal of Information Systems*, 15, 320 – 330.

Weber, R., 2004a. ‘The rhetoric of positivism versus interpretivism: A personal view.’ *MIS Quarterly*, 28(1), iii – xii.

Weber, T., 2004b. *The success of open source.* Harvard University Press, New York, NY.

Weilbach, L & Byrne, E., 2010. ‘A human environmentalist approach to diffusion in ICT policies – A case study of the FOSS policy of the South African Government.’ *Journal of Information Communication & Ethics in Society,* 8(1) 108 – 123.

Weilbach, L. & Byrne, E., 2011. 'Implementing open source software to conform to national policy.' *Journal of Systems and Information Technology*, 13(3), 286 – 302.

Welch, C., Piekkari, R., Plakoyiannaki, E. & Paavilainen-Mantymaki, E., 2011.

'Theorising from case studies: Towards a pluralist future for international business research.' *Journal of International Business Studies*, 42, 740 - 762.

Westerman, M. A. & Yanchar, S. C., 2011. 'Changing the terms of the debate; Quantitative methods in explicitly interpretive research.' *Theory & Psychology*, 21(2), 139 – 154.

Wheeler, D. A., 2005. 'Why open source software/free software? Look at the numbers!' Available from: http://www.dwheeler.com/oss_fs_why.html.

Accessed on 2012/02/11.

Whitman, M. E. & Mattord, H., 2008. *Principles of information security*. 2nd edn., Course Technology, Boston, MA.

Williams, C., 2007. 'Research methods.' *Journal of Business & Economic Research*, 5(3), 65 – 72.

Wiid, J. & Diggines, C., 2013. *Marketing research.* 2nd edn. Juta & Company Ltd., Cape Town, South Africa.

Whittemore, R., Chase, S. K. & Mandle, C. L., 2001. 'Validity in qualitative

Research.' *Qualitative Health Research*, 11(4), 522 – 537.

Witten, B., Lanwehr, C. & Caloyannides, M., 2001. 'Does open source improve system security?' *IEEE Software*, Sept – October, 57 – 61.

Woo, S. W., Alhazmi, O. H. & Malaiya, Y. K., 2006. 'An analysis of the vulnerability discovery process in web browsers.' In *Proceedings of the 10th International Conference on Software Engineering and Applications*, Dallas, TX, Nov. 13 – 15.

Worthington, R. L. & Whittaker, T. A., 2006. 'Scale Development Research: A Content Analysis and Recommendations for Best Practices.' *The Counseling Psychologist*, 34(6), 806 – 838.

Yardley, L.. & Bishop, F., 2008. 'Mixing qualitative and quantitative methods: A pragmatic approach.' In C. Willig & W. Stainton Rogers (ed.), *The Handbook of qualitative research in psychology*, London, UK: Sage, 352 – 369.

Yeo, B., Liu, L. & Saxena, S., 2006. 'When China dances with OSS.' In DiBona, C., Cooper, D. and Stone, M. (eds.), *Open Sources 2.0*, O'Reilly Media, Sebastopol, CA, 197 – 210.

Yin, R. K., 2003. *Case study research, design and methods*. Applied social research methods series. 3rd edn., SAGE Publications, Inc.

Yin, R. K., 2009. '*Case study research design and methods.'* SAGE Publications Inc., Thousand Oaks, CA.

Yu, C. H., 2003.'Misconceived relationships between logical positivism and quantitative research.' Research Methods Forum. Available from http://www.aom.pace.edu/rmd/2002forum.html.

Zwick, W. R. & Velicer, W. F. 1986. 'Factors influencing five rules for determining the number of components to retain.' *Psychological Bulletin*, 99, 432 – 442.

List of Personnal Communications

Phala, S., South African National Department of Arts & Culture, ICT Division, In discussion, 17/07/2013.

Webb, A., State Information Technology Agency, FOSS Programme Office, In discussion, 3/06/2009.

Zwane, B., South African National Public Works, ICT Division, In discussion, 8/08/2013.

APPENDIX A: Definitions of Terms

AICTA: Agency for Information and Communication Technology (based in France)

AJIS: African Journal of Information Systems

BS17799: a Code of Practice for Information Security; renamed to ISO27002.

CDP: Continuous Data Protection

CFA: Confirmatory Factor Analysis

CIA: Confidentiality, Integrity and Availability

CSIR: Council for Scientific & Industrial Research

CSS: Closed Source Software

CMM/CMMi: Capability Maturity Model Integration – is a general model which determines organisation maturity with regard to realization of given goals and enables to improve organisational inner processes in an organised and ordered manner

CobiT: Control Objectives for Information – is a coherent and clear model/set of best practices for IT management, addressed to managers, auditors, and users of information technologies

DeMSET: Design and Modelling in Science, Education and Technology

DLP: Data Loss Protection

DR: Design Research

EFA: Exploratory Factor Analysis

eNaTIS: Electronic National Traffic Information System

ETL: Extract, Transfer and Load

FBI: Federal Bureau of Investigation (USA)

FOSS: Free Open Source Software

FSF: Free Software Foundation

FTC: Federal Trade Commission (USA)

GAO: General Accounting Office (USA)

GITOC: Government Information Technology Officers’ Council

GPL: General Public Licence

GT: Grounded Theory

HTML: Hypertext Markup Language

ICITST: International Conference for Internet Technology and Secure Transactions

IDC: International Data Corporation (USA)

IFMS: Integrated Financial Management System

IS: Information Systems

ISO 17799: International Security Management Standard first published by International Organisation for Standardisation (ISO) in December 2000

IT: Information Technology

ITIL: Information Technology Infrastructure Library – is a set of complex recommendations of IT industry, on the basis of which the international norm for IT service management – (ISO/IEC 2000) was created

KMO: Kaiser-Meyer-Olkin (measure)

NACI: National Advisory Council on Innovation

NaTIS: National Traffic Information System

NHS: National Health Service (UK)

NLSA: National Libraries of South Africa

NIST: National Institute of Standards and Technology (USA)

NSA: US National Security Agency

OS: Operating System

OSD: Open Source Definition

OSI: Open Source Institute

OSS: Open Source Software refers to software that is freely available, accessible, reusable and which the source codes can be modified to make them work as their users need (Sarab *et al.* 2013)

PM: Project Management

PNC: Presidential National Commission

POS: Point of Sale

SA: South Africa

SARS: South African Revenue Services

SAS: Statistical Analysis System

SDLC: Software Development Life Cycle

SEM: Structural Equation Modelling

SITA: State Information Technology Agency

SLA: Service Level Management

SLED: Suse Linux Enterprise Desktop

TCO: Total Cost of Ownership

UML: Unified Modelling Language

US: United States

USA: United States of America

XML: Extensible Markup Language

APPENDIX B: Questionnaires for Quantitative Research

A MANAGEMENT FRAMEWORK TO CONCEPTUALIZE INFORMATION SENSITIVITY DURING MIGRATION OF SOFTWARE PLATFORMS

Dear Respondent
This survey forms part of a doctoral thesis entitled *A Management Framework to Conceptualize Information Sensitivity during Migration of Software Platforms,* for the degree of PhD (Information Systems) at the University of South Africa.
The objectives of this research are to:

- ensure the safety of sensitive information when they are migrated from a closed source platform to an open source platform.
- to define the protection measures that should be undertaken during the migration from a closed source platform to an open source platform.
- to develop a Management Framework for the migration of sensitive information during software platform migrations.

You are kindly requested to complete this survey questionnaire, comprising of seven sections, as honestly and frankly as possible and according to your personal views and experience. You have the right not to complete all questions if so desired.
You are not required to indicate your name but your age, gender, occupation position etc. will contribute to a more comprehensive analysis. All information obtained from this questionnaire will be used for research purposes only.
After completion of the thesis, a summary of the findings of the research will be available to respondents on request.
Any enquiries may be made to Mr. O. A. Ajigini, email: olusega@gmail.com, Prof. J. A. van der Poll, email: vdpolja@unisa.ac.za, or Prof. J. H. Kroeze, email: kroezjh@unisa.ac.za .
Thank you for your cooperation.

Mr. Olusegun Ajigini

A MANAGEMENT FRAMEWORK TO CONCEPTUALIZE INFORMATION SENSITIVITY DURING MIGRATION OF SOFTWARE PLATFORMS

Instructions:

1. Please mark your choice with an **'X'** in the relevant field and select only one option unless otherwise indicated.
2. The questionnaire consists of six sections.

Section A: Biographical Data
Section B1: Employee behaviour: staff awareness of the sensitive nature of company data
Section B2: Training: awareness of the sensitive nature of data
Section B3: Employee accountability: awareness of the sensitive nature of data
Section C1: Organisational strategy: handling sensitive nature of data
Section C2: Organisational Policies and Procedures: handling sensitive nature of data
Section C3: Organisational Data: preparations towards ensuring sensitive data management
Section C4: Organisational Standards (Processes, Hardware & Software): enforcement will ensure proper handling of sensitive data.
Section D1: Data Categories and Business Rules: providing a basis for data classification during migration of sensitive data.
Section D2: Data Classification System: addressing security issues when handling sensitive data during migration of platforms.
Section D3: Data Protection Tools: ensuring sensitive data protection during migration of platforms.
Section D4: Data Sensitivity Assessment: identifying different protection needs for information.
Section D5: Security Models: ensuring protection of sensitive data during migration.
Section E1: Data migration planning: protecting sensitive data during migration.
Section E2: Data migration process: protecting sensitive data during migration.
Section E3: Data migration tools: protecting sensitive data during migration.
Section E4: Data migration controls: protecting sensitive data during migration.
Section E5: Data migration monitoring: protecting sensitive data during migration.
Section F: Further Comments

Section A: Biographical Data

Office use only

Serial no □□□

3

1. **Company**

1	SITA	
2	South African Revenue Services	
3	Presidential National Commission	
4	Department of Public Works	
5	National Library of South Africa	
6	Department of Arts & Culture	
7	Department of Social Development	
6	Other	

□

1

2. **Respondent Post Levels (IT Specialists)**

1	Senior Developer / IT Senior Administrator	
2	Developer/ IT Administrator	
3	Junior Developer/IT Junior Administrator	
4	IT Manager	
5	IT Senior Manager	

□ 2

3. Respondent Post Levels (Non IT Staff)

1	Executive Manager	
2	Senior Manager	
3	Manager	
4	Super End User	
5	End User	

□ 3

4. Sex

1	Male	
2	Female	
3	Other	

□ 4

5. Race

1	Black	
2	Asian	
3	Coloured	
4	White	
5	Foreign	

□ 5

6. Age

1	< 25years	
2	26 – 35years	
3	36 – 45years	
4	46 – 55years	
5	55+ years	

□ 6

7. **Type of work the respondent**

1	Creating sensitive data	
2	Extract, Transfer & Load Data/ETL Migration	
3	Data Security/IT security	
4	Software Coding/Developer	
5	Database Administrator	
6	Storage Administrator	
7	IT Security Administrator	

□ 7

8. **On what basis are you employed?**

1	Permanent	
2	Temporary	
3	Fixed	

□ 8

	term/contract	
4	Other	

9. **Doe 9. Are you aware of a policy on sensitive data management in organisations?**

1	Yes	
2	No	

□ 9

10. Number of years with the current company?

1	2 years or less	
2	3-5 years	
3	6-10 years	
4	10 years+	

□ 10

11. Have you ever been part of a Platform Migration project within your organisation?

1	Yes	
2	No	

□ 11

12. Are you are engaged in any of the following (please check as many of the options that apply): (Perhaps list activities which gives a person data management access/ powers -

1	Giving Access rights to Users	
2	Developing IT Policies & Procedures	
3	Classification of data	
4	Reviewing the security of IT systems	

□
□
□
□ 15

Section B1

Employee Behaviour: Staff awareness of the sensitive nature of company data

(an *employee behaviour* score could indicate how aware are staff of the risk & consequences associated with unauthorized access to some of the data they manage/ work with/ communicate/generate = creating an awareness score for respondents within an organization)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.
Use the following five-point scale in Section B1:

1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Employees should be aware of the concept of sensitive data						□
2. Employees should handle sensitive						

information as outlined in their organisational Data Management Policy						□ □ □ 19
3. Inherently data reflect sensitivity levels should be associated with the (severity of) consequences unauthorized access (or use) of data						
4. It could have dire consequences for any organisation if lapses in information security (e.g. unathourised access, phishing) occurs						

Section B2: Training: awareness of the sensitive nature of data

(a *training score,*- calculated on the response of participants to the questions below - could give an indication of the company's success in empowering staff and management to ensure data security)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.
Use the following five-point scale in Section B2:

1: Strongly disagree   2 : Disagree
3: Neutral   4: Agree   5: Strongly agree

ITEMS						
1. Induction courses at any organisation should cover various aspects of the risks attached to the management of sensitive data						□ □ □ □ 23
2. Training should spell out the consequences of the misuse of sensitive data and also the risk of not protecting sensitive data.						
3. Training for data system analyst/ IT specialists should cover detail policy procedures to protect against information theft						
4.All employees should be educated about the different classification levels, their respective markings and when to apply them.						

Section B3: Employee accountability: awareness of the sensitive nature of data

(an *employee accountability* score could reflect the extent that employees are accountable in ensuring that sensitive information protection is in line with the Policies and Procedures governing sensitive information)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.

Use the following five-point scale in Section B3:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1.Accountability should be valued when handling data						□ □ □ □ 27
2. All staff members should adhere to the sensitive data management policy in their organisations.						
3. The protection of sensitive information should be considered during data migration between platforms						
4. Employees should handle sensitive information with upmost care						

Comments	
Do you have comments on **Employee Issues**?	

Section C1: Organisational Strategy: handling sensitive nature of data

(an *organisational strategy* score could indicate how organisational preparedness in handling sensitive information)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.

Use the following five-point scale in Section C1:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Organisational strategy should incorporate how organisational data will be protected and handled.						□ □ □ □ 31
2. A clear objective on handling sensitive data should be developed and must be aligned with organisational strategy.						
3. Protecting sensitive information should be part of any organisational corporate culture						
4. Data integrity should be the hallmark of any organisation						

Section C2: Organisational Policies and Procedures: handling sensitive nature of data

(an *organisational policies and procedures* score will indicate how policies and procedures are used to ensure protection of sensitive data during migration of platforms)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.

Use the following five-point scale in Section C2:

1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Companies should have a data security policy which lists data security methods and sensitive data management.						□ □ □ □ 35
2. Policies and Procedures on sensitive information management should be regularly communicated and enforced to all staff.						
3. Continual update of data sensitivity policy should be in place						
4. Data security procedures should be explained in detail in the organisational policy on data security						

Section C3: Organisational Data: preparation towards ensuring sensitive data management

(an *organisational data* score will indicate its preparation prior to migration between the platforms)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.

Use the following five-point scale in SectionC3:

1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Organisations should identify the data to be migrated and where it resides during migration of platforms.						□ □ □ □ 39
2. The access to data should be controlled and monitored.						
3. Organisational data should be defined via data discovery and classification						
4. Extremely sensitive data should be accessed only during office hours						

from the organisation Pc or should be accessed after hours only from a specific laptop via a Virtual Private Network (VPN) that uses multi-factor authentication in any organisation.						

Section C4: Organisational Standards (Processes, Hardware & Software): enforcement will ensure proper handling of sensitive data

(an *organisational standard* score will reflect on how standards are used to ensure sensitive data protection during migrations)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.

Use the following five-point scale in Section C4:

1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. There should be a process to identify sensitive information that is worth protecting in organisations.						43
2. Confidentiality, integrity, identifying authorised users and monitoring access should be undertaken to ensure sensitive data protection.						
3. Enforcing hardware and software standards should be performed in other to eliminate unknown factors that assess sensitive information.						
4. Organisations should identify the data to be migrated and where it resides during platform migrations.						

Comments	
Do you have comments on **Organisational Issues**?	

Section D1: Data Categories and Business Rules: providing a basis for data classification during migration of sensitive data

(a *data categories and business rules* score calculated on the responses to the questions below could give an indication to management re the status of a company's 'preparedness' in providing a basis for data classification during migration of sensitive data)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.
Use the following five-point scale in Section D1:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Obligations to protect data should be observed by all who process data, independent of where such processing occurs.						□ □ □ □ □ 48
2. Organisational data should be defined via data discovery and classification						
3. IT Departmental employees should know where the data to be migrated resides, its value to the organisation and who can use it as part of data definition.						
4. Data users should be segregated by limiting/ restricting access to sensitivity categories of data						
5. Examination of business rules should be performed to provide a basis for classification						

Section D2: Data Classification System: addressing security issues when handling sensitive data during migration of platforms.

(a *data classification systemscore* calculated on the responses to the questions below could give an indication to management re the status of a company's 'preparedness' in addressing security issues when handling sensitive data during migration of platforms)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.
Use the following five-point scale in Section D2:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. All the data created by users should be classified or identified and proactively marked before they are migrated						□ □ □ □
2. New data should be classified first and then the legacy data later during						

migration of data between platforms.						52
3. Information creators (e.g. end users and management) should be involved in the classification of data as part of any organisation way of doing business.						
4. Data classification roles and responsibilities (e.g. data creators, data owners, data users, and data auditors) should be clearly defined within any organisation.						

Section D3: Data Protection Tools: ensuring sensitive data protection during migration of platforms

(a *data protection tools*score calculated on the responses to the questions below could give an indication to management re the status of a company’s ‘preparedness’ in protecting sensitive information)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.
Use the following five-point scale in Section D3:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Encryption techniques and management should be in place in organisations.						57
2. Technical controls such as encryption or rights management should be automated at the time of content publication.						
3. Organisations should have data discovery tools and software that can scam endpoints or corporate network assets to identify resources that could contain sensitive information such as hosts, database columns and rows, web applications, storage networks and file shares.						
4. Tools should be used for the automation and enforcement of dynamic data classification						
5. Complementary, endpoint-oriented technologies (e.g. end-user notification, encryption, secure/managed file transfer and rights management) should be used to enforce established security policies						

Section D4: Data sensitivity assessment: identifying different protection needs for information

(a *data sensitive assessment* score calculated on the responses to the questions below could give an indication on how to identify different protection needs for information and how to handle sensitive data)

Please answer each item by ticking off "X" the relevant option.
Use the following five-point scale in Section D4:
Legend:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree
5: Strongly agree

ITEMS						
1. Continual management of data sensitivity management and new risk assessments should be in place						
2. Monitoring the flow of sensitive data communication/ transfer by all staff members should be in place						
3. Monitoring staff managing/ using the sensitive data should be performed by organisations.						61
4. Monitoring database activity to identify and evaluate content in real time across multiple channels should be in place						

Section D5: Security Models: ensuring protection of sensitive data during migration

(a *security model*score calculated on the responses to the questions below could give an indication to management re the status of a company's 'preparedness' in ensuring protection of sensitive information)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.
Use the following five-point scale in Section D5:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Security Models should be developed to support organisational strategy						
2. Basic security models should be						

addressed when handling sensitive data						65
3. Security Models should ensure confidentiality, integrity and reliability of data during protection of sensitive information.						
4. Security Models should be used to protect sensitive information during migrations.						

Comments	
Do you have comments on **Data Issues**?	

Section E1: Data migration planning: Protecting sensitive data during migration

(a *data migration planning*score calculated on the responses to the questions below could give an indication to management re the status of a company's 'preparedness' in ensuring sensitive information protection during migration)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.

Use the following five-point scale in Section E1:

1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. All the applications, functions, host servers and storage impacted by the data migration should be identified during data migration.						69
2. Enough time should be planned for the data migration process.						
3. All the data in the servers, memory and buffers should be de-staged to disc before migration in organisations.						
4. Source data should be backed up prior to data migration to the destination.						

Section E2: Data migration process: protecting sensitive data during migration

(a *data migration process*score calculated on the responses to the questions below could give an indication to management re the status of a company’s ‘preparedness’ in protecting sensitive information)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.

Use the following five-point scale in Section E2:

1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Data migrations should be done only during overnight hours (non-business hours)						□ □ □ □ □ 74
2. The timing of the migration, how long it will take and how long the systems will be down (if necessary) should be determined.						
3. The migration project should be reviewed for issues to correct or improve for the next migration.						
4. Scripts (if used) used to perform the migration should be reviewed for reliability and accuracy.						
5. The administrators should understand the true end-to-end relationships among the platforms being migrated.						

Section E3: Data migration tools: protecting sensitive data during migration

(a *data migration tools*score calculated on the responses to the questions below could give an indication to management re the status of a company’s ‘preparedness’ in protecting sensitive information)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.

Use the following five-point scale in Section E3:

1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Proper data migration tools and data migration strategies should be in place.						□ □ □ □ 78
2. Cipher text encryption on all data migrated should be enforced during migration of platforms.						
3. Organisations should use						

Continuous Data Protection (CDP) technology to protect sensitive information during data migrations.						
4. Organisations should use Data Loss Prevention (DLP) tools to quarantine sensitive data during migration.						

Section E4: Data migration controls: : protecting sensitive data during migration

(a *data migration controls*score calculated on the responses to the questions below could give an indication to management re the status of a company's 'preparedness' in protecting sensitive information)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.
Use the following five-point scale in Section E4:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						
1. Data migration policies should be implemented so that data is moved in an orderly manner.						82
2. The issues of data corruption, missed data or data loss should be considered during migration						
3. Migrated data should be tested and validated after migration in order to ensure data accuracy and integrity.						
4. Controls should be in place to ensure that sensitive data that are migrated are protected.						

Section E5: Data migration monitoring: protecting sensitive data during migration

(a *data migration monitoring*score calculated on the responses to the questions below could give an indication to management re the status of a company's 'preparedness' in protecting sensitive information)

Please indicate your extent of agreement with each of the following statements by ticking the appropriate box.
Use the following five-point scale in Section E5:
1: Strongly disagree 2 : Disagree
3: Neutral 4: Agree 5: Strongly agree

ITEMS						

1. The necessary monitoring systems and risk assessment systems should be in place.						86
2. The Operating systems level permissions, directory structure and share permissions should be recorded.						
3. The network bandwidth needs should be measured before migration and when it will be available.						
4. Verification or comparing migrated data versus source data should be performed.						

Comments	
Do you have comments on **Data Migration Issues**?	

Section F: Further Comments	

Please add some comments (only those that you think to be important and relevant to the aims of this questionnaire).	

Thank you for completing this questionnaire!

APPENDIX C: Qualitative Research Questions

QUALITATIVE RESEARCH QUESTIONS FOR RESEARCH ON INFORMATION SENSITIVITY

1. Do you understand the difference between sensitive information and non-sensitive information in organisations? Can you explain the difference?
2. How do you think that sensitive information should be protected during software migration?
3. What protection mechanisms ought to be implemented during the migration of information from a platform to another e.g. from a Proprietary Platform to an Open Source Platform?
4. What would be the properties of a Management Framework for the migration of sensitive information during platform migrations?
5. Do you think that organisations should have security models to support their organisational strategy? If so, Why?
6. Do you think that the organisational strategy should incorporate how the organisational data will be protected and handled? If so, Why?
7. Do you believe that employees handling organisational data should be trained on how to handle sensitive information? If so, Why?
8. Should employees perform sensitivity assessment as part of the organisational strategy in protecting their organisational data? If so, Why?
9. Should organisations have Policies and Procedures on handling sensitive information? If so, Why? Should they be enforced?
10. Is it important for organisations to control and monitor their data access by their employees? If so, Why?
11. Should Technical Controls be put in place by organisations during data migrations? If so, Why?
12. Should the organisational source data be backed up prior to migration? If so, Why?
13. Is it important to determine the duration of the migration process before the migration of data? If so, why?
14. Should organisations ensure that the necessary Monitoring and Risk Assessment Systems are in place prior to migration of their data? If so, Why?
15. Is it important for organisations to ensure the availability of adequate Network Bandwidth before commencing on a migration process? If so, Why?
16. Should proper Migration Tools and Strategies be provided prior to migration of data? If so, Why?

17. Should the organisational Database Activities be monitored always? If so, Why?
18. Should organisations identify the Applications, Functions, Hosts, Host Servers, Storage impacted by the data migration? If so, Why?
19. Should organisational migrated data be classified prior to migration? If so, Why?
20. Is it important for organisations to clearly define the data classification roles and responsibilities of employees involved in the data migration (e.g. Data Creators; Data Owners; Data Users; Data Auditors)? If so, Why?
21. Should the ETL (Extract, Transfer and Load) Scripts used to perform the migration be reviewed for Reliability and Accuracy? If so, Why?
22. Is it important for the flow of sensitive data be monitored during the migration process? If so, Why?
23. Should the Servers, Memory and Buffers be de-staged before migration? If so, Why?
24. Should Business Rules be examined in order to provide a basis for information categorisation with respect to sensitivity? If so, Why?
25. Is it important to enforce data migration policies during migrations? If so, Why?
26. Should Security Tools such as Continuous Data Protection (CDP) OR Data Loss Prevention (DLP) be used to protect sensitive information during migration of data? If so, Why?
27. Should the migrated data be subjected to Cipher Text Encryption process during migration? If so, Why?
28. Should the encrypted migrated data be decrypted after the migration? If so, Why?
29. Is it important for the migrated data to be tested and validated after the migration in order to ensure data accuracy and integrity? If so, Why?
30. Do you believe that the migrated data should be subjected to a Data Quality Process after the migration process? If so, Why?
31. Do organisations have to consider the perceived reduction in the Cost or TCO (Total Cost of Ownership) of the new software when planning migration? If so, Why?
32. Do you think that Organisational Structure and Culture should be taken into consideration when planning migrations? If so, Why?
33. Do you think that IT Standards such as ISO/IEC 17799 be adhered to during software migrations? If so, Why?
34. Do you think that employee’s attitude and behaviour are some of the important human elements that should be considered during software migrations? If so, Why?
35. Do you think that we have mentioned and discussed all the issues pertaining to the properties of the management framework on information sensitivity during software migrations? If NOT, what remains?

APPENDIX D: Research Participation Form and Ethics Committee Letter

(A) LETTER TO REQUEST PERMISSION FROM GATE-KEEPERS OF PARTICIPATING ORGANISATIONS

RESEARCH BACKGROUND

CONFIDENTIAL

RESEARCHER INFORMATION	SUPERVISOR INFORMATION	CO-SUPERVISOR INFORMATION
Name: Mr. O. A. Ajigini	Prof. J. A. van der Poll	Prof. J.H. Kroeze
E-Mail:olusega@gmail.com	vdpolja@unisa.ac.za	kroezjh@unisa.ac.za
Contact Number: (+27) 082 627 0885	(+27) 011 652 0316	(+ 27) 011 670 9117
Institution: University of South Africa (UNISA)	University of South Africa (UNISA)	University of South Africa (UNISA)

STUDY INFORMATION

Title of Research: A Management Framework to Manage Information Sensitivity during Migration of Platforms

Objectives: (a) to ensure the protection of sensitive information when they are migrated from a Proprietary Platform to an Open Source Platform.
(b) to define the protection measures that should be undertaken during the migration from Proprietary Platform to an Open Source Platform
(c) to develop a Management Framework for the migration of sensitive information between platforms.

Nature: This study has a positive nature and aimed at enhancing the existing body of knowledge related to Information Sensitivity and Management Framework Development research.

Implications: Possible alternations to existing Information Sensitivity Management frameworks. Responses will be confidential and anonymous.

Duration of Study: Until 1st October 2014 (Date of submission)

PERMISSION TO COLLECT DATA FOR RESEARCH

I am a student at the University of South Africa, currently undertaking a PhD research study in Information Systems. As part of the degree requirements, l need to carry out research. I am writing to request your permission to collect data with some of your colleagues in your organization. In order to obtain the relevant information in this study, l will adopt a Mixed Method/Case study approach.

I will distribute questionnaires (attached herewith) to your colleagues in your organization and then later, I will collect data using structured and semi-structured interviews with some selected participants from your organization. For accurate data collection, l will audio-tape the interview proceedings. The plan is to collect data in the period from March 2014 till April 2014. The data so collected and the identity of the participants will be treated with confidentiality. The identities of the participants will be concealed, in any presentation and publication emerging from this research, by use of pseudonyms. The link to their real names will only be accessed by the researcher and the supervisor as well as the co-supervisor. If, however, for any reason, the participants would like their real names to be used in any future presentation, they will need to make written requests to me, as the researcher. There are no known or anticipated risks to the participants who will participate in this study.

Once the research has been completed, a brief summary of the findings will be available to the participants on request. The findings of the study will also be presented in academic conferences and published in National and International academic Journals. The participation of your organization in this research project is completely voluntary. Should your organization wish to withdraw at any stage, or withdraw any unprocessed data that will have been supplied, it will be free to do so without prejudice. The decision to participate or not, or to withdraw, will be completely independent of your organizational dealings with the University of South Africa.

I therefore request your permission to conduct research on the selected participants. I have enclosed a copy of participant consent form for your perusal. Should you have any questions or concerns regarding this letter or my research, please contact me at the address given above. You may also contact my supervisor or the co-supervisor on their email addresses given above.

Sincerely,

____

O. A. Ajigini 07th March 2014
PhD Candidate **Date**

(B) RESEARCH STUDY PARTICIPANT CONSENT FORM

RESEARCH BACKGROUND

CONFIDENTIAL

RESEARCHER INFORMATION	SUPERVISOR INFORMATION	CO-SUPERVISOR INFORMATION
Name: Mr. O. A. Ajigini	Prof. J. A. van der Poll	Prof. J.H. Kroeze
E-Mail: olusega@gmail.com	vdpolja@unisa.ac.za	kroezjh@unisa.ac.za
Contact Number: (+27) 082 627 0885	(+27) 011 652 0316	(+ 27) 011 670 9117
Institution: University of South Africa (UNISA)	University of South Africa (UNISA)	University of South Africa (UNISA)

STUDY INFORMATION

Title of Research: A Management Framework to Manage Information Sensitivity during Migration of Platforms

Objectives: (a) to ensure the protection of sensitive information when they are migrated from a Proprietary Platform to an Open Source Platform.
(b) to define the protection measures that should be undertaken during the migration from Proprietary Platform to an Open Source Platform
(c) to develop a Management Framework for the migration of sensitive information between platforms.

Nature: This study has a positive nature and aimed at enhancing the existing body of knowledge related to Information Sensitivity and Management Framework Development research.

Implications: Possible enhancements to existing Information Sensitivity Management frameworks. Responses will be confidential and anonymous.

Duration of Study: Until 1st October 2014 (Date of submission)

Safety & Health Implications: None

Duration of Participation: One Hour

PARTICIPANT RIGHTS

The participants remain the right to decide to participate in the study. The participant's privacy or dignity will not by violated by using hidden cameras, one-way glass, microphones, sound recordings or any other research devices, without his/her permission. Microphones, sound recordings or any other research devices, may be used where the participant's permission is implied by his/her presence and where it cannot be used elsewhere to their disadvantage. All information will be handled confidentially. The participant's identity will not be revealed and any conclusions derived from the study will be considered anonymous. The results of the study may be used for purposes of publication. Subjects will be provided with a copy of the Participant Information Form, as well as have its contents explained to them, before they consent to participating in the study.

PARTICIPANT PERMISSION FORM

Dear participant

Thank you for showing interest to participate in this PhD research study, which has the objective of developing a Management Framework for the migration of sensitive information between platforms. The study is conducted by Olusegun Ajigini under the supervision of both Prof J. A. van der Poll and Prof Jan Kroeze from the University of South Africa. Your participation is completely voluntary and the results will be treated as both confidential and anonymous, and will only be used for research purposes. The duration is of your participation is not expected to exceed one hour.

During this research, you will be asked to answer some questions relating to Developing a Management Framework for the migration of sensitive information between platforms, the problems you encounter, how you solve them, and measures taken to prevent future occurrence of the problems encountered. This research uses the Case Study method and Mixed Method (Quantitative & Qualitative). Data will be gathered through interviews and a Questionnaire (attached herewith). A number of semi-structured interviews will be conducted during the research. Each interview is designed to last for about an hour in length. The researcher is eager to learn from your practice. Feel free to expand on this subject or talk about related ideas that support your views. You are also free not to answer any questions you feel you cannot answer or that you do not feel comfortable answering. Feel free to indicate this when applicable and l will move on to the next question.

You will be assigned a code number which will protect your identity. All data will be kept in secured files, in accordance with the standards of the University of South Africa. All identifying information will be removed immediately after each interview is completed. Therefore, no one will be able to know your interview responses. Upon completion of this research project, all data and the questionnaires will be destroyed, or stored in a secure location where it can be accessed by the researcher on a need basis.

CONFIDENTIAL

Participant's Agreement

You will be provided with a copy upon signature of this form. The participant also has the right to withdraw their participation at any time.

I,..…………………………………………….. hereby voluntarily grant my permission for participation in the research project as explained to me by the researcher Olusegun Ajigini. The inputs derived from my participation will be interpreted and presented in a confidential nature and anonymous manner. The nature, objective, possible safety and health implications have been explained to me and I understand them. I understand my right to choose whether to participate in the project and that the information furnished will be handled confidentially and anonymously. I am aware that the results of the investigation may be used for the purposes of publication.

I understand the intent and purpose of this research. The researcher has reviewed the individual and social benefits and risks of this project with me. I am aware that data will be used for a dissertation, research paper, and a research presentation. I have the right to review, comment on, and/or withdraw information after giving the researcher reasonable time prior to submission of the research dissertation.

The data gathered in this study are confidential and anonymous with respect to my personal identity unless l specify/indicate otherwise. I grant permission for the use of this information for a:

Dissertation [ ]

Research Paper [ ]

I grant the permission to use one of the following:

My first name only:………………………………………………………………….

My Full name:…………………………………………………………………………..

Just a pseudonym:…………………………………………………………………..

I will be given a copy of the:

Paper [ ]

Audiotape [ ]

Transcribed interview [ ]

Additional conditions for my participation in this research are noted here: ……… ……… ……… …………………………………………

I have read the above form and, with the understanding that l can withdraw at anytime, and for whatever reason, l consent to participate in this interview and complete the questionnaire.

………………………………. ………………………………………..

Participant's signature **Date**

..……………………………… ……………………….. ..……………

Interviewer's signature **Date**

(C) ETHICS COMMITTEE LETTER

Olusegun Ademolu Ajigini (30107474) 2013-11-04

School of Computing

UNISA

Pretoria

Permission to conduct research project

Ref: 048/OG/2013

The request for ethical approval for your PhD in Information Systems research project entitled “A Management Framework to Conceptualise Information Sensitivity during Migration of Platforms” refers.

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

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

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

http://cm.unisa.ac.za/contents/departments/res_policies/docs/ResearchEthicsPolicyapprvCounc_21Sept07.pdf

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

Yours sincerely

Chair: School of Computing Ethics Sub-Committee

APPENDIX E: Tables on Biographical Data Distributions

TABLE E5-1

Frequency Table for the Type/Nature of Respondent Employment

Level	Frequencies	Percentages
SITA	20	22.222
PNC	8	8.889
Dept of Public Works	20	22.222
NLSA	14	15.556
Dept of Arts & Culture	8	8.889
Dept of Social Development	20	22.222
Total	90	100.000

TABLE E5-2

Frequency Table for the Respondent Post Level (IT Specialists*)*

Level	Frequencies	Percentages
Senior developer	13	14.607
Developer	44	49.438
Junior developer	25	28.090
IT manager	4	4.494
IT senior manager	3	3.371

Total	89	100.000

TABLE E5-3
Frequency Table for the Population of Respondents by Gender

Level	Frequencies	Percentages
Male	43	47.778
Female	47	52.222
Total	90	100.000

TABLE E5-4
Frequency Table for the Population of Respondents by Race

Level	Frequencies	Percentages
Black	82	91.111
Asian	1	1.111
Coloured	1	1.111
White	6	6.667
Total	90	100.000

TABLE E5-5
Frequency Table for the Age Distribution of Respondents

Level	Frequencies	Percentages
26 - 35 years	54	60.000
36 - 45 years	31	34.444
46 - 55 years	3	3.333
55+ years	2	2.222
Total	90	100.000

TABLE E5-6
Frequency Table for the Respondent’s Type of Work

Level	Frequencies	Percentages
Creating sensitive data	4	4.444
ETL	19	21.111
Data security	31	34.444
Software developer	7	7.778
Database administrator	12	13.333
Storage administrator	8	8.889
IT security administrator	9	10.000
Total	90	100.000

TABLE E5-7
Frequency Table for the Respondent’s Employment Category

Level	Frequencies	Percentages
Permanent	80	88.889
Fixed term	10	11.111
Total	90	100.000

TABLE E5-8
Frequency Table for the Respondent’s Awareness of Sensitive Data Management Policy

Level	Frequencies	Percentages
Yes	83	92.222
No	7	7.778
Total	90	100.000

TABLE E5-9

Frequency Table for the Respondent’s Number of Years in Service with Company

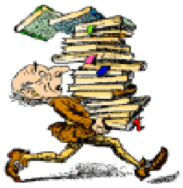
Level	Frequencies	Percentages
0-2 years	15	16.667
3-5 years	49	54.444
6-10 years	20	22.222
10+ years	6	6.667
Total	90	100.000

TABLE E5-10

Frequency Table for the Respondent’s Participation on Platform Migration Projects

Level	Frequencies	Percentages
Yes	85	94.444
No	5	5.556
Total	90	100.000

Letter from the Text Editor



Astute Editing and Research

20 October 2014

To Whom It May Concern

Dear Sir/Madam,

This is to certify that I have fully edited the PhD thesis of Mr. Olusegun Ajigini entitled *A framework to manage sensitive information during its migration between software platforms* for the University of South Africa. The text was checked for style, clarity and ease of reading, grammar and usage, spelling and punctuation, consistency in the use of text and figures in illustrations and tables, completeness and consistency in references, consistency in page numbering, headers and footers and suggestions were offered. I make no pretension to have improved the intellectual content of the thesis and did not rewrite any text. I presumed the text was in final form when I edited it. My suggestions are to be accepted or rejected by the author. The author effected the final changes himself.

Yours sincerely,

C.D. Schutte (D Litt et Phil, Full Member, Professional Editors' Group)

Telephone 012-342-3518 Mobile 083-310-1806

4 Gospel Close, 821 Church Street, Arcadia 0083, Pretoria.

Letter from the Statistician

20 October 2014

RE Statistical analysis of the Doctoral dissertation: "A FRAMEWORK TO MANAGE SENSITIVE INFORMATION DURING ITS MIGRATION BETWEEN SOFTWARE PLATFORMS"

To whom it may concern

This serves to confirm that HJ Gerber was involved in the empirical research efforts of Mr. Olusegun Ajigini for his Doctoral study.

HJ Gerber can vouch for the accuracy of the statistical evaluation undertaken for the empirical chapter of the student's dissertation.

Although every effort was made to ensure that the student presented the statistical results correctly, HJ Gerber cannot accept responsibility for the structure and presentation of the results of this study.

Kind Regards

Hennie Gerber
MCom (Statistics) UP
BCom Hons (Statistics)
UP BCom (Statistics) UP