UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

An investigation into the extent to which South African repositories comply with international trust standards

Mini-dissertation by

Glenn Tshweu

(29346836)

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF INFORMATION TECHNOLOGY (B)

in the

FACULTY OF ENGINEERING, THE BUILT ENVIRONMENT AND INFORMATION TECHNOLOGY

UNIVERSITY OF PRETORIA

Supervisors: Dr MJ van Deventer & Dr H Pienaar

December 2016

Abstract

An institutional repository is seen as a valuable tool to manage digital resources within the organisational context. Repositories can have a positive or negative influence on how an institution manages its digital material in relation to accessibility and dissemination of digital material.

The functionality and status quo of digital repositories can be assessed and measured based on specific guidelines to determine practicality and efficacy. The guidelines used in this regard are known as international repository assessment standards. These standards have been developed by leading organisations that specialise in knowledge creation to develop controlled, consensus-based, market-relevant international standards that can be used to support innovation and provide resolutions to global challenges. In the event where an institution wishes to assess its digital repository using international standards, the underlying purpose of the assessment exercise is for the digital repository to gain trust accreditation.

This study aimed to develop a South African digital repository trust assessment model based on the criteria of international standards. This study investigated the level of trust compliance that a very small sample of South African digital repositories met – using the developed model. The investigation process is also aimed at receiving feedback (in the form of recommendations) from digital repository managers to improve the developed model to make it more useful for South African digital repositories. Furthermore, the study intended to yield further research into the complex topic of digital repository assessment based on international standards. Overall, this research study revealed that South African digital repositories are not far off in complying with the full requirements of international repository assessment standards.

Acknowledgements

The author of this research study would like to thank the following people who contributed to make this mini-dissertation possible:

- My parents, for supporting me in everything that I do and their unconditional love and presence.
- Dr Lucia Lotter for sharing her insights and knowledge with me. Her role as my manager at work was a blessing in my life.
- Dr MJ van Deventer & Dr H Pienaar for acting as my supervisors and walking this journey with me; their assistance made it possible for me to complete this degree.
- Various repository managers who made time to be part of this study and share their experiences in repository management.

Table of Contents

Abstract ii
Acknowledgements iii
List of figures vii
List of tables viii
Glossary ix
CHAPTER 1 INTRODUCTION 1
1.1 Overview 1
1.2 Research statement and research questions 4
1.3 Research methodology 5
1.4 Limitations of the study 6
1.5 Justification for the research 6
1.6 Value of study 7
1.7 Division of chapters 7
1.8 Conclusion 8
CHAPTER 2: LITERATURE REVIEW 9
2.1 Introduction 9
2.2 The Open Archival Information System (OAIS) Reference Model 10
2.3 Organisations interested in repository evaluation 14
2.3.1 World Data System (WDS) 14
2.3.2 The Research Data Alliance (RDA) 15
2.3.3 Research Libraries Group (RLG) and Online Computer Library Center (OCLC) 16
2.3.4 The International Council for Science – World Data System (ICSU-WDS) 17
2.4 Digital repository assessment 18
2.4.1 Trustworthy Repositories Audit & Certification (TRAC) 18
2.4.2 The International Organization for Standardization (ISO) 16363 19
2.4.3 The 'Deutsches Institut für Normung' (DIN) 31644 20
2.4.4 European Framework for Audit and Certification of Digital Repositories 20
2.4.5 Brief comparison of the various international standards 21
2.5 Development of the proposed model 23
2.6 Summary 33
CHAPTER 3: RESEARCH METHODOLOGY 34

3.1 Introduction 34
3.2 Research methodology and design 34
3.2.1 Research approach and methodology 34
3.2.2 The research design 35
3.3 Target population and sampling 36
3.5 Data collection methods 37
3.5.1 Literature review 37
3.5.2 Face-to-face interviews 38
3.5.3 Semi-structured interviews 39
3.6. Data collection instrument – an interview schedule 41
3.6.1 The pilot study 42
3.7 Selection of research location 42
3.8 Data analysis and interpretation 42
3.9 Conclusion 43
CHAPTER 4: DATA ANALYSIS AND FINDINGS 44
4.1 Introduction 44
4.2 Background of the data collection and case study 44
4.3 Themes for data analysis 44
4.3.1 Segment 1: Admin coordination capability infrastructure 45
4.3.2 Segment 2: Ingest capability infrastructure 49
4.3.3 Segment 3: Data management capability infrastructure 51
4.3.4 Segment 4: Metadata management infrastructure 56
4.3.5 Segment 5: Access capability infrastructure 58
4.3.6 Segment 6: Preservation capability infrastructure 61
4.3.7 Segment 7: Financial sustainability infrastructure 63
4.3.8 Segment 8: System security sustainability infrastructure 64
4.4 Responses to additional questions 67
4.4.1 From a manger's point of view, what are the strengths and weaknesses of the model? 67
4.4.2 Which component of the model is most important for your institution? 68
4.4.3 Do you have any recommendations that can be used in redesigning the model? 68
4.4.4 To what extent do you see your institution being willing to comply with the requirements of such a system of trust? 69

4.4.5 What level of trust would your institution be able and willing to comply with, in terms of rating? 69
4.4.6 Which of the levels of the model need revision and what should the revision include? 69
4.4.7 From the developed model, what components were forgotten? 70
4.5 Conclusion 70
CHAPTER 5: RECOMMENDATIONS AND CONCLUSIONS 73
5.1 Introduction 73
5.2 Findings 73
5.2.1 Which international repository assessment standards can be used to assess South African digital repositories for trust status? 73
5.2.2 What will a trust model that has been developed based on international trust standards, look like? 74
5.2.3 To what extent do South African digital repositories comply with the developed OAIS based model in terms of trustiness? 75
5.2.4 How should and/or how could this model be adapted for South Africa as a developing country to make the striving for trustiness more feasible? 78
5.3 Evaluation of the research methodology used 78
5.4 Recommendations 78
5.5 Recommendations for further study 79
5.6 Conclusion 79
6. References 81
Annexure 1: Interview Schedule 87
Annexure 2: Assessment criteria – ISO, DIN and TRAC 93
Annexure 3: Adapted OAIS-based model 102
Annexure 4: Synthesis of international repository standards 103
Annexure 5: Consent form 104

List of figures

Figure 1: OAIS framework indicating functional entities 11

Figure 2: Types of metadata schemas used by South African repositories 53

List of tables

Table 1: Comparison outline offered by the three different accreditation systems mentioned (ISO 16363, DIN 31644, TRAC) 25

Table 2: An overview of the compliance levels 71

Glossary

A	**Archival Information Package** – An information package that is used to transmit archival objects into a digital archival system, store the objects within the system, and transmit objects from the system. An AIP contains both metadata that describes the structure and content of an archived essence and the actual essence itself. It consists of multiple data files that hold either a logically or physically packaged entity. The implementation of AIP can vary from one archive to another archive; it specifies, however, a container that contains all the necessary information to allow long-term preservation and access to archival holdings (International Association of Sound and Audio-visual Archives, 2016).
C	**Certification** – According to the International Organization for Standardization (2015), certification refers to the "the provision by an independent body of written assurance (a certificate) that the product, service or system in question meets specific requirements". Certification can be conducted and granted by different organisations; however, a clear distinction is made in terms of the type of accreditation a digital repository obtains. Internationally recognised organisations will be regarded as fair and reliable accreditation providers, unlike the multiple fictitious formats that exist.
L	**Long-term preservation** – Refers to continued access to digital materials, or at least to the information contained in them, indefinitely (Digital Preservation Coalition, 2016).
M	**Medium-term preservation** – Continued access to digital materials beyond changes in technology for a defined period of time but not indefinitely (Digital Preservation Coalition, 2016). **Model** – WebFinance Inc. (2015) provides a comprehensive and scientific definition of the term 'model': "a model can come in many shapes, sizes, and styles; it is important to emphasize that a model is not the real world but merely a human construct to help us better understand real world systems". The hardship of assessing a digital repository to measure trust compliance requires a modification of different constructs. The developed model will encompass the most basic desires to statutory indicators that assimilate the concept of a trusted digital repository.

O	**Organisation** – The contribution of a digital repository needs to be understood in the context of an organisation, which is defined as "a system of consciously coordinated activities or efforts of social entities that are goal directed, deliberately structured activity systems with a permeable boundary" (WebFinance Inc., 2015). With the absence of working systems, the organisation will not be able to achieve objectives and goals. An organisation is an entity that exists in a technological domain that mainly relies on technology advancements to improve its information and knowledge management.
R	**Repository** – Bentley & Oladiram (2014) explain that a repository is a "digital research archive consisting of accessible collections of scholarly work that represent the intellectual capital of an institution". An institution utilises a repository to manage the "digital scholarship their communities produce, to maximise access to research outputs both before and after publication and also to increase the visibility and academic prestige of both the institution and authors" (Bentley & Oladiram, 2014). The institutions that were investigated in this research use their repositories for different organisational reasons; hence the need to evaluate them based on international repository assessment standards.
S	**Short-term preservation** – Access to digital materials either for a defined period of time while use is predicted but which does not extend beyond the foreseeable future and/or until it becomes inaccessible because of changes in technology (Digital Preservation Coalition, 2016).
T	**Trusted digital repository** – According to Smith (2009) and RLG-OCLC (2002), "a trusted digital repository is one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future"; a trusted digital repository has to be able to demonstrate fiscal responsibility and sustainability. A trusted digital repository should have practices, policies and performance that can be audited and measured (Dobratz et al., 2007). The ideal management of digital resources in a trusted digital repository is conducted in respect to adhering to particular standards. A trusted digital repository is commonly used for open access research outputs and regarded as an immediate and valuable complement to the existing scholarly publishing model (Crow, 2002) (as cited by Nicholas et al., 2013).

CHAPTER 1 INTRODUCTION

1.1 Overview

In South Africa, the National Research Foundation (NRF) "encourages its stakeholder community, including NRF's Business Units and National Research Facilities, to support public access to repositories through web search and retrieval according to international standards and best practice" (National Research Foundation, 2014). In order to establish an open-access publication culture, the NRF has a mandate of requiring stakeholders and the broader South African research community to promote and enable open access publication (also of data) from public funded research.

There are different international standards for the assessment of digital repositories for their trustiness. These standards provide an overarching compliance framework. Ross & McHugh (2005:1) explain that "controls must exist to protect and provide a guarantee for the authenticity and integrity of stored materials; accessibility must be maintained; and documentation, metadata, and assets must all be self-contained and maintained in-house or in other trusted repositories". Several reference frameworks that organisations can use to guide the development of trusted repositories exist. A unified certification framework was established in 2010. This framework is known as the European Framework for Audit and Certification of Digital Repositories. The collective union was led by David Giaretta in his capacity as Chair of the CCSDS (Consultative Committee for Space Data Systems)/ISO Repository Audit and Certification Working Group (RAC), Henk Harmsen in his capacity as Chair of the Data Seal of Approval (DSA) Board and Christian Keitel in his capacity as Chair of the DIN Working Group Trusted Archives – Certification (Higgins, 2015). The unified framework consists of a sequence of three certification levels, each providing for an increase in trustworthiness:

- Basic Certification is granted to repositories which obtain DSA certification. It is known as the basic 'bronze' level and comprises 16 criteria that may be self-assessed or peer reviewed (Higgins, 2015). The DSA comprises 16 criteria that are used for self-assessment, and that together determine whether the repository is fully compliant, partially compliant, or not compliant (Houghton,

2015). The peer-review approach involves receiving 'quality control' evaluation from an external entity using the 16 DSA criteria (Higgins, 2015).

- Extended Certification is granted to Basic Certification repositories which in addition perform a structured, externally reviewed and publicly available self-audit based on ISO 16363 or DIN 31644. The externally reviewed and publicly available self-audit requires an organisation to notify Nestor Seal (DIN 31644) and nominates two contact persons. After nominating, NESTOR confirms the start of the review and appoints a reviewer thereof; the archives conduct the self-assessment and submit it to the Nestor reviewer (Qasim, 2012). This type of self-assessment originates from NESTOR and comprises 34 criteria, representing the 'silver' level (Higgins, 2015).
- Formal Certification is granted to repositories which in addition to Basic Certification obtain full external audit and certification based on ISO 16363 or equivalent DIN 31644 (Higgins, 2015). Formal certification originates from Trustworthy Repositories Audit & Certification (TRAC) and comprises over 100 criteria/metrics. The gold' level or standard comprises an external audit performed using either the DIN or the ISO standard (Higgins, 2015). The external audit is an arms-length process that "reviews the repository, requiring evidence of compliance and testing to see that the repository is functioning as a Trusted Digital Repository" (Houghton, 2015).

The above-mentioned organisations are continuously making efforts to improve the assessment practice of repository assessment. In the fourth quarter of 2016, the ICSU World Data System (WDS) and the Data Seal of Approval (DSA) Board released the first version of their universal and unified *Core Trustworthy Data Repository Requirements (WDS Members Forum, 2016; WDS Annual Report, 2015–16). These new requirements are an addition to the above-mentioned levels of certification. These initiatives, done by international organisations, are efforts to create and mobilise useful and effective repository certification processes.

The importance of international accreditation using the above-mentioned standards has different reasons and advantages for a digital repository. Obtaining such certification, to a large extent, is associated with establishing trust for a digital

repository. Whyte et al. (2014) highlight key reasons to why international accreditation is important:

- Firstly, accreditation offers assurances for the peer review of data by using approved structures and procedures (Whyte et al. 2014).
- Secondly, accreditation based on international standards offers auditable checks to demonstrate that a repository performs technical reviews when data is deposited. It also verifies that there are appropriate standards for checking metadata completeness, and the authenticity and integrity of data is assessed (Whyte et al. 2014).

According to Down & Chan (2012:1), the needs of a user community must be served efficiently, "repositories with collections that serve diverse and interdisciplinary user communities will need to ensure that their collections can be used by the designated user communities". In order for this to be accomplished, digital repository custodians must have knowledgeable skills and expertise. This is a challenge in South Africa because digital repository managers and staff are still learning how to develop and improve their institutional repositories.

Another challenge that developing countries (i.e. South Africa) are faced with is "inadequate information and communication technology infrastructure – a major problem in this area is the high cost of internet bandwidth in various regions" (Christian, 2008:3). South African repositories are faced with the issue of sustainability when it comes to finances. The Consultative Committee for Space Data Systems (2011) explains that, "constant monitoring, planning, and maintenance, as well as conscious actions and strategy implementation will be required of repositories to carry out their mission". In South Africa (a developing country), it is a difficult task for an institution to source funds for the maintenance and sustainability of digital repositories. A challenging reality faced by organisations in developing countries (such as South Africa) is that they "continue to grapple with percentage decline in budgetary allocation. Funding therefore constitutes another major obstacle to the development of institutional repositories in a developing country's institutions" (Christian, 2008:3).

Organisations that utilise repositories require guidance that is aligned with best practices of international standards. Guidelines are an essential gateway tool that

enables digital repository managers to be cognitively aware of the administrative decisions a repository manager needs to implement to ensure that the repository is seen as a trusted digital repository.

This research intended to have a detailed look at international repository assessment standards and, based on that review, a model was developed that could be used by South African digital repositories to evaluate whether they are ready for compliance measurement. The research was guided by the research questions provided in the next section.

1.2 Research statement and research questions

In essence, digital data archiving and research data management have become increasingly important for institutions in South Africa (Koopman, 2016:1). A repository is an important aspect in managing and curating data so that the substantial investments in preparing and presenting the content and tools will not be lost (ICSU-World Data System, 2016). In this regard, certification is fundamental in guaranteeing the trustworthiness of digital repositories, and thus in sustaining the opportunities for long-term data sharing and corresponding services (ICSU-World Data System, 2016).

Trust assessment of a digital repository involves formal processes that are informed by international standards. In recent years, a number of certification standards and accreditation procedures have been developed worldwide: The Data Seal of Approval (DSA)1, the Network of Expertise in Long-term Storage and Accessibility of Digital Resources in Germany (NESTOR) seal/German Institute for Standardization (DIN) standard 316442, the Trustworthy Repositories Audit and Certification (TRAC) criteria/International Organization for Standardization (ISO) standard 163633, and the International Council for Science – World Data System (ICSU-WDS) certification of WDS Members (ICSU-World Data System, 2016). The objective of this research study was to develop a digital repository trust assessment readiness model that is based on international standards. Thus, the following research statement was formulated:

To develop a South African digital repository trust assessment model based upon international standards to measure trustiness of digital repositories.

In order to determine how South African digital repositories comply with international trust standards, the following research questions needed to be answered:

i. Which international repository assessment standards can be used to assess South African digital repositories?
ii. To what extent do SA research data repositories comply with this model in terms of trustiness?
iii. What model should be developed for SA as a developing country in striving for trustiness?

To effectively answer the above research questions, the research methodology explained in the next section was used.

1.3 Research methodology

The research study was primarily a qualitative study that used a case study design and a very structured interview schedule as the data collection instrument. The data collection methods, research paradigms and sampling methods are discussed in the next sections.

The research questions (refer to section 1.2) required the researcher to conduct a literature review of existing literature, and to have face-to-face interviews with research participants. The literature review consisted of secondary sources to help the researcher become familiar with the field of digital repository assessment; the types of assessment available for organisations and how international standards are used (refer to Chapter 2 of this document). The literature review serves as a good reference to the various international repository assessment standards, frameworks, and trends in digital repository assessment. Furthermore, a review of organisations that are interested in repository evaluation was conducted to gain a wider understanding of the dynamics involved in repository assessment.

Due to the time and nature element of the research study (mini-dissertation), only a small purposively selected sample was used and not an entire population. The case study was based upon feedback/input from digital repository managers from South African institutions that have implemented functional digital repositories.

The sampling that was used in this research was purposive sampling, which is a non-probability sampling technique (Michael, 2008). Non-probability sampling is one that does not "attempt to select a random sample from the population of interest, rather subjective methods are used to decide which elements are included in the sample" (Michael, 2008).

1.4 Limitations of the study

The following limitations were identified in the research study:

i. The research was limited to the Gauteng province.

ii. The sample was very small.

1.5 Justification for the research

The rationale for carrying out this research study lies in the unclear understanding of statutory requirements that digital repositories are assessed on for the purpose of establishing trust, especially in the South African context. This is due to the theoretically challenging nature of international repository assessment standards.

The objective of this research was to develop an assessment model that can be used as a starting point to measure trustiness of digital repositories. It is hoped that the developed model would improve developments and assessment of digital repositories to meet all the criteria of the international standards.

The purpose here is to investigate the possibility of using a developed model to indicate and outline trust requirements for digital repositories. Adopting criteria from international repository assessment standards will inform the development of a model that can be implemented for South African digital repositories to ensure work towards meeting full trust compliance is done. Furthermore, the developed model was cross-examined by means of interviewing a selection of digital repository managers. The model was adjusted and refined when such an adjustment contributed to better understanding of the model.

1.6 Value of study

Producing a theoretical framework regarding the assessment of digital repositories (in South Africa) highlights the need for further investigation into this aspect of assessment. The findings of this research study will yield a broader understanding of how repository managers understand the 'trust' aspect of their organisation's digital repositories. This understanding can be used to pave the way for further research concerning effective digital repository management practice. The study will contribute to developing a viable digital repository audit practice that is transparent and adheres to criteria of international standards in South Africa. The developed model and the findings thereof may assist data repository and IT managers to create best practice guidelines for their institutions.

Generally it is anticipated that the developed model will offer a realistic, viable framework to constitute trustworthiness within the context of digital repository challenges and opportunities for South African organisations.

1.7 Division of chapters

This document was subdivided into five chapters. Chapter 1 presents the introductory framework of the study. This chapter clearly explains the focus of the research that was undertaken, and comprehensive background information was given in this chapter to clarify the main reason for the study. Chapter 1 consolidates the main problem statements of the study, as explained by Bloom & Trice (2007) "the introduction of a research paper begins with a broader perspective of the problem and will become narrower as the introduction proceeds" – this is the foundation upon which Chapter 1 was based.

Chapter 2 presents an important aspect of the study, namely the literature review. Reviewing relevant literature provides an analytical framework that conceptualises the area of developing a South African-based assessment model. Relevant literature was linked to model theories and the latest developments of model concepts, which broadened the context of the background that was discussed in Chapter 1.

Chapter 3 informs the reader about the research approach and design, the collection of data, and how the data was controlled and analysed.

Chapter 4 is the chapter in which all the results of the study are presented. Predominantly this chapter presents the main findings and results of the data analysis.

Chapter 5 represents an overview of the study. The chapter provides detailed information about each section of the study. This chapter provides the most important findings, conclusions reached and the recommendations following the research findings. It also provides insight into further research that is needed.

1.8 Conclusion

The purpose of Chapter 1 was to provide some background and an overview of the study. The main research statement was provided, including the research questions. This chapter also briefly introduced the research methodology. In addition, the value of the study was mentioned, including the limitations and justification of the study. The next chapter discusses the literature reviewed for this research.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

Research offers an investigation arena that allows for objective scrutiny and assessment to be conducted into an area of interest. Kothari (2004) supports this notion by stating that, "the purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has not been discovered as yet" (Kothari, 2004:2). The purpose of this research is to develop a working framework/model that could be used to establish a holistic understanding of what 'trust' means in terms of research data repositories for organisations in South Africa. It is envisioned that the developed model will become a hands-on tool that South African organisations can use to assess, reflect and know where change is needed for their research data repository to attain 'trust' status.

The process of appraising and certifying a research data repository for its trustiness has been debated since the advent of repositories as information management platforms (Qasim, 2012). The chapter begins by focusing on the Open Archival Information System (OAIS) functional model, and the importance of each OAIS functional entity is discussed. This is followed by a discussion on organisations that are interested in repository assessment and how these organisations operate to elucidate a broader understanding of repository assessment, after which digital repository assessment will be discussed. This chapter concludes with a section discussing the development of an assessment model (based on the literature discussed) that can be used to assess South African repositories for trust status. The literature review in this study was conducted based on the research questions (presented in section 1.2). The theoretical background of repository assessment using international standards informed the research questions, which guided the how the literature review was done.

Various international organisations have developed (in the form of standards) tools to assess if repositories comply with the set standards. The Consultative Committee for Space Data System (CCSDS) is one such organisation. The CCSDS have "defined recommended practice upon which to base an audit and certification

process for assessing the trustworthiness of digital repositories" (The Consultative Committee for Space Data Systems, 2011:15). The main focus of this research is to state what is meant by 'trust' status for South African research data repositories. It is thus important to understand what the definition of 'trust' is in the context of a research data repository.

According to Ambacher (2007:1), "claims of trustworthiness and trust are easy to make but have thus far been difficult to justify or objectively prove". To address the issue of asserting trust status for a research data repository, standards are used for a systematic and formal evaluation. These trust evaluation standards will be discussed in section 2.3.1. The discussion will form part of a synthesis (based upon the OAIS framework) that was used to develop a comprehensive overview of the trust evaluation systems against which a South African research data repository trust assessment model could be developed.

Wikipedia (2016) provides a very good overview for the novice to gain some understanding of the framework. From this piece it is possible to report that the OAIS framework model is platform agnostic (it does not require any specific application/ software system). Section 2.2 provides a more detailed explanation of the OAIS reference model and its various components.

2.2 The Open Archival Information System (OAIS) Reference Model

The development of the OAIS Reference Model was led by the Consultative Committee for Space Data Systems (CCSDS) (Day, 2007). An OAIS is a concept that is determined by a variety of different terms. Day (2007) explains that the term 'open' means that "documents are developed in an open way, and does not imply that access to any OAIS should be unrestricted". This statement relates to the open nature of information stored in an OAIS, which requires long-term preservation (please refer to the glossary for a definition of long-term preservation). In the context of OAIS, the term 'archive' is used to define "an organization that intends to preserve information for access and use by a designated community" (Day, 2007). In essence, an OAIS is understood to mean "any organization or system charged with the task of preserving information over the long term and making it accessible to the designated community" (Lavoie, 2004).

The OAIS Reference Model is not a repository evaluation model; it is a framework that provides common definitions of terms and means of comparison (Day, 2007). To help facilitate the preservation of information for the long term, the OAIS framework has six entities that contribute to the framework's mission. The entities of the OAIS framework are represented in Figure 1 below and discussed in the following section.

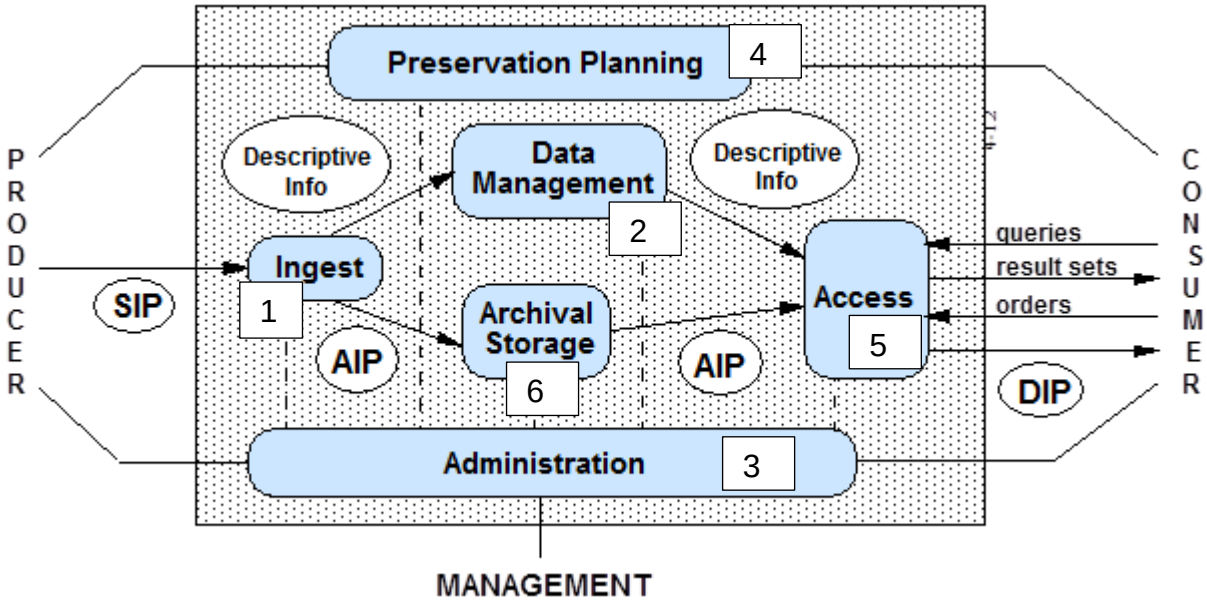


Figure 1: OAIS framework indicating functional entities

1) Ingest Functional Entity

The ingest entity of the OAIS framework focuses on the management of Submission Information Packages (SIPs). To understand what an SIP is, Day (2007) provides a systematic explanation of this concept: "SIP is an information package that is delivered to the repository and digital storage system for ingest – SIPs include the data to be stored and all the necessary related metadata about the object and its content". The ingest entity of the OAIS framework is the process of accepting content and all its related metadata (Day, 2007). Once content and related metadata is accepted, the ingest entity is responsible for providing the services and functions to accept "SIPs from relevant producers and prepare the contents for storage and management within the archive" (Sawyer. 2002:43).

2) Data Management Functional Entity

There are different technical factors that the OAIS framework regards as being important for effective data management. These factors include "administering the archive database functions, performing database updates, performing queries on the data management data to generate result sets, and producing reports from these result sets" (Consultative Committee for Space Data Systems, 2012). The assessment of these factors facilitates a consistent repository data management framework that is measurable and controlled. The data management entity of the OAIS framework "supports search and retrieval of archived content through the use of descriptive metadata" (Michael, 2016).

3) Administration Functional Entity

This OAIS entity "provides the services and functions for the overall operation of the archive system, which includes soliciting and negotiating submission agreements with producers, and auditing submissions to ensure that they meet archive standards" (Consultative Committee for Space Data Systems, 2012). Best practices are a crucial component in the management of a digital repository. These best practices include infrastructure requirements, human capacity, and maintaining the services and functions for the overall operation of a system (Consultative Committee for Space Data Systems, 2012). The administration entity of the OAIS framework manages the day-to- day operation of the archive by being the interface between the archive and two components of the OAIS environment: Management and the designated community (Lavoie, 2015).

4) Preservation Planning Functional Entity

Computing environments are inclined to experience changes and ultimately become obsolete with time. This entity of the OAIS framework "provides the services and functions for monitoring the environment of the OAIS and provides recommendations to ensure that the information remains accessible to the designated user community over the long term, even if the original computing environment becomes obsolete" (CCSDS, 2012). To a large extent, preservation planning requires an institution to regularly monitor risks "that could prevent them from properly preserving and maintaining access to digital objects" (Michael, 2016). This entity ensures that an

archive has updated policies and procedures in place to keep up with any technology or environment changes in order to mitigate risk (Michael, 2016).

5) Access Functional Entity

The OAIS framework puts emphasis on communication abilities to improve the manner in which a system "supports consumers in determining the existence, description, location and availability of information stored in the OAIS, and allows consumers to request and receive information products" (Consultative Committee for Space Data Systems, 2012). This entity enables the OAIS to adequately receive requests, apply controls to limit access to specially protected information, and coordinating the execution of requests to successful completion. Sawyer (2002:43) further explains that the OAIS access entity "supports consumers in determining the existence, description, location and availability of information stored in the OAIS".

6) Archival Storage Functional Entity

A key component of the OAIS framework is the Archival Information Packages (AIPs) (please refer to the Glossary for a definition of AIP). The Archival Storage Functional Entity incorporates a faculty that "provides the services and functions for the storage, maintenance and retrieval of AIPs" (Consultative Committee for Space Data Systems, 2012). This entity allows the OAIS to manage AIPs hierarchically from the point of ingest until AIPs are provided access. The responsibilities that are executed by the archival functional entity include receiving new AIPs from the Ingest function and assigning them to permanent storage according to various criteria (media requirements, expected utilisation rates, etc.); migrating AIPs to new media as required; error checking; implementing disaster recovery strategies; and providing copies of requested AIPs to the Access function (Lavoie, 2015).

Once OAIS compliance has been introduced, it does not automatically ensure sustainability, and therefore assessment systems were developed. Assessment systems aim to make it clear to what extent the system is in fact doing the 'right things' (ingest, data management, administration preservation, access and archival storage) (Consultative Committee for Space Data Systems, 2012). The OAIS framework provides context that is conducive to designing a sustainable environment for any collection of objects. At the very least, the OAIS framework "is laying

important foundations for a coordinated and widely applicable solution to the challenges of digital preservation" (Lavoie, 2015). The OAIS framework played a key role in the development of the proposed South African digital repository assessment model in that the entities are the foundational pillars of the proposed model. The assessment criteria of various international repository standards are derived from the OAIS framework (Consultative Committee for Space Data Systems, 2012). The OAIS framework informed this study to ensure that the main entities that are regarded when assessing a digital repository also form part of the proposed South African repository trust assessment model. Unlike other models the OAIS framework offers a fundamental context to base and align repository assessment practice.

Various organisations are interested in repository evaluation for trust agreement. These standards do not make exceptions in relation to compliance and the geographic region of an institution (Sawyer, 2002). The following section provides a discussion on some organisations that are interested in repository evaluation.

2.3 Organisations interested in repository evaluation

At least two associations are very interested in finding a solution that is inclusive for all repository owners. These associations/organisations are the World Data System (WDS) and the Research Data Alliance (RDA). Both of these are described in more detail below. Other organisations that are also included below is the Research Libraries Group (RLG) and the Online Computer Library Center (OCLC). The reason for including these organisations is that they offer a retrospective overview of how various organisations agreed on standards, criteria and mechanisms needed to certify repositories of digital information as archives (RLG-OCLC, 2002).

2.3.1 World Data System (WDS)

As part of combined efforts, different organisations have come together to develop processes that would improve certification procedures. One such group is the World Data System (WDS), which has been created "by a decision of the 29th General Assembly of the International Council for Science (ICSU)" (Creative Commons Attribution, 2015). The main aim and concept of this group is to build a world data certification membership approach that caters for organisations at a global scale,

and to enhance the "potential offered by advanced interconnections between data management components for disciplinary and multidisciplinary applications" (World Data System, 2012). The background that is used by the WDS is based on the key principles of transparency and openness whereby an organisation's digital repository is measured by "striking the right balance between simplicity and robustness of the work and effort involved" (Genova, 2015).

The certification and accreditation procedure offered by WDS is provided to interested participants (organisations) for one or multiple roles. The roles may include "data collection and processing (including quality assurance), long-term data repository (e.g. data library), data publisher (including periodic compilation of data products), community related service, data analysis service" (World Data System, 2016). The criteria that WDS uses to evaluate organisational roles take into consideration the context of the institution, its mission, priorities, and stated commitments (World Data System, 2016); making the WDS certification process an iterative one.

2.3.2 The Research Data Alliance (RDA)

The World Data System is not the only organisation that has a mandate to facilitate repository audit certification. The Research Data Alliance (RDA) has a mission of building "social and technical bridges that enable data sharing through the creation, adoption and use of the social, organizational, and technical infrastructure needed to reduce barriers to data sharing and exchange" (Wood, 2015:3). In relation to trust certification of research data repositories, RDA has acknowledged that when an institute considers using worldwide certification standards (i.e. DSA, ISO), a level of rigidness occurs. This is because "the primary focus of the Data Seal of Approval (DSA) has been on digital repositories in the Humanities and Social Sciences, while the focus of ICSU-WDS has been on Earth and Space Sciences for historical reasons" (Wood, 2015). RDA offers a more proactive approach of certification assessment; it operates on an efficient and simplified assessment approach of using a "harmonized catalogue of criteria for basic certification of repositories drawn from the DSA and WDS requirements, as well as a set of common procedures for repository assessment" (Wood, 2015).

The Research Data Alliance (2016) explains that the core functionality of RDA is that "it allows members to come together through self-informed, volunteer-focused Working Groups, exploratory Interest Groups to exchange knowledge, share discoveries, discuss barriers and potential solutions, explore and define policies and test including harmonizing standards to enhance and facilitate global data sharing".

Hanahoe (2014) provides some examples and theories that the RDA is based on, which include:

- *Persistent identifiers are regarded as the core of proper data management and access.*
- *Designing and implementing an API (a set of functions and procedures that allow the creation of applications which access the features or data of an operating system, application, or other service) for interaction with typed information.*
- *Automated data management across disciplines and repositories benefit from standardized types.*

RDA aims to reduce the certification barriers that organisations experience when solely using DSA or WDS independently.

2.3.3 Research Libraries Group (RLG) and Online Computer Library Center (OCLC)

The Research Libraries Group (RLG) and Online Computer Library Center (OCLC) digital repository have identified additional attributes that are regarded as important. The RLG-OCLC collaboration was created "to establish attributes of a digital repository for research organizations, building on and incorporating the emerging international standard of the reference model for an OAIS" (RLG & OCLC, 2002). The RLG and OCLC attributes affirm the requirements that are needed for a digital repository to receive 'trust' accreditation and to a large extent is based on the OAIS Framework.
The RLG & OCLC (2016) have defined criteria that assist a digital repository to improve the management of its digital assets. These include:

- *Define the characteristics of reliable archiving services for heterogeneous research collections in a document to be made available electronically to the community at large.*

- *Produce a rational set of criteria for archives that can hold the full range of digital collections and datasets (including both "born digital" and "born-again digital" information) requiring long-term storage and access systems.*
- *Identify tools that support research institutions as they seek either to build their own archiving capacity or contract with third-party services for archiving functions* (RLG & OCLC, 2002).

2.3.4 The International Council for Science – World Data System (ICSU-WDS)

The practice and field of data repository assessment and certification can become a complex one that requires interoperability and knowledge from different backgrounds. The International Council for Science – World Data System (ICSU-WDS) is one such entity because it "brings together much expertise and experience in the area of certification over a broad range of disciplines and with a global reach" (Kowalczyk & Shankar, 2013:249). The collaboration of these two organisations broadens the scale of data repository certification by combining critical "mass of stakeholders to offer a certification service that is of high quality, efficiency and agility" (Kowalczyk & Shankar, 2013:249). This level of collaboration aims to deliver research data repository accreditation that is of high quality in an age where information and knowledge management tools are fast growing.

As part of developing a model to assess the level of trust a research data repositories adheres to, the WDS outlines specific areas that, if they form part of the model, will aid in addressing issues of "policies, organizational framework, network framework, management of data, metadata, and services, and technical infrastructure – where appropriate" (World Data System, 2012). WDS does not only provide relevant assessment criteria, but also offers a monitoring framework that ensures that proper mechanisms are implemented "to monitor the overall performance of the system as well as the performance of member facilities" (World Data System, 2012). The WDS certification "offers a basic certification standard for trusted digital repositories" (Genova, 2015).

The WDS certification has in place different themes required to establish basic trust certification; important areas in this regard (on face value) are technical infrastructure, policies and organisational framework.

The WDS offers basic certification that is only viable to a certain level and not in accordance with international standards.

The above discussion demonstrated a single view of how a digital repository can be assessed. The focus of the developments was definitely Northern hemisphere/developed country and it is therefore already clear that the developing country perspective would need to be investigated. The following section discusses formal digital repository assessment, including the use of international repository assessment standards.

2.4 Digital repository assessment

Assessment of a data repository puts an organisation (specifically repository managers) in a better position to identify non-compliant areas that need to be addressed. In this regard, the ability to conduct relevant corrections requires having a measuring indicator that will assist a research data repository manager to understand the level of compliance being targeted.

Qasim (2012) is of the impression that "an organisation should conduct self-reflection assessments to meet requirements set out by the criteria of international standards". In order to comply with these standards, it is critical for an organisation to know what these standards entail.

There are various international standards – such as the International Organization for Standardization (ISO) 16363 and the 'Deutsches Institut für Normung' (DIN) 31644. These standards have a list of criteria that have to be complied to. For novices it will in all probability be very difficult to comply with the criteria of these standards. In order to reduce the level of difficulty of complying with international standards, certain frameworks have been developed to facilitate the process.

The next section discusses the international standards that were considered for the purpose of this study, and what the standards entail in terms of digital repository trust assessment.

2.4.1 Trustworthy Repositories Audit & Certification (TRAC)

The Trustworthy Repositories Audit & Certification (TRAC) defines clear criteria that, if used consistently, can certify a research data repository with trust status. The

TRAC criteria originate from the OAIS Framework, but deliver a comprehensive audit framework, to assess functionality, and ultimately aid potential certification (Ross & McHugh, 2006).

TRAC uses 19 criteria and 3 main groups that relate to basic concepts of repository assessment (see Annexure 2 for the full TRAC criteria). The groups include:

- Organisational infrastructure
- Digital object management
- Technologies, technical infrastructure, and security.

The TRAC criteria that are used to assess the capabilities of a research data repository to meet trust status, include technical and administration assessment. This accreditation can be given to different institutions such as "academic institutional preservation repositories to large data archives and from national libraries to third-party digital archiving services" (Hitchcock & Donnelly 2010).

2.4.2 The International Organization for Standardization (ISO) 16363

According to the International Organization for Standardization (ISO) 16363 (2012), the ISO 16363 standard "demonstrates a reliable framework that is based on trustworthiness and responsible data management and stewardship" (Consultative Committee for Space Data Systems, 2012). This standard can be used by an organisation to facilitate trust assessment procedures based on international practice.

The main characteristics of the standard are the depth and rigorousness of the criteria against which to measure. The ISO 16363 standard does not only provide directions, but essentially incorporates criteria that are mandatory for adherence. The ISO 16363 standard used TRAC as the point of departure, but added further criteria – in total more than 100 criteria to be evaluated. These criteria aim "to be relevant to all kinds of repositories, including those for commercial and cultural heritage, as well as scientific purposes" (Callaghan et al., 2014:155). This combination between ISO 16363 and TRAC creates a holistic framework that will undoubtedly ensure trust status. However, it is anticipated that the comprehensiveness and strictness of the ISO 16363 standard will create a stumbling block for South African repositories. This is because "the criteria outlined in the ISO

16363 standard are difficult to comply with – especially for under-resourced organizations" (Downs & Chen, 2013).

2.4.3 The 'Deutsches Institut für Normung' (DIN) 31644

The DIN 31644 working group has developed a practical and composed repository assessment that is based on 34 core requirements (UK Data Archive, 2016) (see Annexure 2 for the full TRAC criteria). The 34 requirements were developed to cater for institutions that have an academic and preservation mandate, such as libraries and museums.

The 34 requirements are organised into 3 categories:

- Organisation
- Management of intellectual entities and their representations
- Infrastructure and security (UK Data Archive, 2016).

As compared to the ISO 16363 and TRAC standards, the DIN 31644 uses a condensed criteria assessment framework. The process of using the DIN 31644 standard involves an institution conducting extended self-assessment to verify the level of trust that a research data repository complies with. Doorn (2014:19) explains that, "if the reviewed assessment yields a positive result they are entitled to publicize this by using the Nestor Seal for Trustworthy Digital Archives". The Nestor Seal for Trustworthy Digital Archives is "a self-assessment process for digital archives developed and offered by Nestor on the basis of the DIN 31644 standard" (Nestor, 2013).

2.4.4 European Framework for Audit and Certification of Digital Repositories

The European Framework for Audit and Certification of Digital Repositories is an Audit and Certification Working Group that was founded in 2010 by the Consultative Committee for Space Data System (CCSDS) and the DIN Working Group. The establishment of the Framework is intended to provide certification for digital repositories based on auditing and certifying mechanisms which are based on "a tiered approach to certification, allowing an entry-level self-assessment and peer review based on the Data Seal of Approval, a more extensive self- assessment

(based on DIN 31644 or ISO 16363), and a full scale external audit based on ISO 16363" (University of Glasgow, 2016).

It has already been acknowledged that it is necessary to make provision for different levels of certification when evaluating a research data repository. To comprehend these different levels, a Memorandum of Understanding was signed in the year 2010 and came to be known as the European Framework for Audit and Certification of Digital Repositories. This framework offers three different levels (or tiers) of certification, which are categorised in the following manner:

- Basic Certification – is granted by obtaining the Data Seal of Approval (DSA) (Schumann, 2012:24), which an organisation can obtain by complying with a total of 16 criteria guidelines that are focused on "data producers, repositories, and users" (Qasim, 2012:5). To fully obtain basic certification, an organisation has to conduct a self-assessment using the DSA guidelines.
- Extended Certification – this certification is a continuation process whereby a repository that has basic DSA certification "can obtain an extended certification by getting an externally reviewed self-assessment based on either ISO 16363 or DIN 31644" (Qasim, 2012:5).
- Formal Certification – once a repository has earned a DSA certification, a "full external audit is done in accordance with either the International Organization for Standardization (ISO) 16363 or Deutsches Institut für Normung (DIN) 31644 standards" (Qasim, 2012:5), and the successful results thereof awards the repository with 'Formal Certification'.

The certification offered by the European Framework for Audit and Certification of Digital Repositories presents a unique approach of certification (TRAC, 2007). This certification is not necessarily based on the broader meaning of what 'trust' entails according to international standards. The following section discusses what trust entails according to international standards.

2.4.5 Brief comparison of the various international standards

The different international standards have their own attributes and purposes, which make them differ from one another. The first apparent difference is the affiliation aspect that is required. Kowalczyk & Shankar (2013) explain that, "data services

consist of a number of components (data centres, analysis centres, product centres, etc.) and have their own organisational structure (e.g. central bureau or governing board)". A digital repository can be used as an agent to facilitate these services. The International Council for Science – World Data System (ICSU-WDS) "has a membership focus, and includes not only data centres (mainly repositories) but also data services" (Kowalczyk & Shankar, 2013), which is different to the other standards.

The European Framework for Audit and Certification of Digital Repositories is a standard that is not primarily focused on organisational association, but on a tiered approach (taken from the Data Seal of Approval)"based on the size, objectives, and available resources" (Qasim 2012:5). The European Framework for Audit and Certification of Digital Repositories is based on three types of data repository certification, namely basic certification, extended certification and formal certification. The ISO 16363 and DIN 31644 standards do not use the approach of adopting a criterion and methodology from existing standards; instead the two standards defined their own criteria that are unique and accepted internationally for the purpose of accrediting a research data repository with 'trust' status. The OAIS Framework is aimed at providing "a framework, including terminology and concepts, for describing and comparing architectures and operations of existing and future archives" (Consultative Committee for Space Data Systems, 2011). The OAIS Framework is different from other assessment standards because it compares by describing the various long-term preservation techniques and strategies – an issue that other standards do not comprehensively address.

After having evaluated all the possible options, it was decided to use the OAIS Framework to serve as the foundation for developing the proposed trust assessment model for South Africa. A central OAIS point of reference, as the basis for using the OAIS Framework as the foundation for the developed model, is that every component of the framework is used to integrate the criteria of international standards for evaluating the trust status of a research data repository. It is important to explain the function of each entity of the OAIS Framework to demonstrate how the integration of international standards will facilitate trust assessment.

The above discussions of organisations that are interested in repository assessment (including digital repository assessment based on international standards) provided an overview of what repository assessment entails. The starting point of this section was a discussion about the OAIS Framework. The following section discusses how a proposed model was developed to assess South African digital repositories for trust status. This model assesses a repository up to a certain level and not the full requirements as required by the criteria (see Annexure 2 for the full criteria of each international standard) of international standards.

2.5 Development of the proposed model

The previous section explained the different international repository assessment standards. Based on the comprehensive nature of each international standard, the proposed model to assess South African digital repositories is going to be developed by synthesising the criteria of the ISO 16363, DIN 31644, TRAC and RLG-OCLC into the OAIS Framework. Each OAIS entity and what it entails has been explained in the previous section. The following section describes how the OAIS entities are synthesised with international standards to form various infrastructures that will assess a South African research data repository.

The five discussed OAIS Framework entities are the foundation pillars of the proposed model to assess South African digital repositories for trust status. As mentioned in the previous section, the five entities have been used to accommodate the different metrics of international standards (ISO 16363, DIN 31644, TRAC, RLG-OCLC) to form the model that will be used in the South African context. From the studied literature, it was discovered that the OAIS Framework does not cover aspects of system security and financial sustainability. A decision was made to adapt the OAIS Framework, but to then include these two aspects. The reasons to include these two aspects originate from the RLG-OCLC explanation as to why they are important (see Annexure 3 for the adapted OAIS model).

Taking into consideration the different trust assessment standards (i.e. TRAC, ISO and DIN), it becomes clear that developing an evaluation model that will meet criteria requirements and address the level in which compliance can be achieved of research digital repositories in Low to Middle Income Countries (LMIC) is not an easy task to complete. Many factors have to be considered for the task to be viable

and attainable. The attempt to aid the construction of a model to assess the level of trust a digital repositories adheres to, has to be comprehended in respect to different theoretical frameworks that are available for consideration (Day, 2007). Adopting these theories requires a critical and specialised focus on which areas are most important and to what degree. The following table provides a comparison outline offered by the three different accreditation systems mentioned (ISO 16363, DIN 31644, TRAC).

Table 1: Comparison outline offered by the three different accreditation systems mentioned (ISO 16363, DIN 31644, TRAC)

OAIS Category	Explanation	ISO 16363	DIN 31644	TRAC
Administration entity/infrastructure	The repository needs to have sufficient numbers of appropriately qualified staff and updated job descriptions need to exist that set out the required qualifications of the digital archive personnel and contain an organisational chart and/or a staff development plan based on the tasks and objectives of the digital archive (Nestor, 2013).		**C9 Personnel** - Does the organisation have sufficient and qualified staff members available to manage the repository? - Are the required qualifications set out, including an organisational chart? - Is there a staff development plan that outlines the tasks and objectives of the repository?	
	Obtains technical authority over the representations being ingested, allowing it to transform them into archival information packages and, if necessary, to carry out long-term preservation measures (Nestor, 2013).		**C20 Technical Authority** - Does the repository have processes in place that ensure authority control on a permanent basis without technical restrictions (e.g. encryption, copy and print protection)?	
	The repository shall track and manage intellectual property rights and restrictions on use of repository content as required by deposit agreement, contract, or license (Downs & Chen, 2013).	**Contracts, Licenses, and Liabilities** - Does the repository have and maintain appropriate contracts or deposit agreements for digital materials that it manages, preserves, and/or to which it provides access? - Are formal deposits and contracts legitimate, i.e. are they countersigned and current?		

OAIS Category	Explanation	ISO 16363	DIN 31644	TRAC
	The type of digital information for which the digital archive is responsible must be clear both internally and externally (Nestor, 2013).		**C1 Selection of information objects and their representations** - Does the repository have a criterion that defines the selection of information objects and their representations?	
	Criteria have been laid down for selecting the information objects and their Representations (Nestor, 2013).		**C1 Selection of information objects and their representations** - Is the selection of digital information transparently documented on the basis of criteria, guidelines and profiles?	
	To what extent has the digital archive ensured that the information objects are preserved even after the archive itself has ceased to exist.		**C21 Submission information packages** - Does the repository adequately specify the composition of data packages for data transfer?	
	The institution has to have plans in place in the event of a crisis (Keitel, 2014). This is necessary to ensure business continuity.		**C21 Submission information packages** - Which specifications does the digital archive have regarding content data that is accepted, and the metadata required?	
	A repository shall have access to the necessary tools and resources to provide authoritative representation Information for all of the digital objects it contains (Day, 2007).	**Ingest: Acquisition of content** - Can the repository demonstrate (to funders, depositors, and users) what responsibilities it is taking on and what aspects are excluded? - Can the repository determine and check what the characteristics and properties of preserved items will be over the long term?		

OAIS Category	Explanation	ISO 16363	DIN 31644	TRAC
Data management infrastructure	A repository needs to specify minimum information requirements to enable the designated community to discover and identify material of interest (Day, 2007).	**Information Management: Specify minimum information requirements to discover and identify material of interest** - Is the repository able to deal with the different types of information requests made by users from the designated community? - Does the repository have adequate retrieval and descriptive information, discovery metadata (such as Dublin Core), and other documentation describing information objects to be retrieved?		
	The repository should be able to produce minimum descriptive information that was either received from the producer or created by the repository (Day, 2007).	**Information Management: Capture or create minimum descriptive information associated with the AIP** - Can the repository deal with the types of requests that come from a typical user from the designated community?		
	Repositories must implement procedures to establish and maintain relationships to associate descriptive information for each AIP, and should ensure that every AIP has some descriptive information associated with it and that all descriptive information must point to at least one AIP (Book, 2012).	**Information Management: Maintain bi-directional linkage between each AIP and its descriptive information** - Can all the Archival Information Packages (AIPs) be located and retrieved? Are there adequate measures (such as descriptive metadata; unique, persistent identifier) in place to ensure that AIPs are located and retrieved? - Is there a procedure in place that notifies when the relationship between the data and the associated descriptive information is temporarily broken to ensure that it can be restored?		

OAIS Category	Explanation	ISO 16363	DIN 31644	TRAC
	A repository needs to monitor the physical and legal control over the existence, authenticity, location, and accessibility of records (Nicholson & Dobreva, 2009).			**Chain of Custody: Have physical and legal control over the existence, authenticity, location, and accessibility of records** - Can the repository demonstrate the chain of custody for all of its digital content from the point of deposit?
	It is crucial to know the chain of custody for digital content that resides within a research data repository. This ensures a track record that outlines the timescale from the point of deposit (TRAC, 2007).			**Chain of Custody: Have physical and legal control over the existence, authenticity, location, and accessibility of records** - Can the repository demonstrate that the content it has matches the content it received?
Metadata management infrastructure	The repository needs to have "basic preconditions for appropriate use now and in the future, including the interpretability of both content data and metadata" (Nestor, 2013).		**C5 Interpretability** - Does the repository have a metadata schema that it uses? - Can the repository ensure long-term interpretability of at least one representation of content data and metadata? - Does the repository have methods to allow the user community to check interpretability on a regular basis?	

OAIS Category	Explanation	ISO 16363	DIN 31644	TRAC
Access infrastructure	"The repository shall comply with Access Policies" (Downs & Chen, 2013) – and "the repository shall follow policies and procedures that enable the dissemination of digital objects that are traceable to the originals, with evidence supporting their authenticity.	**Access Management: Comply with Access Policy** - Is the repository able to produce evidence to demonstrate that it has fully addressed all aspects of usage which might affect the trustworthiness of the repository?		
		Access Management: Dissemination of original digital objects with evidence - Does the repository follow policies and procedures that enable the dissemination of digital objects that are traceable to the originals, with evidence supporting their authenticity?		
	A repository needs to ensure that authorised users in the designated communities can access the representations" (Nestor, 2013). A repository will have to declare "its conditions of use and any costs which may arise, listing these in a transparent manner" (Nestor, 2013).		**C4 Access** - Can the repository ensure authorized access to information for the designated community? - Does the repository have appropriate search possibilities, which indicate the terms of use and restriction? - Does the repository declare its conditions of use and costs that may arise?	

OAIS Category	Explanation	ISO 16363	DIN 31644	TRAC
Preservation infrastructure	Preservation Period Determination Phase: Nominate institution to offer long-term preservation - Can the repository offer long-term preservation service? If answer is 'yes', ask respondent to elaborate.			
	The institution has to have "a mission statement that reflects a commitment to the preservation of, long-term retention of, management of, and access to digital information" (Downs, & Chen, 2013).	**Governance and Organizational Viability** - Does the parent organisation or the repository's mission statement explicitly address preservation?		
	A repository needs to "assume responsibility for the long-term preservation of the information objects on the basis of legal requirements or its own objectives" (Li & Banach, 2011).		**C18: Authenticity - Preservation measures** - Does the repository deploy methods that ensure the authenticity of the objects during implementation of the long-term preservation measures and document the degree of authenticity? - Does the repository ensure that relevant information objects retain their authenticity while undergoing preservation processes and that all measures are transparently and permanently documented?	
	The repository needs to be able to assert provenance and authenticity of a digital object - guaranteeing that the object is stored intact as it had been created (TRAC, 2007).			**Authenticity** - As part of preservation, does the repository assert provenance of digital objects by guaranteeing that objects ae stored intact as it has been created?

OAIS Category	Explanation	ISO 16363	DIN 31644	TRAC
Financial sustainability infrastructure	The business planning correlate with organisational objectives that a repository facilitates the successful achievement of it. In addition to business planning processes the ISO 16363 Financial Sustainability criteria requires A research data repository needs to have financial practices and procedures which are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements (Downs, & Chen, 2013).	**Financial Sustainability** **-** Does the repository have a short and long-term business planning process in place to sustain the repository over time? (An annual business planning process is commonly accepted as the standard for most Organizations).		
System security sustainability infrastructure	This criteria requires that a repository "must implement suitable measures to protect its own integrity and that of its archive assets to ensure that the assets remain intact and to fulfil its legal or contractual obligations" (Nestor, 2013). This means that an institution has to identity components of the repository that require serious protection and "an analysis of any potential threat to the specific archive and a risk assessment of the damage scenarios" (Nestor, 2013).		**C34 Security** **-** Can the organisation and the infrastructure protect the repository and its information objects and representations to ensure integrity? - Does the repository have suitable measures to protect its own integrity and that of its digital assets to ensure that the assets remain intact and to fulfil its legal or contractual obligations?	
	This criterion outlines a comprehensive framework that ensures that a repository operates within a secure environment to ensure that digital material are accessible and usable over time (Downs, & Chen, 2013	**Security Risk Management** - Does the repository conduct regular risk assessments and maintain adequate security protection in order to provide expected and contracted levels of service? - Can the repository show how it deals with its security requirements: If some digital materials pose a higher vulnerability of being attacked the repository will need to provide more protection?		

The above table is a comparison of what the different standards address in relation to the OAIS Framework. The table above also represents an adapted OAIS Framework that forms the basis of this study. The adapted model includes the Financial and Security sustainability aspects as explained by the Research Libraries Group (RLG) and Online Computer Library Center (OCLC). Adding to section 2.3.3 of this chapter, the RLG-OCLC introduced a collaboration" to establish attributes of a digital repository for research organizations, building on and incorporating the emerging international standard of the Reference Model for an Open Archival Information System (OAIS)" (RLG & OCLC, 2002). The RLG and OCLC attributes affirm the requirements that are needed for a digital repository to receive 'trust' accreditation. Building on the OAIS components discussed in section 2.2, the 'Financial Sustainability' and 'System Security' attributes are included in the OAIS Framework. The section below discusses the two RLG and OCLC aspects.

Financial Sustainability

Substantial existence of a digital repository requires financial consideration to sustain the repository over time – The RLG and OCLC (2002) report explains that "trusted repositories will adhere to all good business practices and should have a sustainable business plan in place". A repository's financial outlook needs to be reviewed at least once a year (RLG and OCLC, 2002). This review helps repository managers to understand operating budgets to be able to produce a balance of risks benefit, investment, and expenditure (RLG and OCLC, 2002). Financial sustainability in the proposed model ensures that an inclusion and viable approach of administering every aspect of a repository is maintained.

System Security

Security is an important dimension in managing digital resources; if a repository does not have security measures in place, the repository faces risks associated with unauthorised access and usage, damage and losing valuable information. The RLG and OCLC system security attribute ensures that "all systems used in the operation of a trusted digital repository will be designed to assure the security of the digital assets" (RLG and OCLC, 2002). In order to guarantee reliable and consistent system security, a repository needs to have policies and practices to ensure that community needs are satisfied, "particularly those pertaining to copying processes, required

redundancy of data, authentication systems, firewalls, and backup systems" (RLG and OCLC, 2002). The policy content needs to have clear statements about issues pertaining to which steps to follow in the event of disaster preparedness and recovery, and staff will be trained appropriately (RLG and OCLC, 2002).

See Annexure 3 for the adapted OAIS model to assess South African digital repositories for trust status.

2.6 Summary

The evaluation of a digital repository for trust accreditation is a formal process that involves different phases. Institutions who want their digital repositories evaluated have to use international repository assessment standards (such as ISO, TRAC). These standards provide an informed framework that assesses a digital repository by systematically covering important aspects. However, taking the developing status of South Africa, organisations find it challenging to fully comply with these international standards. The available literature shows the vastness and formality of international repository assessment standards.

The proposed OAIS-based model presented in this chapter aims to address this dilemma that South African institutions face. The adapted OAIS-based model intends to further understand the shortfalls that South African organisations face in relation to complying with international assessment standards. Based on the feedback that organisations provide, recommendations are welcomed of how the model can be redesigned.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

The purpose of Chapter 3 is to explain the type of research design and methodology used to develop a research data repository trust assessment model for South African organisations. The chapter starts by providing an overview and the goal of the study, and is then followed by discussing the research design selected for the study, including the data collection methods and sampling deployed for data collection. The chapter concludes by explaining how data analysis and interpretation were done on the collected data.

3.2 Research methodology and design

In order to conduct a systematic and formal research project, it is important to firstly comprehend the methodology and design that the research is based upon. Burns and Grove (2003:195) define a research methodology as "a blueprint for conducting a study with maximum control over factors that may interfere with the validity of the findings". The type of research methodology and approach used in this study are discussed in sections 3.3.1 and 3.3.2 respectively.

3.2.1 Research approach and methodology

According to Kothari (2004:5) "there are two basic approaches to research, viz. the quantitative approach and the qualitative approach". When these two are used in combination it is known as 'mixed methods research'. Creswell (2008) further defines mixed method research as "both a method and methodology for conducting research that involves collecting, analysing, and integrating quantitative and qualitative research in a single study or a longitudinal program of inquiry". A key advantage of the mixed method research is that it provides a holistic understanding of a research problem or issue than either research approach alone (Creswell, 2008).

Authors such as Burns and Grove (2011:19) describe a qualitative approach as "a systematic subjective approach used to describe life experiences and situations to give them meaning". Kothari (2004:5) stated earlier that, "the qualitative approach to research is concerned with subjective assessment of attitudes, opinions and behaviour". In contrast, quantitative research focuses on "gathering numerical data and generalising it across groups of people" (Sibanda, 2009:3). Generating data in

the form of using rigorous quantitative methods means that a researcher will at one point manipulate numbers to arrive at certain claims that are backed up by evidence. "In such research, it is the numbers of a phenomenon, an opinion, or the results of an experiment that provide evidence for a researcher to make claims" (Sibanda, 2009). A quantitative researcher has objectives of using measurement means to be able to generalise or replicate findings, which is different from a qualitative researcher because he/she pays specific attention to analysing the intrinsic nature of people being investigated, "and emphasis is on context and flexibility" (Fernihough, 2011). The decision on which approach to use depends on how and why a researcher wishes to collect information. For this research, the researcher decided to primarily make use of a qualitative approach. Once the approach has been determined it is then necessary to select the appropriate research design. The research design needs to be aligned with the research approach. When a researcher undertakes a qualitative research study, the main goal is to find out about people's feelings, opinions, viewpoints and behaviour and to highlight how these sentiments came about (Burgess, 2001:3). However, to efficiently achieve this, it is necessary to acknowledge that a qualitative research study has various data collection methods. These include interviews, observations, past records and documents. As discussed earlier in this chapter, the research design of this study is an embedded one, which combines quantitative and qualitative approaches.

The following section explains the research design selected for in this study.

3.2.2 The research design

Several options were available when selecting the research design. The research design articulates what data is required, what methods are going to be used to collect and analyse this data, and how all of this is going to answer the research question (Van Wyk, 2008:4). Based on this definition, the research design that was used for this study is a case study using a semi-structured interview schedule as the data collection instrument. As was stated above, the researcher intended to utilise a qualitative approach as the primary method to answer the questions on compliance with international trust assessment standards.

Case study research is a methodology that places emphasis on the term 'case'. Hsieh and Shannon (2005) explain that, "the word 'case' means 'an instance of' and

the central feature of case study research design is the investigation of the one or more specific 'instances of' something that comprises the cases in the study". For this research, the 'case' under investigation is the South African repository community.

A case study is useful when collecting new data about a phenomenon that is poorly understood or unknown. A researcher uses a case study to thoroughly study a particular event or programme to find new information about that event/programme and identify which features are common, not common or unique to the specific case (Leedy & Ormrod, 2013:141) (as cited by De Wee, 2013:54). There is very little evidence that the South African repository community understands trust requirements or that there is an awareness of the extent to which our repositories comply with trust standards.

A case study may involve a tangible entity such as an in-depth study of a small number of cases, often longitudinally (prospectively or retrospectively). Multiple sources of data, including interviews, observation, archival documents and even physical artefacts, are used to allow triangulation of findings, and data is collected and analysed about a large number of features of each case (Gomm et al., 2000; Yin, 2013). Due to the nature of mini-dissertations, the researcher did not make use of multiple data collection tools. It was therefore not possible to use triangulation, but this would be essential for further studies.

This section discussed the research approach and strategy used for this research study. The section further explained the research methodology and provided a description of the research design. The following section describes the tools that were used for data collection, as well as the target population and the sampling method used in this study.

3.3 Target population and sampling

The target population for this study was identified to be the following: research data repository and data centre managers at different research institutions and academic organisations in South Africa.

Sampling is defined as "drawing a representative sample from the population, so that the results of studying the sample can then be generalized back to the population"

(Marshall, 1996:522). Sampling "is also done to save time, money and effort while conducting the research" (Kothari, 2004:116). However, a researcher who conducts sampling is faced with the challenge of not being able to test each and every individual in the entire population. The sampling that was used in this research was purposive sampling, which is a non-probability sampling technique (Michael, 2008). Non-probability sampling does not "attempt to select a random sample from the population of interest, rather subjective methods are used to decide which elements are included in the sample" (Michael, 2008). The non-probability sampling technique that was used is known as purposive sampling. In purposive sampling, "the researcher uses his or her own judgment about which respondents to choose, and picks those who best meets the purposes of the study" (Kothari, 2004). This type of sampling was the best choice for this study because of the number of available digital repository managers that could participate in the study. Due to the nature of the research study (mini-dissertation), the cost, availability, and lack of substantial research data repository managers in South Africa, no sampling was conducted.

3.5 Data collection methods

3.5.1 Literature review

In order to develop a research data repository trust assessment model, the researcher used the OAIS Framework to synthesise international repository assessment standards into a trust assessment model. The development of the model required that an extensive literature review was conducted before research data could be collected. There is comprehensive and detailed literature on international standards to assess research data repositories for trust compliance. The literature review (refer to Chapter 2 of this report) assisted in building an informed understanding about the process of evaluating a research data repository for trust status (Snap & Spencer, 2003).

The examined literature also assisted in selecting an already existing model (OAIS) to guide the synthesising process. This synthesis was used to develop a South African research data repository trust assessment model.

It was important to systematically comprehend the different levels of certification that a research data repository can meet, including organisations that are interested in repository assessment. A review of the research data repository assessment tools

and the organisations interested in repository assessment was therefore included in the literature review. The review on research data repository assessment comprised three different levels of certification (basic, extended, and formal) offered by the European Framework for Audit and Certification of Digital Repositories. The reason for reviewing these different levels of certification was to contextualise the ranks of formal certification that is available. The reviewed literature about organisations that are interested in repository assessment included the World Data System (WDS), the Research Data Alliance (RDA), the Research Libraries Group (RLG) and the Online Computer Library Center (OCLC). The reason for selecting these organisations is that each one covers a different domain pertaining to trust accreditation requirements.

To validate the above, a decision was made to use a data collection method that will allow the researcher the opportunity to discuss the collected data with interested parties. This method allowed the researcher to get a first-hand feel and objective response. The assessment was used to gather feedback on the viability of the South African research data repository trust assessment model from research data repository and data centre managers.

This study is founded on a qualitative approach; hence, the face-to-face, structured interview was selected for data collection. The following section explains face-to-face structured interviews. It also explains the interview schedule used and the pilot study that was conducted.

3.5.2 Face-to-face interviews

Face-to-face interviews are "purposeful discussions between two or more people" (Morse, 2005). A researcher would use face-to-face interviews "to collect valid and reliable data through personal communication" (Morse, 2005). The researcher has several options when choosing the face-to-face interview as the data collection tool (semi-structured, in-depth or structured interviews). It is required to design the face-to-face interview schedule keeping these options in mind. The chosen option often determines whether open-ended or closed-ended questions or both could be included.

In addition to offering ample time for respondents, face-to-face interviews permit a researcher to use "visual aids to illustrate points or identify issues s/he is addressing"

(Burgess, 2001). The researcher is also provided with a direct platform to explain any misunderstandings or confusing terms to the respondent. Another reason for using interviews was that the interview schedule allowed for a direct engagement with participants regarding their recommendations to redesign the developed OAIS-based framework.

The face-to-face interview schedule used in this research is provided in Annexure 1. The schedule contains several closed-ended questions each with "structured answers to guide the interviewer". This was done in accordance with advice gained from Farooq (2013). The interview schedule consisted of 8 sections with 37 questions and 7 additional questions. Section 3.6 discusses the sections and themes used in the interview schedule. It is necessary to first discuss structured interviews in more detail.

3.5.3 Semi-structured interviews

According to Harrel & Bradley (2009:27), in semi-structured interviewing, "a guide is used, with questions and topics that must be covered – the interviewer has some discretion about the order in which questions are asked, but the questions are standardized, and probes may be provided to ensure that the researcher covers the correct material". The main reason for using semi-structured interviews is that the researcher wanted to collect information in a conversational manner in order to probe deeply into each topic and question of the interview schedule.

The interview schedule that was used in this study incorporated closed- and open-ended questions. According to Farrell (2016), "open-ended questions are questions that allow someone to give a free-form answers". In contrast, closed-ended questions "can be answered with "Yes" or "No," or they have a limited set of possible answers" (Farrell, 2016). Because both open- and closed-ended questions were asked, it was possible to do both qualitative and quantitative analysis of the data collected from the respondents.

3.5.3.1 Advantages of semi-structured interviews

There are various advantages of semi-structured interviews. The list below provides the advantages by different authors:

- Many researchers like to use semi-structured interviews because questions can be prepared ahead of time (Cohen & Crabtree, 2006). This allows the interviewer to be prepared and appear competent during the interview.
- Semi-structured interviews also allow informants the freedom to express their views in their own terms (Cohen & Crabtree, 2006).
- Semi-structured interviews can provide reliable, comparable qualitative data (Cohen, & Crabtree, 2006).
- Everyone gets the same key questions asked, but there is flexibility in how they are asked (Van Teijlingen, 2014).
- Particularly useful for exploring the views of a person towards something (Van Teijlingen, 2014).

3.5.3.1 Disadvantages of structured interviews

- The skills of the interviewer have an impact on the interview because of factors such as "the ability to think of questions during the interview, for example and articulacy of respondent" (Van Teijlingen, 2014).
- The interviewer may give out unconscious signals/cues that guide respondent to give answers expected by interviewer (Van Teijlingen, 2014).
- Prejudices, stereotypes, appearances and/or perceptions of researcher may alter response (Van Teijlingen, 2014).
- 'Equivalence of meaning' difficulties may arise (Van Teijlingen, 2014), which may cause problems when analysing the data for meaning.

The researcher experienced most of the disadvantages listed above during the pilot testing phase. A decision was then made to also add open-ended questions so that it would be possible to explain the OAIS-based model questions in the coming interviews.

A consent form (see Annexure 5) was used for every interview. Participants gave their consent to be voice-recorded. A clause that allows for the 'curation of the data' was stated in the consent form.

The in-depth nature of interviews takes up a lot of time. To determine the amount of time the interviews would take and the efficiency of the interview schedule, the

researcher conducted a pilot study (refer to section 3.6.1) to be able to prepare accordingly.

3.6. Data collection instrument – an interview schedule

The semi-structured interview schedule (see Annexure 1) that was used in this study incorporated both closed-ended and open-ended questions. The open-ended questions that were used in the interview schedule, aimed at gathering anecdotal information about the functionalities of the different research data repositories. The closed-ended questions were used to collect statistical information required to compare the responses received. The following is an overview of the categories used in the interview schedule.

The questions were divided into the following themes, each with a set of research questions (Annexure 1):

1) Administration capability – investigating the administrative coordination capability of a research data repository. The purpose was to have a better understanding of the control exercised by institutions to manage their research data repositories.
2) Ingest capability – investigating the selection process that a research data repository deploys, including responsibilities of a research data repository.
3) Data management functional capability – scrutinising the experimental group on the basis of how a research data repository manages data and how it delivers information resources to the designated community.
4) Metadata management – investigating the capability of a research data repository to describe content and apply proficient metadata control.
5) Access capability – investigating the capability of a research data repository to offer effective and correct access to resources.
6) Preservation capability – investigating a research data repository's preservation outlook based on guidelines prescribed in organisational policies and the wider organisation mission statement.
7) Financial sustainability – examining how a repository ensures financial sustainability within business continuity.
8) System security sustainability – investigating how a research data repository addresses security issues to maintain a secure working environment.

9) General – additional questions for clarification.

3.6.1 The pilot study

A pilot study is defined as "is a mini-version of a full-scale study or a trial run done in preparation of the complete study – it can also be a specific pre-testing of research instruments, including questionnaires or interview schedules" (Graham, 2001). The pilot study was carried out with a senior library innovation specialist at the University of Pretoria Library. The researcher selected the specialist based on her informed knowledge of library systems and information systems, and her relevant knowledge in the study's research area. The interview schedule was put to use, and it was discovered that some of the questions were confusing and needed to be rephrased using simple terms. Many of the questions were changed after the pilot study. To a large extent, the pilot study assisted the researcher to have an estimated time frame that respondents would need to answer all the questions.

3.7 Selection of research location

The research location is Pretoria, Gauteng province of South Africa. Interviews were conducted at the workplace of each of the respondents. The next section explains how the data was analysed and interpreted.

3.8 Data analysis and interpretation

Data analysis and interpretation is "the process by which sense and meaning are made of the data gathered in qualitative/quantitative research" (Hsieh & Shannon, 2005). The main reason for analysing and interpreting data is to (but not limited) "describe and summarise the data, identify relationships between variables, compare variables, identify the difference between variables, and forecast outcomes" (Hsieh & Shannon, 2005).

All interviews were transcribed before it was possible to start with the data analysis.

The in-depth data analysis and interpretation of this research study is reported in Chapter 4, including the inferences that were drawn from the open-ended questions.

The type of analysis that was used to analyse the quantitative data is a scale of measurement, specifically an ordinal scale. An ordinal scale is where "the data can be classified into non-numerical or named categories, and an inherent order exists

among the response categories – ordinal scales are seen in questions that call for ratings of quality (for example, very good, good, fair, poor, very poor) and agreement (for example, strongly agree, agree, disagree, strongly disagree)". The ordinal scale was represented through data tabulation, which is "the systematic arrangement of the statistical data in columns or rows. It involves the orderly and systematic presentation of numerical data in a form" (Miles & Huberman, 1994). The qualitative data was analysed using data coding, which is defined as "an analytical process in which data is categorised to facilitate analysis" (Miles & Huberman, 1994).

3.9 Conclusion

This chapter has outlined and summarised the primary and secondary data collection methods, research designs and methodology that the author used to conduct the study so that the research questions could be answered. This chapter also explained the data collection instrument, the research location used, the type of sampling used, as well as the pilot study. The chapter concluded with describing the data analysis and interpretation that was used. The next chapter reports on the data analysis and interpretation respectively.

CHAPTER 4: DATA ANALYSIS AND FINDINGS

4.1 Introduction

This chapter focuses on describing the data collected from the semi-structured interviews that were conducted. The collected data was analysed to gain interpretation and meaning. The data analysis was completed by grouping data into different themes, as explained in Chapter 3. This chapter explains the themes and how they are interpreted, including the differences and similarities of the structured interviews. This chapter also discusses the answers of the additional questions that formed part of the interview schedule.

4.2 Background of the data collection and case study

The research study used purposive sampling (a non-probability sampling technique) as discussed in Chapter 3. The digital repository managers that were interviewed comprised two managers of two different repositories at one research institution; to make distinction, they are referred to as Repository Manager A and Repository Manager B, respectively. The other two managers are from different institutions; an academic and a research institution, respectively. The next paragraph explains themes that were used to cluster question for data analysis.

4.3 Themes for data analysis

The themes that were used to group the interview questions were discussed in Chapter 3. The themes were divided in the following manner:

- Segment 1 investigated the administration coordination capability of a digital repository, which was answered by the experimental group – to have a better understanding of the control exercised by institutions to manage their research data repositories.
- Segment 2 investigated the selection process that a digital repository deploys. This included responsibilities of a digital repository. The experimental group provided examples of the typical ingest processes that it uses.
- Segment 3 scrutinised the experimental group on the basis of how a digital repository manages data and how it delivers information resources to its designated community.

- Segment 4 investigated the capability of a digital repository in relation to its metadata capability.
- Segment 5 investigated the capability of a digital repository to offer effective and correct access to resources.
- Segment 6 investigated a digital repository's preservation outlook based on guidelines prescribed in organisational policies and the wider organisation mission statement.
- Segment 7 examined how a digital repository ensures financial sustainability within business continuity.
- Segment 8 investigated how a digital repository addresses security issues to maintain a secure working environment.

The outputs of these segments, including the questions that were asked in the semi-structured interviews, are described in the following section.

4.3.1 Segment 1: Admin coordination capability infrastructure

This section explains the administration competency of digital repositories as compared to the requirements of international trust assessment standards.

DIN 31644: C9 Personnel

The first criteria of the administration coordination infrastructure investigated if digital repositories have adequate numbers of qualified staff members to manage the repository. From the reviewed literature it became clear that repository staff need to have the right expertise and skills to oversee the administration requirements of a digital repository. The questions below were asked to each respondent.

Does the organisation have sufficient and qualified staff members available to manage the repository?

The majority of repository managers indicated that their unit does not have sufficient staff members to run the repository. Repository Manager B from Institute 1 explained that where there are sufficient staff members, the staff members have good skills from a data side, but that they battled with Information Technology (IT) skills. The shortage of IT skills spanned across all the institutions that were interviewed, although it was reported that IT support is provided by the various IT units in all institutions. Due to the nature and specialisation of one repository, the repository

manager from Institution 1 explained that data processing skills are needed by repository staff members. The same manager reported that the repository has qualified staff members but due to organisational constraints, the number of staff members is not sufficient. The repository manager from Institute 3 reported that their unit has a sufficient team of indexers that are being trained to have the skills of managing the repository. The repository manager from Institute 2 elaborated that their digital repository does not have sufficient staff members to run the repository; they only receive technical support from the IT department.

Are the required qualifications set out (including an organisational chart)?

It was discovered that the required qualifications are set out differently from institution to institution. The repository manager from Institute 2 explained that the qualification requirements have changed for the level of working with the institution's repository over the years. The same repository manager believes that a research background (at a master's level) is needed to run a repository, but that this cannot happen due to the institution's policy. The repository manager from Institute 1 eluded that the required qualifications to manage the repository are set out in the job advertisement for staff; there is an organisational chart but qualifications are not set out in the organisational chart. Repository Manager B from Institute 1 explained that there is no organisational chart in the library that set out the required qualification and that there are no specific people in the organisation that have specifically studied to manage a library repository. The repository manger from Institute 3 confirmed that all the required qualifications are set out in the job descriptions and are represented in the institution structure. The job descriptions are approved by the Human Resource department and executives responsible for information services.

Is there a staff development plan that outlines the tasks and objectives of the repository?

It was found that staff development plans are documented differently across the institutions. The repository manager at Institute 2 clarified that development plans are specified in a policy, directive, and job descriptions – there is also a work plan (with performance agreement) for the Personal Development Plan. The repository managers from Institute 1 gave diverse answers with regard to having a staff development plan that outlines the tasks and objectives of the repository. Repository Manager A confirmed that there is a development plan, while Repository Manager B

from Institute 1 disagreed to having a staff development plan. Different feedback was provided by the repository manager from Institute 3 in that the institution uses a performance contract that specify specific targets that should be reached. Every year each staff member must submit a personal development skills plan to address gaps that they have identified within their skills sets. There is also an organogram that sets out the structure.

DIN 31644: C20 Technical Authority

The second criteria of the admin coordination infrastructure investigated how a repository allows users to use its digital information.

Does the repository have processes in place that ensure authority control on a permanent basis without technical restrictions (e.g. encryption, copy and print protection)?

The repository manager from Institute 2 confirmed that the repository has mechanisms in place that allow access rights to be given to users. In contrast, one of the repository managers from Institute 1 said that their repository does not have processes in place that ensure authority control. Repository Manager B at Institute 1 agreed that the library repository has control tools to allow users to use digital objects for what they are intended for. At Institute 3, a comprehensive process is used which involves having no limit on the format of the bit maps. Copyright restriction does not allow the repository to make full articles available, only published on an abstract level. The same repository manager adds records, does the quality control, and verifies the bit maps. Copyright is also verified against the SHAPE/RoMEO guidelines.

ISO 16363: Contracts, Licenses, and Liabilities

The third and final criteria of the administration infrastructure questioned respondents if their institution's repository operates according to set guidelines that set out how digital material are managed, preserved and accessed.

Does the repository have and maintain appropriate contracts or deposit agreements for digital materials that it manages, preserves, and/or to which it provides access?

The repository manager from Institute 2 explained a key and unknown (externally) dimension about their institution's repository: Archival material is not hosted by the

repository any longer. Currently, a terms and conditions (with copyright owners) agreement is signed for archival material and legislation, and requirements of various publishers in the case of articles are followed. At Institute 1, both repository managers explained that internally there are no depositor agreements, and that processes are managed/governed by policies. Work done for external clients is managed by the research contract. At Institute 3, the repository manager explained that deposits are stated clearly in the conditions of service that relate to copyright conformance.

Are formal deposits and contracts legitimate, i.e. are they countersigned and current?

It was discovered that both repositories at Institute 1 do have legitimate contracts and deposits, which are current and countersigned. However, Repository Manager B gave a different response in that everything that is captured on the organisational database goes to the repository. Users sign an 'authorship sign-off form' that accompanies every output that is in the library repository. The repository manager from Institute 2 verified that all deposits are conducted using a submission form to grant approval for the repository to use the document. In addition, a risk compliance certificate is followed for the repository. At Institute 3, the repository manager stated that contracts are stated clearly in the conditions of service that relates to copyright conformance.

The responses provided by the research participants demonstrate a medium to full compliance level to the criteria requirements of the DIN 31644 and ISO 16363 standards respectively, which relate to the administration capability of digital repositories. An apparent but negative finding is that digital repositories do not have organisational charts that set out the qualifications required of staff members for the digital repository. All institutions have contracts or deposit agreements that inform the operation of the digital repository. These contracts and agreements are updated either quarterly or annually by the different institutions.

4.3.2 Segment 2: Ingest capability infrastructure

This section explains the ingest competency of digital repositories as compared to the requirements of international trust assessment standards.

DIN 31644: C1 Selection of information objects and their representations

The first criterion of the ingest capability infrastructure investigated topics related to the criteria that a repository has in place and uses to select which information objects are to be ingested.

Does the repository have a criterion that defines the selection of information objects and their representations?

The repository managers gave positive feedback with regard to the criteria that define the selection of information objects. The Institution 2 repository manager explained that the scope of their repository is clearly defined in the form of research outputs (thesis/dissertations, conference papers, inaugural lectures, etc.). Both of the repository managers at Institution 1 confirmed that there are criteria of what gets added to their repository and what does not. The repository manager at Institution 3 agreed that their institution does have selection criteria. These criteria define the selection of information objects and their representations; however, a formal workflow system is used and is also stated in performance contracts.

Is the selection of digital information transparently documented on the basis of criteria, guidelines and profiles?

The guidelines that each organisation uses differ. The repository manager from Institution 2 confirmed that their repository uses submission guidelines. Repository Manager A at Institution 1 clarified that the criteria that inform the selection of digital information are documented in a policy. Repository Manager B at Institution 1 also confirmed that there are specific rules of what gets added to the repository and what becomes uploaded in electronic formats in the library repository. The repository manager at Institution 3 reported that authors request a publication number in the workflow whereby they have full control of the process. This is an auditable process conducted annually.

DIN 31644: C21 Submission information packages

The second criterion of the ingest capability infrastructure investigated how a digital repository manages the content it receives.

Does the repository adequately specify the composition of data packages for data transfer?

This question specifically asked respondents if their institution's repository specified what kind of data it receives. The repository manager from Institution 2 gave a 'not applicable' answer to this question. Both of the repository managers at Institution 1 confirmed that their repositories have guidelines that specify the composition of data packages. Their repositories are always up to date and aware of what content the repository ingests. Repository Manager B added that users know what they must send to the administrator that will be uploaded on the repository. At Institution 3, this process is still under investigation.

Does the digital archive have specifications regarding content data that is accepted, and the metadata required?

Research participants were asked about the specifications regarding content data that is accepted by their institution's repository. The repository managers from Institution 1 agreed that their repositories have specifications in this regard; which are documented in a policy and guidelines. The repository manager from Institution 2 responded 'not applicable' to this question and the Institution 3 repository manager confirmed that this process is still under investigation due to the diverse nature of their institution.

ISO 16363 Ingest: Acquisition of content

The third criterion of the ingest capability infrastructure focused mainly on management practicality of digital objects and stakeholders.

Can the repository demonstrate (to funders, depositors, and users) what responsibilities it is taking on and what aspects are excluded?

The repository managers of the different institutions gave varied answers to this criteria question. The Institution 2 repository manager mentioned that there are certain tasks that the library repository does (manage embargo dates, support and

train clients) as part of the evidence required by this criterion. Repository Manager A from Institution 1 confirmed that their department's repository can produce evidence, while Repository Manager B from Institution 1 reported that the library repository does not yet have a policy that stipulates that evidence should be produced in this regard, and that there's only a policy on research outputs. At Institution 4, the repository manager explained that there is a procedural document and not a policy per se. At Institution 3, the repository is also used for marketing to showcase the type of work the repository is doing. There is no formal document.

Can the repository determine and check what the characteristics and properties of preserved items to be used over the long term will be?

The repository manager at Institution 2 highlighted that PDF format is the preferred format, but it is not always practical because some documents that the repository receives do not have certain qualities to be converted to PDF format. Repository Manager A at Institution 1 reported that their department's repository keeps a track record of preserved items. Repository Manager B of Institution 1 responded that it might be possible to achieve this on the software side, but it has never been tried or tested because the software is not managed locally – it is done by an external service provider. At Institution 3, the repository manager confirmed that this type of analytics does not exist.

The feedback that the repository managers provided for the ingest capability infrastructure demonstrates a high compliance level to the DIN 31644 and ISO 16363 standards, respectively. It was discovered that digital repositories have effective workflow systems and other processes that are used for the process of ingesting. Policies, guidelines and research outputs are the used formats when ingesting takes place. However, where compliance criteria are not met, the criteria question is not applicable to the digital repository.

4.3.3 Segment 3: Data management capability infrastructure

This section explains the data management competency requirements of digital repositories as compared to the requirements of international trust assessment standards. The assessment criteria of the data management capability infrastructure were changed after the first interview with the Repository Manager A of Institution 1. It became clear that it was necessary to probe repository managers if their

institutions contains or intends to integrate and manage research data in the future. Consequently, the following question was adapted and included in the subsequent interviews.

Does the repository contain data or intend to integrate and manage research data in the future?

The Institution 2 repository manager confirmed that the integration of data into the repository is still under investigation and pilot projects are being conducted. Repository Manager A from Institution 1 manages a research data repository and highlighted that research data is the crux of their repository. Repository Manager B of Institution 1 said that the library is looking into integrating all the available institution data into the repository. The repository manager from Institution 3 explained that the institution's plan is not to include data in the repository, but to provide links from where the data can be obtained. The large size of the data (terabytes) won't be handled effectively by the repository. Investigations into storing data in a cloud are being made. The repository manager from Institute 3 further explained that the institution is investigating cloud storage services provided by the Data Intensive Research Initiative for South Africa (DIRISA). DIRISA provides the digital storage infrastructure and services in South Africa, to reliably and persistently store and share research data. Data will be soundly and reliably managed throughout their life cycle, from upload/deposit to their preservation or expunction (DIRISA, 2016).

ISO 16363 Information Management: Specify minimum information requirements to discover and identify material of interest

The first criteria of the data management functionality infrastructure probed respondents on the ability of a digital repository to adequately handle user information needs and the technicality involve.

Is the repository able to deal with the different types of information requests made by users from the designated community?

All the repository managers agreed that their institution's repository is able to deal with different types of information requests. The general feedback was that

interruptions occur when systems are down due to maintenance or repair action that takes place occasionally at scheduled periods.

Does the repository have adequate retrieval and descriptive information, discovery metadata (such as Dublin Core), and other documentation describing information objects to be retrieved?

Conformance was agreed to by all of the repository managers that their institution's repository deploys discovery metadata (i.e. Dublin Core). The figure below represents the usage ratio of metadata schema by digital repositories who make up the sample of this study.

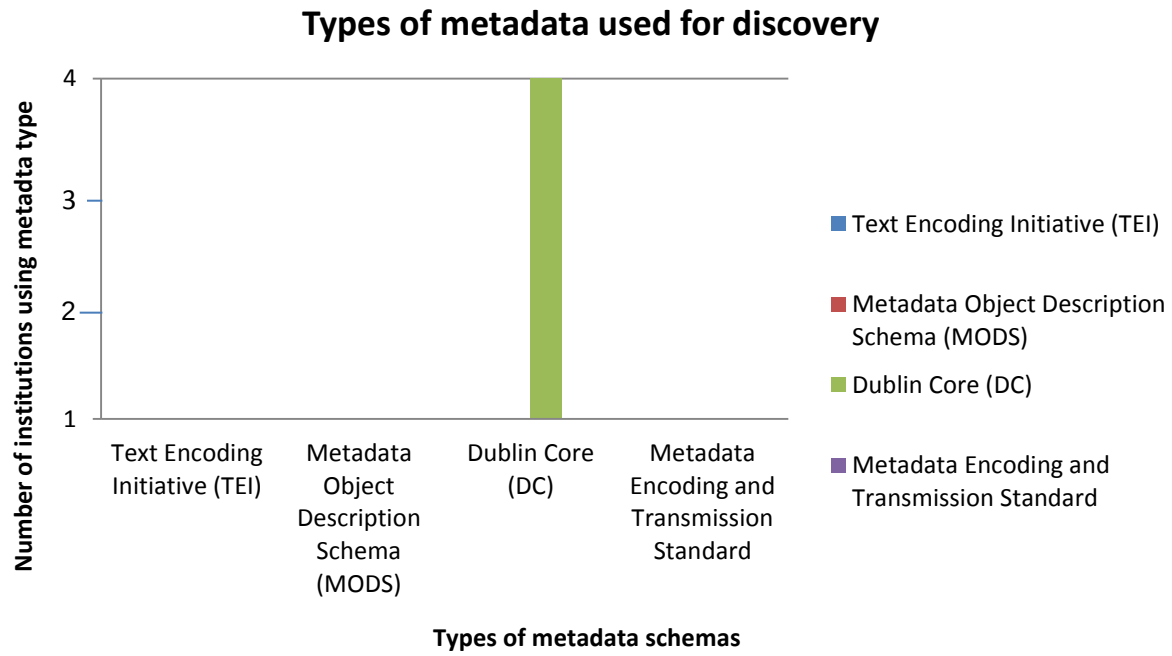


Figure 2: Types of metadata schemas used by South African repositories

ISO 16363 Information Management: Capture or create minimum descriptive information associated with the Archival Information Package

The second criterion of the data management functionality infrastructure probed respondents on the ability of a digital repository to provide relevant information to the user community.

Can the repository deal with the types of requests that come from a typical user from the designated community?

This criterion question probed if the repository is able to deal with the types of requests that come from users. A distinctive response was provided by the repository manager from Institute 2. The manager explained that due to the distance-learning aspect of the institution, remote areas are a factor that the repository considers and tries by all means to cater for the needs of the user community. Both repository managers of Institute 1 confirmed that their institution repositories can deal with the types of requests that come from a typical user from the designated community. Due to the nature and magnitude of Institute 3, the repository manager highlighted that if an item is not in the department's repository, other repositories of the institution are contacted to check for the item and make it available if it is found.

ISO 16363 Information Management: Maintain bi-directional linkage between each AIP and its descriptive information

The third criterion of the data management functionality infrastructure probed respondents on the ability of a digital repository to use tools to retrieve the relevant information and the repository's responsiveness.

Can all the Archival Information Packages (AIPs) be located and retrieved? Are there adequate measures (such as descriptive metadata; unique, persistent identifier) in place to ensure that the AIPs are located and retrieved?

The Institute 2 repository manager explained that their institution's repository uses more than one measure to verify that archival information packages are located and retrieved. The repository uses descriptive metadata, persistent identifier and digital object identifier. Repository Manager A of Institute 1 confirmed that their department's repository is able to locate and retrieve AIPs. Repository Manager B of Institute 1 reported that descriptive metadata is being used but they are still in the process of acquiring the handle and the persistent identifier. However, at the moment there are no such tools. The Institute 3 repository manager explained that the repository software (DSpace) can handle this. The unit is busy looking into the implementation of ORCID ID.

Is there a procedure in place that notifies the repository when the relationship between the data and the associated descriptive information is temporarily broken to ensure that it can be restored?

A digital repository will occasionally experience difficulties in accessing information as and when it is needed. This criterion probed repository managers if their institution's repository has procedures in place that notify them when information cannot be accessed. Of the four repository managers, only one repository (Institute 1: Repository Manager B) has an email system that notifies the administrator when information cannot be accessed by users.

TRAC – Chain of Custody: Have physical and legal control over the existence, authenticity, location, and accessibility of records

The fourth criterion of the data management functionality infrastructure probed respondents on the ability of a digital repository to manage all digital information in a workflow process.

Can the repository demonstrate the chain of custody for all of its digital content from the point of deposit?

The repository managers from Institute 1 agreed that their institution's repositories utilise a workflow process to demonstrate the chain of custody of its digital content. However, the Institute 2 repository manager added that their institution's repository also captures description provenance metadata. The repository manager from Institute 3 elaborated that the system that is in place is not visible to the end-user, but visible to the administrator (notifies who did what, the ownership, the ingest workflow).

Can the repository demonstrate that the content it has matches the content it received?

It was discovered that there are different processes that the various institutions use in terms of demonstrating that the content that resides in the repository matches the content it received. The Institute 2 repository manager eluded that users of the repository enter metadata which goes into the workflow that is reviewed by the library staff. At Institute 1, Repository Manager B confirmed that the repository cannot demonstrate that the content it has matches the content it received; therefore

the administrator operates on an ad-hoc basis. There is no real workflow process that users use to submit documents. The administrator receives an email and has to make sure that it is uploaded. Repository Manager A of Institute 1 reported that their department's repository can demonstrate that the content it has matches the content it received by being able to produce reports. A different feedback was provided by the Institute 3 repository manager, in that the repository cannot demonstrate such a match, but the supporting workflow is able to validate. It is not part of the repository process.

The responses provided by the research participants demonstrate a medium to full compliance level to the criteria requirements of the ISO 16363 and TRAC standards, respectively, which relate to the data management capability of digital repositories. It was discovered that the inclusion of managing data in the digital repositories is regarded as important, but will need a theoretical and practical understanding of how to manage research data. In terms of addressing the user needs of the designated community (both on a technical and practical level), all the digital repositories have operative tools in place to effectively achieve this.

4.3.4 Segment 4: Metadata management infrastructure

This section explains the metadata management competency of digital repositories as compared to the requirements of international trust assessment standards.

DIN 31644: C5 Interpretability

The first criterion of the metadata management infrastructure investigated if digital repositories have metadata schema in place and the level of interpretability thereof.

Does the repository have a metadata schema that it uses?

It was discovered that all the repositories do have a metadata schema that they use. The common finding is that all the repositories utilise Dublin Core as a metadata schema. However, the repository manager from Institute 3 explained that although their repository uses Dublin Core, with various subject areas additional metadata might be integrated in the future to address the dynamic needs of the organisation.

Can the repository ensure long-term interpretability of at least one representation of content data and metadata?

It was established that all repositories that were investigated use different methods to ensure long-term interpretability of content and content data. The Institute 2 repository manager said that their digital repository uses a persistent URL, even when migration is done from one system to another. The repository managers of Institute 1 confirmed that both repositories can ensure long-term interpretability. However, the answer provided by Repository Manager B was based on an assumption that the repository software (DSpace) should be able to ensure long-term interpretability. This was also the response provided by the repository manager from Institute 3.

Does the repository have methods to allow the user community to check interpretability on a regular basis?

Repository managers gave different answers to this question. The repository manager from Institute 2 responded that repository users can check their workflow in the space where they submit, but after submission, users cannot monitor their workflow. Repository Manager A of Institute 1 said that their department's repository does not have methods in place to allow the user community to check interpretability on a regular basis. In contrast, Repository Manager B of Institute 1 stated that the user community have access to what they see; there is a way that users can contact administrators if something is not right. The Institute 3 repository manager gave a one-dimensional response to this criterion in that there is no formal process in place; it either works or it does not. There is no process in place; if users cannot perform a task, it is reported to the repository.

The responses provided by the research participants demonstrate a relatively full compliance level to the criteria requirements of the DIN31644 standard, which relates to the metadata management capability of digital repositories. It must be noted that the digital repositories that were investigated have to improve on the degree of checking interpretability on a regular basis. Metadata is an important aspect in the operation of a digital repository. It was clear that the digital repositories that were used for this study comprehend this notion.

4.3.5 Segment 5: Access capability infrastructure

This section explains the access management competency of digital repositories as compared to the requirements of international trust assessment standards, and how digital repositories facilitate contact to their digital content.

ISO 16363: Access management – comply with access policy

The first criterion of the access management infrastructure investigated if digital repositories address usage aspects in a legitimate and controlled manner.

Is the repository able to produce evidence to demonstrate that it has fully addressed all aspects of usage which might affect the trustworthiness of the repository?

It was revealed that all repositories, except one, do have guidelines and procedures in place that can be used as evidence to demonstrate that it fully addresses all aspects of usage that might affect the trustworthiness of the repository. The repository manager from Institute 2 reported that their repository uses and regards their workflow process, guidelines, procedures and standards as main evidence. Repository Manager B at Institute 1 agreed that the library repository is able to produce evidence in that there are specific rules of what can be accessed and what cannot be accessed. This is implemented if digital material is open access and made available. If not, a 'contact us' message is given to the user for enquiry purposes. Repository Manager A of Institute 1 said their department's repository cannot produce evidence of this nature. At Institute 3, the repository manager explained that they receive emails from clients when clients cannot access certain material on the repository. There is a daily check-up process on usage statistics, but evidence cannot be produced.

ISO 16363 Access management – dissemination of original digital objects with evidence

The second criterion of the access management infrastructure investigated if digital repositories follow guidelines for the manner in which digital material is disseminated.

Does the repository follow policies and procedures that enable the dissemination of digital objects that are traceable to the originals, with evidence supporting their authenticity?

It was discovered that the management of digital objects differs from one institution to another. The repository manager from Institute 2 stated that the articles of original published articles are captured and a link is created to the original. However, some articles cannot be traced, including the original publication. Both the repository managers of Institute 1 agreed that there are policies and procedures as to where information is stored and where it gets linked to, and these policies are followed consistently. This is also the case at Institute 3 with the use of a records management policy, which is audited annually.

DIN 31644: C4 Access

The third criterion of the access management infrastructure investigated if digital repositories are capable of operating in a legal framework.

Can the repository ensure authorised access to information for the designated community?

A common element amongst the different digital repositories is that each repository specifically puts the information needs of its parent organisation staff members first. To achieve this, the type of authorised control is achieved through assigning username and passwords for repository users. This method was reported by all the repository managers.

Does the repository have appropriate search possibilities, which indicates the terms of use and restrictions?

The Institute 2 repository manager explained that they have a copyright statement that users have to agree to on the various collections and on each item they want to access. The statement appears on every item in the repository. This is the method used to indicate terms of use and restrictions. Repository Manager B of Institute 1 confirmed that their repository provides both an 'advanced' and 'general/simple' search functionality. Repository Manager A of Institute 1 confirmed that their department's repository does have search possibilities that users can utilise. This

was also the case at Institute 3; however, the repository manager said that the search possibilities can be improved based on functionality.

Does the repository declare its conditions of use and costs that may arise?

It was revealed that the digital repository of Institute 2 has no costs involved. The repository manager responded that their repository only provides open access and restricted access. Repository Manager A of Institute 1 confirmed that their department's repository does declare its conditions of use and costs that may arise with regard to material that has restricted access rights. Repository Manager B of Institute 1 on the other hand, stated that there is no cost because it is publicly funded research and is made available free of charge. The repository of Institute 3 does not declare its conditions of use and costs that may arise. The repository manager clarified that if digital material is available it can be used, and there are no restrictions – there is no grey area in this regard.

The responses provided by the research participants demonstrate a full compliance level to the criteria requirements of the ISO 16363, DIN 31644 and TRAC standards respectively, which relate to the access capability of digital repositories.

Nevertheless, it would be ideal if two of the digital repositories were to introduce the process of declaring conditions of use and costs that may arise, so that users can be informed as to how they are using the repository. The meaning and classification of what 'evidence' entails in regard to demonstrating that the repository has fully addressed all aspects of usage differs from one institution to the next. The different document types (i.e. workflow, guideline rules, and emails) can create confusion as to the degree of the usefulness the document serves as tangible evidence. The usage of a policy when disseminating digital objects is a common trend amongst all the digital repositories. This compliance ensures that the investigated digital repositories operate in legitimate boundaries. The policies that the digital repository uses also enable the digital repositories to have procedures in place to allow authorised access to information for the designated community. The usage aspect of the digital repositories is an area that all investigated digital repositories excel in.

Every digital repository of this study offers users appropriate search possibilities, which indicate the terms of use and restrictions that apply to the usage of the repository. Due to the open access nature of the digital material housed by the

repositories, only one of the four repositories (Institute 1: Repository Manager A) declared its conditions of use and costs that may arise. Access provision of digital material is effectively provided by all the repository that were investigated. The open access dimension spans across all institutions which eliminates the barrier of accessing digital information.

4.3.6 Segment 6: Preservation capability infrastructure

This section explains the preservation capability of digital repositories as compared to the requirements of international trust assessment standards.

In order to fully scrutinise the preservation capability of the different digital repositories, it was compulsory to first determine if the repositories provided a long-term preservation service. The infrastructure question below was adapted by the researcher and does not originate from an international repository assessment standard.

Can the repository offer a long-term preservation service?

There are different approaches that each institution utilises for long-term preservation. The Institute 2 repository manager reported that there are plans to use an external hosting source to preserve material. Persistent URL is used in the preservation process whereby only backups are made. Repository Manager B of Institute 1 explained that long-term preservation is conducted through the repository software (DSpace) that is updated regularly. But without the software, the repository cannot provide long-term preservation. In contrast, Repository Manager A from Institute 1 explained that preservation is only done on a medium-term basis and not in the long term. At Institute 3, digital material is kept indefinitely, and it is backed up regularly and stored off-site. No deletion takes place because there is a deletion policy in place that prevents this from happening.

ISO 16363: Governance and Organizational Viability

The first criterion of the preservation capability infrastructure investigated if preservation is regarded as a key objective and the processes that govern the preservation mission of the digital repository.

Does the parent organisation or the repository's mission statement explicitly address preservation?

From a wider organisational perspective, it was discovered that most organisations do not address preservation in their mission statement. In support of this finding, the repository manager from Institute 2 explained that preservation is not addressed. The Institute 1 repository managers explained that preservation is clearly addressed in policies (research output policy). In contrast, the Institute 3 repository manager explained that their records management policy addresses all preservation issues.

DIN 31644 C18: Authenticity – Preservation measures

The second criterion of the preservation capability infrastructure investigated the type of preservation strategies that the digital repositories deploy. However, this question was only asked if a repository provides long-term preservation.

Does the repository deploy methods which ensure the authenticity of the objects during implementation of the long-term preservation measures and document the degree of authenticity?

All repository managers (except the Institute 3 manager) responded 'not applicable' to this question, since long-term preservation cannot be provided. However, the repository manager from Institute 3 explained that before the workflow process ends during indexing, the owner of the document must verify that the bit stream added and all the metadata included is an accurate reflection of the item. Ad-hoc verification is done to ensure that metadata added is correct. A high level of quality assurance is maintained.

Does the repository ensure that relevant information objects retain their authenticity while undergoing preservation processes and that all measures are transparently and permanently documented?

Three repository managers responded 'not applicable' to this question. The Institute 3 repository manager reported that the problems encountered are noted and looked into.

As part of preservation, does the repository assert provenance of digital objects by guaranteeing that objects are stored intact as they have been created?

Repository Manager B from Institute 1 was the only manager who gave positive feedback to this question. Repository Manager B reported that digital objects are saved in a document management system. All the other remaining repository managers stated that their repositories cannot assert provenance of digital objects by guaranteeing that objects are stored intact as they have been created. The repository manager at Institute 3 elaborated that the owner of digital material is asked to sign off and approve the process of preservation.

The responses provided by the research participants demonstrate a low to medium compliance level to the criteria requirements of the ISO 16363, DIN 31644 and TRAC standards, respectively. However, a key concern is the lack of preservation strategies that digital repositories have in place. Not having preservation strategies causes institutions to lack an informed understanding of what long-term preservation is. Long-term preservation as prescribed by the international standards used for this infrastructure differs from what is regarded as long-term preservation by the investigated institutions. This realisation creates compliance confusion because of using different meanings of long-term preservation (refer to the Glossary for the definition of long-term preservation). Nevertheless, two out of four repository managers agreed that their institution's repository does offer long-term preservation. However, the follow-up question relating to information objects retaining authenticity while undergoing preservation processes was answered negatively, which proves that the concept 'long-term preservation' is not mutually understood.

4.3.7 Segment 7: Financial sustainability infrastructure

This section explains the financial sustainability of digital repositories as compared to the requirements of an international trust assessment standard.

ISO 16363: Financial Sustainability

This criterion of the financial sustainability infrastructure investigated if a digital repository has adequate financial business plans to sustain the repository over time.

Does the repository have a short- and long-term business planning process in place to sustain the repository over time?

The findings related to the financial processes of digital repositories revealed a common aspect that institutional funds are not being allocated to digital repositories separately. Funds are shared with other units – IT and ICT being the main beneficiaries. The repository manager of Institute 2 highlighted that the financial processes of the institution create problems; they take a long time to approve. The manager further confirmed that the financial management of the repository comes from the ICT budget. Repository Manager A at Institute 1 confirmed that their department's repository does have short- and long-term business planning processes in place, and these are used accordingly. Repository Manager B at Institute 1 stated that there are neither plans nor processes in place; things are implemented as they go along. At Institute 3, the repository manager explained that there is a medium- to long-term business planning process documented in a business plan which is reported on quarterly and annually.

The responses provided by the research participants demonstrate a low to medium compliance level to the criteria requirements of the ISO 16363. A common underlying challenge that most of the institutions face is that the digital repositories are not given their own fiscal programme to sustain the repository over time. This in turn causes digital repositories to adopt a reliance and secondary approach of receiving funds.

4.3.8 Segment 8: System security sustainability infrastructure

This section explains the system security sustainability of digital repositories as compared to the requirements of international trust assessment standards.

DIN 31644: C34 Security

The first criterion of the system security infrastructure investigated if there are security measures in place to protect the integrity of a digital repository. Without security measures, a digital repository becomes very vulnerable to malicious cyber-attacks.

Can the organisation and the infrastructure protect the repository and its information objects and representations to ensure integrity?

It was discovered that a key factor behind security measures is associated with the services received from the various IT departments. The repository manager from Institute 2 reported that their institution's IT department is responsible for the security services at the repository, and that 'backups' are made as a solution. The difference between the responses given by the two repository managers from Institute 1 is that the one repository is hosted onsite and receives support from the IT unit, whilst the other repository is hosted by a server that is outside of the organisation. Repository Manager B stated that as long as payment is made, hosting will be there. The repository manager of Institute 3 explained that security protection is part of ICT, hence system security is a formal process that the repository receives.

Does the repository have suitable measures to protect its own integrity and that of its digital assets to ensure that the assets remain intact, and to fulfil its legal or contractual obligations?

A common similarity amongst the responses that repository managers gave is the dependency that the repository has on its IT support. The Institute 2 repository manager confirmed this realisation by stating that their repository's continuous functionality is dependent on the IT unit. Both repositories at Institute 1 do have measures to protect their integrity. In addition, Repository Manager B of Institute 1 reported that there is a contract with DSpace software to maintain a certain level of service. At Institute 3, the repository manager explained that the ICT infrastructure manages this aspect.

ISO 16363 Security risk management

The second criterion of the system security infrastructure examined if there are security counter measures in place that a digital repository deploys for unexpected threats.

Does the repository conduct regular risk assessments and maintain adequate security protection in order to provide expected and contracted levels of service?

Risks are dealt with in different ways at different institutions. The repository manager from Institute 2 explained that risks are identified but a risk assessment is not

conducted. Repository Manager B at Institute 1 reported that a risk assessment has never been done, whilst Repository Manager A from Institute 1 reported that the IT department does a risk assessment every quarter, and that the department has a person who specialises in security risk. This was also the case at Institute 3, as risk assessments form part of the manner in which the ICT infrastructure is managed.

Can the repository show how it deals with its security requirements?

There are different responses to the manner in which digital repositories handle their security requirements. The Institute 2 repository manager replied that the repository does not have specific security requirements that it conforms to; the repository follows the wider organisational policy of ICT. Both repository managers at Institute 1 confirmed that their department's repository can deal with security requirements. However, the Repository Manager B explained that the security requirements are stipulated in contracts, and the contracts specify who has access to digital material and who does not. The repository manager at Institute 3 confirmed that ICT also manages this component.

The responses provided by the research participants demonstrate a medium to high compliance level to the criteria requirements of the ISO 16363. A key concern is the absence of conducting regular risk assessments and maintaining adequate security protection measures. Taking into consideration the changing and volatile nature of the computing environment, regular risk assessments are crucial for minimising potential threats.

The above section provided the findings of the data that was collected using the developed OAIS-based model to assess South African digital repositories for trust status. Each infrastructure of the developed model probed repository mangers using specific and relevant questions originating from international repository trust assessment standards. Apart from these criteria questions, repository managers were asked additional questions to gain further understanding of the usefulness of the developed model. The next section provides an overview of the answers given by repository managers.

4.4 Responses to additional questions

4.4.1 From a manger's point of view, what are the strengths and weaknesses of the model?

Weaknesses:

The repository manager from Institute 3 explained that the human element, as a key component of processes, becomes a weakness. This is because subjectivity is an issue. There is no way to eliminate the human component (i.e. people still need to submit information into the repository and interpret information). In any model, the role of the human being is the weakest link because it is not an automated process.

Repository Manager A from Institute 1 stated that the weaknesses and strengths will become clearer when the model has been applied and tested in different organisations. However, some of the questions of the model (coming from international standards) are not really clear. Some of them have one evaluation of two aspects in the same question. The rating: a quantitative measure is fine but evaluation to get accreditation should not just be in terms of the numeric measurements. The comments are the most important input, because this is where reasons are provided for the rating that was selected. There must be minimum standards to the ratings. The questions need specific explanations and examples for clarification.

Repository Manager B from Institute 1 and the repository manager from Institute 2 did not find any weaknesses in the model.

Strengths:

Two of the repository managers (Institute 3 and Institute 2) both agreed that the developed model is a great departure point to correct the unaddressed areas of repository sustainability in a South African context (Phase 1). The developed model is a very good starting point for establishing a digital repository. The model is straightforward and to the point; it is neither too long nor too concise. It offers a good coverage and it is realistic for South Africa.

Repository Manager A from Institute 1 emphasised that the model simplifies things, and makes the accreditation process less complicated to understand. The model helps to determine if a repository is worthy of accreditation. Repository Manager B of

Institute 1 said that the model includes all the important data management questions (i.e. metadata). The qualification requirements of the library staff are also an important aspect because library staffs do not have IT skills and expertise.

4.4.2 Which component of the model is most important for your institution?

Repository managers provided a subjective answer to this question. The repository manager of Institute 3 responded that if the administration and preservation planning do not improve, the repository can fall flat. These two components require additional attention to ensure sustainability in their institution. Repository Manager A from Institute 1 highlighted that one cannot say that one part/section is more important than the other, e.g. looking after digital objects requires one to do everything. All the components are important. Repository Manager B of Institute 1 explained that the most important component is the access infrastructure and making information available to users, as this is the mandate of the library. However, the access infrastructure also goes hand in hand with preservation and data management infrastructure. The repository manager of Institute 2 identified preservation infrastructure as the most important component, because it is not taken care of in the institution.

4.4.3 Do you have any recommendations that can be used in redesigning the model?

Repository Manager B from Institute 1 disagreed with this question and explained that this is great starting point. The manager could not provide a recommendation because as the model gets used, things will evolve over time. Repository Manager A from Institute 1 agreed that there is no recommendation to make at this point. Decisive recommendations were provided by the repository manager from Institute 3 in that the Descriptive information component needs improvement. The producer of the information must contribute towards the terminology and the phrases used in the OAIS Framework so that the Access aspect is enhanced. Version control (cannot obtain pre-print version) can be included in the model (maybe at the ingest aspect). The repository manager from Institute 2 suggested that the Authority control aspect needs to incorporate using other services like research tools (e.g. APIS and ORCHID).

4.4.4 To what extent do you see your institution being willing to comply with the requirements of such a system of trust?

Repository managers gave positive responses to this question. The importance of complying with international standards for assessment purposes was emphasised by all the repository managers.

The repository manager from Institute 3 explained that their institute is willing to comply fully; that there is no reason for deviation. The entire model can be used to identify gaps and improve on services. Both repository managers from Institute 1 explained that the institution would be very willing. It is essential for them to comply with a system of trust; however, funding remains the biggest limiting factor. Repository Manager B of Institute 1 explained that for the level of trust that the proposed model aims for, the repository library is willing to comply 100%. The repository manager from Institute 2 explained that in principle, full compliance is desired, but in practice it is going to take some time because of the challenges with ICT and finance.

4.4.5 What level of trust would your institution be able and willing to comply with, in terms of rating?

The repository manager from Institute 3 replied that this is difficult to answer because the model requirements have not yet been tested. A similar answer was given by the repository managers from Institute 1 in that the institution cannot comply with the fullest level of trust because of limited resources (funds and human capital). Infrastructure also needs to be developed to comply with such a system of trust. In contrast, the repository manager from Institute 2 said that their institution would be willing to comply with an 'Agree' rating as used in the interview schedule.

4.4.6 Which of the levels of the model need revision and what should the revision include?

The answers that the repository managers gave to this question were based on their experience of digital repository management. The repository manager from Institute 3 responded that for the Ingest level (legal control) – compliance; administration and planning – a better understanding of the issues involved is needed. Repository Manager A from Institute 1 said that the model makes provision only for the dissemination of data, and the repository is actually also playing an important role in

terms of promoting the use of its holdings, providing assistance and training, and interacting with both producers and consumers. Repository Manager B from Institute 1 mentioned that the model helps the library to identify what needs to be revised (such as policies licences). The developed model does not need revision. The repository manager from Institute 2 said that the Authority control management level needs revision.

4.4.7 From the developed model, what components were forgotten?

As with the previous question, the repository managers used the experience they have of working with a digital repository to answer this question. The repository manager from Institute 3 explained that more examples are necessary; people look at concepts with a narrow view of their interpretation. Repository Manager A from Institute 1 said that the model makes provision only for the dissemination of data. The repository is also playing an important role in terms of promoting the use of its holdings, providing assistance and training, and interacting with both producers and consumers. In contrast, Repository Manager B from Institute 1 and the repository manager from Institute 2 explained that at this point not much has been forgotten. The basic concepts are included, which might change as repositories evolve.

4.5 Conclusion

The data collected using the developed OAIS-based model to assess South African digital repositories indicates that compliance can be achieved. The criteria of international repository assessment standards (used for the developed model) probed repository managers on critical areas. Participants in the study demonstrated a positive attitude to knowing which areas of their digital repository needed improvement. It was found that repository managers strive for full trust compliance, but currently this is not possible due to organisational policies and financial processes. Repository managers indicated that some features of the model can be changed to make it more inclusive in addressing data management areas. The table below illustrates the compliance level of each digital repository that was assessed and is based on every segment (each criterion) that was used to probe the digital repository managers.

Table 2: An overview of the compliance levels

Segments	Institution	Level of compliance (i.e. Full; >50%; <50%)
1. Admin coordination capability infrastructure	Institution 1 – Manager A	>50%
	Institution 1 – Manager B	>50%
	Institution 2	>50%
	Institution 3	>50%
2. Ingest capability infrastructure	Institution 1 – Manager A	Full
	Institution 1 – Manager B	Full
	Institution 2	<50%
	Institution 3	>50%
3. Data management capability infrastructure	Institution 1 – Manager A	>50%
	Institution 1 – Manager B	>50%
	Institution 2	>50%
	Institution 3	>50%
4. Metadata management infrastructure	Institution 1 – Manager A	>50%
	Institution 1 – Manager B	>50%
	Institution 2	Full
	Institution 3	>50%
5. Access capability infrastructure	Institution 1 – Manager A	Full
	Institution 1 – Manager B	>50%
	Institution 2	Full
	Institution 3	>50%
6. Preservation capability infrastructure	Institution 1 – Manager A	<50%
	Institution 1 – Manager B	<50%
	Institution 2	<50%
	Institution 3	>50%
7. Financial sustainability infrastructure	Institution 1 – Manager A	Full
	Institution 1 – Manager B	<50%
	Institution 2	<50%
	Institution 3	Full
8. System security sustainability infrastructure	Institution 1 – Manager A	Full
	Institution 1 – Manager B	>50%
	Institution 2	<50%
	Institution 3	Full

Based on this analysis, it is very clear that although there are problems here and there, South African digital repositories are complying with the requirements of international standards. Nonetheless, the digital repository of Institute 2 has to address its Ingest capability because it is not complying with the Ingest criteria of the international standards. The Preservation capability is a worrisome area for three of the four digital repositories that were investigated. Preservation is a concept that is

addressed differently by the different institutions. It will be best for institutions to work towards addressing preservation according to the requirements of international standards. One out of the four digital repositories that were investigated does not have efficient security sustainability measures. This can create serious problems for the institution, as the repository is vulnerable to unethical attacks. Improved security measures need to be implemented to ensure that these types of attacks do not happen. Overall, this research study revealed that South African digital repositories are not far off in complying with the full requirements of international repository assessment standards.

CHAPTER 5: RECOMMENDATIONS AND CONCLUSIONS

5.1 Introduction

The main objective of this chapter is to review the research questions and objectives of the entire study as presented in Chapter 1, to form a synthesis of Chapter 2, and to merge that with the findings of Chapter 4. This chapter further discusses in detail the recommendations that arise from the collected data and the analysis thereof – reflecting the milestones of the literature review, and providing an overall conclusion to the research study.

5.2 Findings

The main purpose of this research was to investigate to what extent South African digital repositories comply with international trust standards.

The research questions below were used to answer this research question:

i. Which international repository assessment standards can be used to assess South African digital repositories?
ii. What will a trust model that has been developed based on international trust standards, look like?
iii. To what extent do SA research data repositories comply with this model in terms of trustiness?
iv. How should and/or how could this model be developed for SA as a developing country to make the striving for trustiness more feasible?

The next section reports the answers to the above research sub-questions, including reference to the literature review that was conducted.

5.2.1 Which international repository assessment standards can be used to assess South African digital repositories for trust status?

As discussed in Chapter 2 (section 2.1), different international organisations have developed (in the form of standards) tools to assess if repositories comply with set standards and their criteria. It was discovered that every international repository assessment standard covers similar topics when assessing for trust compliance.

In Chapter 2 (section 2.2), the OAIS Framework was used as a departure point to provide a common definition of what digital preservation of information entails for the

long term. The definitions of the OAIS model are contextualised into six entities that contribute to the framework's mission of facilitating preservation. The OAIS Framework offers a basic blueprint in terms of incorporating concepts that other international repository assessment standards have borrowed. As was discussed in Chapter 2 (section 2.4), there are various international standards to assess a digital repository for trust status. These international standards present a strict framework with which novices need to comply. Based on a rallying and assimilation perspective, the entities of the OAIS Framework were used to incorporate criteria from three standards to develop a South African-based repository assessment model. The three international standards that were used are listed below:

- Trustworthy Repositories Audit & Certification (TRAC)
- The International Organization for Standardization (ISO) 16363
- The 'Deutsches Institut für Normung' (DIN) 31644.

The following section discusses the features and functionality (based on international standards) of the proposed model.

5.2.2 What will a trust model that has been developed based on international trust standards, look like?

As explained in the previous section, three international standards were used in conjunction with the OAIS-based model to develop a model of assessing South African digital repositories for trust status. The five OAIS Framework entities were used as the foundation pillars of the proposed model. Each OAIS entity was discussed in Chapter 2 (section 2.2) whereby a linkage between the OAIS Framework and the criteria of the three international standards was made explicit.

Based on the approach of incorporating the criteria of international standards with the OAIS Framework, a synthesised framework was developed. This synthesis is represented in Chapter 2 (section 2.5). Based on this synthesis, it became convenient to adapt the OAIS Framework so that it can assess South African digital repositories on essential aspects that international standards use, but to a large extent are based on the OAIS Framework. The developed OAIS-based model had to incorporate two aspects from the Research Libraries Group (RLG) and Online Computer Library Center (to assess the financial and security sustainability of a digital repository) to form two new entities of the adapted OAIS Framework.

5.2.3 To what extent do South African digital repositories comply with the developed OAIS based model in terms of trustiness?

The developed OAIS based model is a synthesized prototype that comprises of entities that originate from 3 international repository assessment standards (i.e. DIN 31644, ISO 16363, and TRAC). The OAIS based model offers an extension of the original OAIS framework entities; it includes a system security and financial sustainability entities which originate from RLG & OCLC (2002). The inclusion of these entities allows an institution to reflect its security and financial conditions as required by international repository assessment standards. Every entity of the developed model intends to offer a collective and reliable assessment paradigm that can be used by South African repositories. The data that was collected using the developed model were presented in Chapter 4. Unique findings surfaced in the different themes used to classify the findings. The following section discusses these findings according to each segment in relation to compliance with the developed OAIS-based model.

5.2.3.1 Administration coordination infrastructure

From the data collected it can be observed that the digital repositories that were studied do, to a large extent, comply with the criteria of the various international repository assessment standards. There are shortfalls that were discovered, which are mainly caused by institutional structures and processes. These challenges could be overcome but that would require wider organisational involvement and participation. Championing the changes required is also a deciding factor that organisations can look into to get broader organisational buy-in. Based on the data collected using the OAIS-based model, South African digital repositories also fared well when evaluating the infrastructure requirements.

5.2.3.2 Ingest capability infrastructure

It can be concluded that the South African digital repositories that were used for this study have adequate tools and processes to manage the ingest processes. The level at which the various digital repositories handle user engagement complies with the requirements of international standards. The use of strategy documents, by all the repositories, is a clear indication that compliance requirements are being met.

5.2.3.3 Data management capability infrastructure

All of the digital repositories (except the data curation repository) intend to include and manage research data in the future. This will require a paradigm shift in the manner in which processes are followed; research data requires different methods of management. The question regarding research data opens up a window for further research to understand the organisational changes required for a repository to be assessed in this manner. The information needs of the designated communities vary and can change with time. It was discovered that all the digital repositories meet the information user needs effectively. It can be concluded that all the digital repositories that were studied do comply with the data management capability infrastructure of the developed OAIS-based model.

5.2.3.4 Metadata management infrastructure

It was expected by the author that metadata management would be an area that is fully addressed by the digital repositories. This was the case for all the digital repositories used for this study. However, the different organisational structures of the institutions used for the study may require some institutions to reassess their metadata schemas in order to cater for the structural complexity of the parent organisation.

5.2.3.5 Access capability infrastructure

The data collected for this infrastructure points to a conclusion of full compliance with only a few changes required from the side of the digital repositories. It is evident (from Chapter 4) that the majority of the digital repositories deploy adequate measures and processes to facilitate access of digital material. The only grey area that was discovered is the type of evidence that each repository produces to demonstrate that it has fully addressed all aspects of usage, which might affect the trustworthiness of the repository. What may be regarded as evidence in one institution may not be regarded as such in another institution. This may also be the case with respect to the standard criteria (ISO 16363) used to assess this aspect. Apart from this minor stumbling block, it can be concluded that the studied digital repositories achieve full compliance to the access capability infrastructure.

5.2.3.6 Preservation capability infrastructure

It was discovered that preservation is viewed differently by various institutions. It can be concluded that the majority of the digital repositories investigated depend largely on the repository software for long-term preservation. This dependency caused the repository manager to give negative responses to the preservation questions that followed. This, in essence, creates a compliance gap. Preservation infrastructure is largely based on misunderstanding.

5.2.3.7 Financial sustainability infrastructure

The financial well-being and management of the investigated digital repositories are covered in the broader business planning process of the various IT units. Financial management is one of the biggest concerns that the investigated digital repositories face. Two of the repositories receive allocated funds on an ad hoc basis. This consequently has a negative implication on planning, because financial processes take long to be approved. It can be concluded that non-compliance to the ISO 16363 financial sustainability criteria could be expected in as many as 50% of the sample population.

5.2.3.8 System sustainability infrastructure

System security is an important aspect to ensure continuous viability of a digital repository. From the data gathered it is clear that all institutions rely solely on their IT unit's security support services. Where this is not the case, system security is relied upon and obtained from the repository software (DSpace), which is not an ideal mechanism to use as required by international repository assessment standards. A worrying element that was discovered in this assessment category is that digital repositories do not conduct regular risk assessments and maintain adequate security protection. Two out of the four repositories do conduct risk assessments, but the risk assessments do not take place at scheduled times – they happen haphazardly. Over and above all, the digital repositories do have security measures in place, meaning that they comply with the requirements of international repository assessment standards.

5.2.4 How should and/or how could this model be adapted for South Africa as a developing country to make the striving for trustiness more feasible?

From the data that was collected it is clear that although the South African digital repositories do not fully comply with the assessment standards; they do comply with international repository assessment standards at a certain level. Work needs to be done by the various institutions to introduce new modules to their repositories for them to reach full trust compliance. Most of the repository managers made recommendations on how the developed model can be modified to be more comprehensive of the tasks for which a digital repository is responsible. Key components that were identified were preservation and access capability.

5.3 Evaluation of the research methodology used

The main research methodology (as explained in Chapter 3) was an embedded research design that combined the quantitative and qualitative research approaches. This design intended to clearly understand the nature of South African digital repositories in terms of what international repository assessment standards require. Evaluating this study after the study had been completed revealed that it was the best methodology to use. This is because:

- The face-to-face interaction with digital repository managers helped to better explain concepts.
- The researcher was able to modify and change components of the study (the developed model) based on participant feedback and literature combined.
- The conclusions reached are based on actual feedback from participants and not assertions based on literature.

5.4 Recommendations

The following section provides a summary of the recommendations received from interview participants.

- South African digital repositories need to review the processes in place that inform sustainability.
- The financial and human resources elements of digital repositories need improvement from an organisational perspective.

- Training in the use of international repository assessment standards should be introduced in institutions.
- Digital repository institutions could conduct a feasibility study based on complying with the criteria of the developed OAIS based model.
- Similar study need to be conducted at intervals to assess the level of digital repositories working towards compliance of international repository assessment standards.

Contrary to popular belief it was established that it would not be that difficult for South African repositories to meet the international standard requirements for trustiness. It is therefore recommended that the four participating repositories should at least attempt to do a formal evaluation of their trust status.

5.5 Recommendations for further study

The following recommendations are made in relation to further studies:

- A study may be conducted to determine how full compliance with international repository assessment standards can be achieved.
- A study should be undertaken to determine what measures institutions (with digital repositories) are deploying to acclimatise compliance with international repository assessment standards.
- A similar study may be conducted to adapt and improve the OAIS-based model to thrive for the integration of complete criteria of international repository assessment standards.
- Based on the international developments of repository assessment (section 1.1) that are being done, it is necessary to investigate if the developed model can assist (as a precursor) repositories to meet the Data Seal of Approval accreditation.

5.6 Conclusion

Trust certification processes require a repository to have certain processes in place that can be critically assessed. The European Framework for Audit and Certification of Digital Repositories defined three levels of repository certification. Only the highest level of certification requires that a formal, external audit is completed and only then is certification granted in accordance with either the International Organization for

Standardization (ISO) 16363 or the Deutsches Institut für Normung. This study showed that South African institutions may not be ready for full accreditation, but that they may be closer to the target than what was previously anticipated. Repository managers should therefore consider completing the lower levels of trustiness evaluation so that they could establish what weaknesses to address.

As was seen in Chapter 2, the OAIS Framework provides common definitions of terms and means of comparison that were informed by the available international evaluation standards. Through this study it was discovered that South African institutions comply with most of the technical requirements of international repository assessment standards. The study also found that South African digital repositories are not far off in meeting the majority of the criteria and requirements set by the international standards. Finally, the study has shown that the developed OAIS-based model is a good starting point to establish a benchmark when repository managers consider conducting trust audits of their repositories.

6. References

Ambacher, B. I. 2007. Government Archives and the Digital Repository Audit Checklist, *Journal of digital information*.

Bentley, P. J. G. & Oladiram, M. 2014. The Role of Institutional Repository in Digital Scholarly Communications. [Online]. http://www.ais.up.ac.za/digi/docs/jain_paper.pdf [Accessed 11 November 2016].

Bloom, J., Sanders, B. & Lankenau, S.E.J. 2010. Putting in work: Qualitative research on substance use and other risk behaviours among gang youth in Los Angeles, *The American journal of managed care*, 45(5), pp.736-753.

Bryman, A. & Bell, E. 2007. Business Research Methods, 2n d ed. New York: *Oxford University Press*.

Burgers, T. 2001. A general introduction to the design of questionnaires for survey research, *Information Systems Services*.

Burns, N. & Grove S. K. 2011. Understanding nursing research building an evidence-based practice, 5th edi. *Elsevier Saunders*, U.S.A.

Callaghan, S., Tedds, J., Kunze, J., Khodiyar, V., Lawrence, R., Mayernik, M.S., Murphy, F., Roberts, T. & Whyte, A. 2014. Guidelines on recommending data repositories as partners in publishing research data*, International Journal of Digital Curation*, 9(1), pp. 152-163.

Christian, G. E. 2008. Issues and challenges to the development of open access institutional repositories in academic and research institutions in Nigeria. [Online]. https://idl-bnc.idrc.ca/dspace/bitstream/10625/36986/1/127792.pdf [Accessed 19 October 2016].

Cohen, D. & Crabtree B. 2006. Qualitative Research Guidelines Project. [Online]. http://www.qualres.org/HomeSemi-3629.html [Accessed 17 October 2016].

Creative Commons Attribution. 2015. WDS Certification. [Online]. https://www.icsu-wds.org/services/certification [Accessed 24 February 2016].

Crow, R. (2002). The Case for Institutional Repositories: A SPARC position paper. [Online]. http://www.arl.org/sparc/bm~doc/ir_final_release_102.pdf [Accessed 24 February 2016].

Day, M. 2007. The Reference Model for an Open Archival Information System (OAIS), *Planets and Nestor training event Vilnius*, Lithuania.

De Wee, J. A. 2013. An investigation into how mobile technologies can advance service delivery for library users at the university of Pretoria library services, Masters Thesis, Dept. of Information Technology, *University of Pretoria*.

Digital Preservation Coalition. 2016. Digital Preservation Handbook. [Online]. http://handbook.dpconline.org/glossary [Accessed 15 July 2016].

Dobratz, S., Schoger, A. & Strathmann, S. 2007. The Nestor Catalogue of Criteria for Trusted Digital Repository Evaluation and Certification, *Journal of Digital Information* 8 no. 2.

Doorn, P. 2014. Repository certification & Trusted Digital Archives, Brussels: Data Archiving and Networked Services, *Information Society Technologies*.

Downs, R.R. & Chen, R.S. 2012. Independent Evaluation of a Scientific Data Center for Compliance with the ISO 16363 Requirements for Audit and Certification of Trustworthy Digital Repositories, *Information Society Technologies*.

Farooq, U. 2013. What is Interview Schedule, Definition & Types. [Online]. http://www.studylecturenotes.com/social-research-methodology/what-is-interview-schedule-definition-types [Accessed 07 July 2016].

Farrell, S. 2016. Open-Ended vs. Closed-Ended Questions in User Research. [Online]. https://www.nngroup.com/articles/open-ended-questions/ [Accessed 03 September 2016].

Fernihough, S. 2011. e-Research: an implementation framework for South African organisations, Master's Thesis, *University of South Africa*.

Genova, F. 2015. DSA/WDS Certification of Repositories. [Online]. https://confluence.csc.fi/pages/viewpage.action?pageId=60131746 [Accessed 18 February 2016].

Gomm, R., Hammersley, M. & Foster, P., (eds.) (2000). Case study method. London: *Sage*.

Graham, W. 2001. The importance of conducting and reporting pilot studies: the example of the Scottish Births Survey, Journal of Advanced Nursing 34: 289-295.

Hanahoe, H. 2016. The growth and impact of the Research Data Alliance. [Online]. https://www.rd-alliance.org/hilary-hanahoe-growth-and-impact-research-data-alliance [Accessed 10 August 2016].

Harrel. M.C. & Bradley, M. A. 2009. Data collection methods - Semi-Structured Interviews and Focus Groups, *The Rand Corporation*.

Higgins, S. 2015. European Framework for Audit and Certification of Digital Repositories. [Online]. http://www.trusteddigitalrepository.eu/Trusted%20Digital%20Repository.html [Accessed 12 August 2015].

Hitchcock, S & Donnelly, M. 2010. Digital Preservation Tools for Repository Managers 5: Trust. At Keep it course module 5, *University of Northampton*.

Holloway, I. & Biley, F.C. 2011. Being a qualitative researcher, *Qualitative health research*, 21(7), pp. 968-975.

Houghton, B. 2015. Trustworthiness: Self-assessment of an Institutional Repository against ISO 16363-2012. [Online]. http://www.dlib.org/dlib/march15/houghton/03houghton.html [Accessed 14 September 2015].

Hsieh, H.F. & Shannon, S.E. 2005. Three approaches to qualitative content analysis, *Qualitative Health Research*, 15(9), 1277-1288.

International Association of Sound and Audio-visual Archives (IASA). 2016. Guidelines on the Production and Preservation of Digital Audio Objects. [Online] http://www.iasa-web.org/tc04/submission-information-package-sip [Accessed 26 September 2015].

Koopman, M. M. & Jager, K. 2016. Archiving South African digital research data: How ready are we?. [Online]. http://sajs.co.za/system/tdf/publications/pdf/SAJS%20112_7-8_Koopman_Research%20Article.pdf?file=1&type=node&id=35242&force= [Accessed 21 November 2016].

Kothari, C. 2004. Research Methodology Methods and Techniques, *New Age International Publishers*.

Kowalczyk, S. & Shankar, K. 2013. Data Sharing in the Sciences, *Annual Review of Information Science and Technology*, 45(1), 247-294.

ICSU-WDS. 2016. WDS Certification. [Online]. http://www.icsu-wds.org/services/certification [Accessed 21 November 2016].

International Organization for Standardization. 2015. Certification. [Online]. http://www.iso.org/iso/home/standards/certification.htm [Accessed 24 August 2015].

Lancy, D.F. 1993. Qualitative research in education: An introduction to the major traditions. New York: Longman.

Lavoie, B. R. 2004. The Open Archival Information System Reference Model: Introductory Guide, *Microform & Imaging Review*, Volume 33, Issue 2.

Leedy, P. D. & Ormrod, J. E. 2010. Practical research: Planning and design, 9th edi. *Pearson Education International*: Boston.

Marshall, M.N. 1996. Sampling for qualitative research, *Oxford University Press*.

Michael, B. 2008. Nonprobability Sampling, *Encyclopaedia of Survey Research Methods*.

Michael, C. J. 2016. 7 Key Functions of OAIS Reference Model for Digital Archiving. [Online]. http://www.consultparagon.com/blog/functions-of-oais-reference-model-for-digital-archiving [Accessed 14 July 2016].

Miles, M. & Huberman, A.M. 1994. Qualitative Data Analysis. Thousand Oaks, CA: *Sage Publications*.

Morse, J.M. 2005. What is qualitative research?, *Qualitative Health Research*, 15(7), pp. 859-860.

National Research Foundation. 2014. Statement on Open Access to Research Publications from the National Research Foundation (NRF)-Funded Research. [Online]. http://www.nrf.ac.za/media-room/news/statement-open-access-research-publications-national-research-foundation-nrf-funded [Accessed 28 August 2015].

Nestor. 2013. Nestor Seal for Trustworthy Digital Archives. [Online]. http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html [Accessed 16 June 2016].

Nicholas, D., Rowlands, I., Watkinson, A., Brown, D., Russell, B. & Jamali, H. 2013. Have digital repositories come of age? The views of library directors. [Online]. http://www.webology.org/2013/v10n2/a111.pdf [Accessed 11 November 2016].

Qasim, U. 2012. Repository Assessment Methods. [Online]. http://plnwiki.lockss.org/wiki/uploads/a/a3/Repository_Assessment_Methods.pdf [Accessed 12 January 2016].

RLG-OCLC. 2002. Trusted Digital Repositories: Attributes and Responsibilities, Mountain View, CA: RLG, May 2002.

Ross, P.S. & McHugh, M.A. 2005. The role of evidence in establishing trust in repositories. *D-Lib*, 12(7/8).

Sawyer, D. M. 2002. Framework for digital archiving: OAIS reference model, Presentation delivered at the *OCLC Steering by Standards Teleconference on the OAIS Imperative*: Enduring Record or Digital Dust.

Sibanda, N. 2009. Quantitative Research, *Victoria University Wellington Press*.

Smith, I. 2009. A Digital Preservation Strategy for the University of Pretoria using DSpace Open Source Software, African Digital Scholar-ship & Curation Conference, CSIR.

Snap, D. & Spencer, L. 2003. Qualitative Research Practice: A Guide for Social Science Students and Researcher, *SAGE Publications*.

Strauss, A.L. & Corbin, J. 1998. Basics of Qualitative Research: Grounded Theory Procedures and Techniques, 2nd edition. Thousand Oaks, CA: *Sage Publications*.

The Consultative Committee for Space Data Systems. 2011. Audit and Certification of Trustworthy Digital Repositories. [Online]. public.ccsds.org/publications/archive/652x0m1.pdf [Accessed 14 April 2016].

The Consultative Committee for Space Data Systems. 2012. Reference Model For An Open Archival Information System (OAIS). [Online]. https://public.ccsds.org/pubs/650x0m2.pdf [Accessed 14 April 2016].

The Research Data Alliance. 2016. About RDA. [Online]. https://www.rd-alliance.org/about-rda [Accessed 23 September 2016].

TRAC. 2007. Trustworthy Repositories Audit & Certification: Criteria and Checklist, *The Center for Research Libraries* (CRL).

Turner, E. O. 2015. Districts' Responses to Demographic Change: Making Sense of Race, Class, and Immigration in Political and Organizational Context, *American Educational Research Journal*.

UK Data Archive. 2016. Standards of Trust/DIN 31644. [Online]. http://www.data-archive.ac.uk/curate/trusted-digital-repositories/standards-of-trust?index=3 [Accessed 06 January 2016].

University of Glasgow. 2016. Audit and certification. [Online]. http://handbook.dpconline.org/institutional-strategies/audit-and-certification [Accessed 14 June 2016].

Van Teijlingen, E. 2014. Semi-structured interviews. [Online]. https://intranetsp.bournemouth.ac.uk/documentsrep/PGR%20Workshop%20-%20Interviews%20Dec%202014.pdf [Accessed 14 October].

Van Wyk, B. 2008. Research design and methods Part I, *University of Cape Town Press.*

WebFinance, Inc. 2015. Business Dictionary. [Online]. http://www.businessdictionary.com/ [Accessed 12 August 2015].

Whyte, A., Molloy, L., Beagrie, N. & Houghton, J. 2014. What to measure? Toward metrics for Research data management. Research data management: practical

strategies for information professionals, *Purdue University Press*, West Lafayette, 275-300.

Wikipedia. 2016. Open Archival Information System. [Online]. https://en.wikipedia.org/wiki/Open_Archival_Information_System [Accessed 07 October 2016].

Wood, J. 2015. Research Data Alliance. [Online]. www.rd-alliance.org [Accessed 12 January 2016].

World Data System. 2015. WDS Members Forum 2016. [Online] https://www.icsu-wds.org/community/wds-members-forum/2016-members-forum [Accessed 26 November 2016].

World Data System. 2016. Trusted Data Services for Global Science. [Online]. https://www.icsu-wds.org/community/membership/community/membership/regular-members [Accessed 14 June 2016].

World Data System. 2015. WDS Annual Report. [Online] https://www.icsu-wds.org/news/news-archive/annual-report-2015-16 [Accessed 26 November 2016].

Yin, R.K. 2013. Case study research: Design and methods. *Sage publications*.

Annexure 1: Interview Schedule

To develop a South African Digital Repository Trust Assessment Model based on International Standards to measure trustiness of Digital Repositories Introduction by the researcher:

This interview is conducted as a requirement for a Master's Degree in Information Technology presented by the University of Pretoria's Department of Information Science. The researcher, Glenn Tshweu, is investigating a working framework (model) that could be used to assess South African repositories for trust status according to international standards, based on the internationally accepted OAIS-model. It is anticipated that adhering to the proposed model could ensure that a repository is seen as a trusted one. This interview is intended to gain feedback from you about the developed model and to establish if you have suggestions for improvement of the model.

Base principal: EVERY repository being established should be evaluated for its level of trustiness.

Questions for each infrastructure aspect as requirements for full accreditation

Every infrastructure aspect of the proposed model addresses a different concept; the table below provides a roadmap of the questions that an institution will answer when assessing whether or not their repository complies with the criteria of the given standard. The OAIS-model will be shown and discussed with the respondents.

The following questions originate from the three standards (Trustworthy Repositories Audit & Certification (TRAC), the International Organization for Standardization (ISO) 16363, and the German Institute of Standardization (DIN) 31644) that were used to develop the proposed model. Each question is grouped in the standards where it originates. Annexure 1 is the synthesis of the OAIS-based model with the linkages of the TRAC, ISO 16363 and DIN 31644 standards. Annexure 2 is a synthesis of the literature, together with the synthesis model of the questions.

Please rate your organisation on a scale between 1– 4, and provide a comment to support each rating given.

(1 = Strongly Disagree; 2 = Disagree; 3 = Agree; 4 = Strongly Agree):

Standard	Question(s)	Rating	Comment
1. Questions relating to the repository's Administration capability:			
1.1) DIN 31644: C9 Personnel	- Does the organisation have sufficient and qualified staff members available to manage the repository? - Are the required qualifications set out including an organisational chart? - Is there a staff development plan that outlines the tasks and objectives of the repository?	1 2 3 4 1 2 3 4 1 2 3 4	
1.2) DIN 31644: C20 Technical Authority	- Does the repository have processes in place that ensure authority control on a permanent basis without technical restrictions (e.g. encryption, copy and print protection)?	1 2 3 4	
1.3) ISO 16363: Contracts, Licenses, and Liabilities	- Does the repository have and maintain appropriate contracts or deposit agreements for digital materials that it manages, preserves, and/or to which it provides access? - Are formal deposits and contracts legitimate, i.e. are they countersigned and current?	1 2 3 4 1 2 3 4	

2. Questions relating to the repository's Ingest capability:			
2.1) DIN 31644: C1 Selection of information objects and their representations	- Does the repository have a criterion that defines the selection of information objects and their representations? - Is the selection of digital information transparently documented on the basis of criteria, guidelines and profiles?	1 2 3 4 1 2 3 4	
2.2) DIN 31644: C21 Submission information packages	- Does the repository adequately specify the composition of data packages for data transfer? - Does the digital archive have specifications regarding content data that is accepted, and the metadata required?	1 2 3 4 1 2 3 4	
2.3) ISO 16363 Ingest: Acquisition of content	- Can the repository demonstrate (to funders, depositors, and users) what responsibilities it is taking on and what aspects are excluded? - Can the repository determine and check what the characteristics and properties of preserved items will be used over the long term? Example: Are there definitions of the information properties which should be preserved; submission agreements/deposit agreements, preservation policies, written processing procedures, workflow documentation?	1 2 3 4 1 2 3 4	
3. Questions relating to the repository's Data Management Functional capability:			
	Does the repository contain data or intend to integrate and manage research data in the future?	1 2 3 4	
3.1) ISO 16363 Information Management: Specify minimum information requirements to discover and identify material of interest	- Is the repository able to deal with the different types of information requests made by users from the designated community? - Does the repository have adequate Retrieval and descriptive information, discovery metadata (such as Dublin Core), and other documentation describing information objects to be retrieved?	1 2 3 4 1 2 3 4	
3.2) ISO 16363 Information Management: Capture or create minimum descriptive information associated with the Archival Information Package.	- Can the repository deal with the types of requests that come from a typical user from the designated community?	1 2 3 4	

3. Questions relating to the repository's Data Management Functional capability (concluded):			
3.3) ISO 16363 Information Management: Maintain bi-directional linkage between each AIP and its descriptive information	- Can all the Archival Information Package (AIPs) be located and retrieved? Are there adequate measures (such as descriptive metadata; unique, persistent identifier) in place to ensure that AIP are located and retrieved. (Please see Glossary for the AIP definition.) - Is there a procedure in place that notifies when the relationship between the data and the associated descriptive information is temporarily broken to ensure that it can be restored?	1 2 3 4 1 2 3 4	
3.4) TRAC - Chain of Custody: Have physical and legal control over the existence, authenticity, location, and accessibility of records	- Can the repository demonstrate the chain of custody for all of its digital content from the point of deposit? - Can the repository demonstrate that the content it has matches the content it received?	1 2 3 4 1 2 3 4	
4. Questions relating to the repository's Metadata management:			
4.1) DIN 31644: C5 Interpretability	- Does the repository have a metadata schema that it uses? - Can the repository ensure long-term interpretability of at least one representation of content data and metadata? - Does the repository have methods to allow the user community to check interpretability on a regular basis?	1 2 3 4 1 2 3 4 1 2 3 4	
5. Questions relating to the repository's Access capability:			
5.1) ISO 16363: Access Management - Comply with Access Policy	- Is the repository able to produce evidence to demonstrate that it has fully addressed all aspects of usage which might affect the trustworthiness of the repository?	1 2 3 4	
5.2) ISO 16363 Access Management – Dissemination of original digital objects with evidence	- Does the repository follow policies and procedures that enable the dissemination of digital objects that are traceable to the originals, with evidence supporting their authenticity?	1 2 3 4	
5.3) DIN 31644: C4 Access	- Can the repository ensure authorized access to information for the designated community? - Does the repository have appropriate search possibilities, which indicates the terms of use and restriction? - Does the repository declare its condition of use and costs that may arise?	1 2 3 4 1 2 3 4 1 2 3 4	

6. Questions relating to the repository's Preservation capability:			
6.1) Preservation Period Determination Phase; Nominate institution to offer Long-term preservation	- Can the repository offer long-term preservation service? (Long-term preservation refers to continued access to digital materials, or at least to the information contained in them, indefinitely. Please refer to the Glossary for the definition of short, medium and long-term preservation.) (If answer is 'yes', ask respondent to elaborate.)	1 2 3 4	
6.2) ISO 16363: Governance and Organizational Viability	- Does the parent organisation or the repository's mission statement explicitly address preservation?	1 2 3 4	
6.3) DIN 31644 C18: Authenticity - Preservation measures	[*Ask this question only if a repository provides long-term preservation*] - Does the repository deploy methods which ensure the authenticity of the objects during implementation of the long-term preservation measures and document the degree of authenticity? - Does the repository ensure that relevant information objects retain their authenticity while undergoing preservation processes and that all measures are transparently and permanently documented?	1 2 3 4 1 2 3 4	
6.4) TRAC: Authenticity	- As part of preservation, does the repository assert provenance of digital objects by guaranteeing that objects are stored intact as it has been created?	1 2 3 4	
7. Questions relating to the repository's Financial sustainability:			
7.1) ISO 16363: Financial Sustainability	- Does the repository have a short and long-term business planning process in place to sustain the repository over time? (An annual business planning process is commonly accepted as the standard for most Organizations.)	1 2 3 4	
8. Questions relating to the repository's System Security sustainability:			
8.1) DIN 31644: C34 Security	- Can the organisation and the infrastructure protect the repository and its information objects and representations to ensure integrity? - Does the repository have suitable measures to protect its own integrity and that of its digital assets to ensure that the assets remain intact and to fulfil its legal or contractual obligations? Example: Measures have to be implemented that are informed by IT security systems.	1 2 3 4 1 2 3 4	
8.2) ISO 16363 Security Risk Management	- Does the repository conduct regular risk assessments and maintain adequate security protection in order to provide expected and contracted levels of service? - Can the repository show how it deals with its security requirements?	1 2 3 4 1 2 3 4	

General - additional questions for clarification:

Please consider the model that I have developed (refer to Annexure 2).

1) From a manger's point of view, what are the strengths and weaknesses of the model?
2) Which component of the model is most important for your institution?
3) Do you have any recommendations that can be used in redesigning the model?
4) To what extent do you see your institution being willing to comply with the requirements of such a system of trust?
5) What level of trust would your institution be able and willing to comply with, in terms of rating?
6) Which of the levels of the model need revision and what should the revision include?
7) From the developed model, what components did I forget to include?

Annexure 2: Assessment criteria – ISO, DIN and TRAC

Standards	Criteria
TRAC	**Section A: Organizational Infrastructure**
	A1. Governance & organizational viability
	A1.1 Repository has a mission statement that reflects a commitment to the long-term retention of, management of, and access to digital information.
	A1.2 Repository has an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.
	A2. Organizational structure & staffing
	A2.1 Repository has identified and established the duties that it needs to perform and has appointed staff with adequate skills and experience to fulfil these duties.
	A2.2 Repository has the appropriate number of staff to support all functions and services**.**
	A2.3 Repository has an active professional development program in place that provides staff with skills and expertise development opportunities.
	A3. Procedural accountability & policy framework
	A3.1 Repository has defined its designated community(ies) and associated knowledge base(s) and has publicly accessible definitions and policies in place to dictate how its preservation service requirements will be met.
	Evidence: Mission statement; written definitions of the designated community(ies); documented policies; service-level agreements.
	A3.2 Repository has procedures and policies in place, and mechanisms for their review, update, and development as the repository grows and as technology and community practice evolves.
	Evidence: Written documentation in the form of policies, procedures, protocols, rules, manuals, handbooks, and workflows; specification of review cycle for documentation; documentation detailing review, update, and development mechanisms. If documentation is embedded in system logic, functionality should demonstrate the implementation of policies and procedures.
	A3.3 Repository maintains written policies that specify the nature of any legal permissions required to preserve digital content over time, and repository can demonstrate that these permissions have been acquired when needed.
	A3.4 Repository is committed to formal, periodic review and assessment to ensure responsiveness to technological developments and evolving requirements.
	A3.5 Repository has policies and procedures to ensure that feedback from producers and users is sought and addressed over time.
	A3.6 Repository has a documented history of the changes to its operations, procedures, software, and hardware that, where appropriate, is linked to relevant preservation strategies and describes potential effects on preserving digital content.

Standards	Criteria
TRAC	A3.7 Repository commits to transparency and accountability in all actions supporting the operation and management of the repository, especially those that affect the preservation of digital content over time.
	A3.8 Repository commits to defining, collecting, tracking, and providing, on demand, its information integrity measurements.
	A3.9 Repository commits to a regular schedule of self-assessment and certification and, if certified, commits to notifying certifying bodies of operational changes that will change or nullify its certification status.
	A4. Financial sustainability
	A4.1 Repository has short- and long-term business planning processes in place to sustain the repository over time.
	A4.2 Repository has in place processes to review and adjust business plans at least annually.
	A4.3 Repository's financial practices and procedures are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements.
	A4.4 Repository has ongoing commitment to analyze and report on risk, benefit, investment, and expenditure (including assets, licenses, and liabilities).
	A4.5 Repository commits to monitoring for and bridging gaps in funding.
	A5. Contracts, licenses, & liabilities
	A5.1 If repository manages, preserves, and/or provides access to digital materials on behalf of another organization, it has and maintains appropriate contracts or deposit agreements.
	A5.2 Repository contracts or deposit agreements must specify and transfer all necessary preservation rights, and those rights transferred must be documented.
	A5.3 Repository has specified all appropriate aspects of acquisition, maintenance, access, and withdrawal in written agreements with depositors and other relevant parties.
	A5.4 Repository tracks and manages intellectual property rights and restrictions on use of repository content as required by deposit agreement, contract, or license.
	A5.5 If repository ingests digital content with unclear ownership/rights, policies are in place to address liability and challenges to those rights.

Standards	Criteria
TRAC	**Section B: Digital Object Management**
	B1. Ingest: acquisition of content
	B1.1 Repository identifies properties it will preserve for digital objects.
	B1.2 Repository clearly specifies the information that needs to be associated with digital material at the time of its deposit (i.e. SIP).
	B1.3 Repository has mechanisms to authenticate the source of all materials.
	B1.4 Repository's ingest process verifies each submitted object (i.e. SIP) for completeness and correctness as specified in B1.2.
	B1.5 Repository obtains sufficient physical control over the digital objects to preserve them.
	B1.6 Repository provides producer/depositor with appropriate responses at predefined points during the ingest processes.
	B1.7 Repository can demonstrate when preservation responsibility is formally accepted for the contents of the submitted data objects (i.e. SIPs).
	B1.8 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Ingest: content acquisition).
	B2. Ingest: creation of the archival package
	B2.1 Repository has an identifiable, written definition for each AIP or class of information preserved by the repository.
	B2.2 Repository has a definition of each AIP (or class) that is adequate to fit long-term preservation needs.
	B2.3 Repository has a description of how AIPs are constructed from SIPs.
	B2.4 Repository can demonstrate that all submitted objects (i.e. SIPs) are either accepted as whole or part of an eventual archival object (i.e. AIP), or otherwise disposed of in a recorded fashion.
	B2.5 Repository has and uses a naming convention that generates visible, persistent, unique identifiers for all archived objects (i.e. AIPs).
	B2.6 If unique identifiers are associated with SIPs before ingest, the repository preserves the identifiers in a way that maintains a persistent association with the resultant archived object (e.g. AIP).
	B2.7 Repository demonstrates that it has access to necessary tools and resources to establish authoritative Representation Information of the digital objects it contains.
	B2.8 Repository records/registers Representation Information (including formats) ingested.
	B2.9 Repository has documented processes for acquiring preservation metadata (i.e. PDI) for its associated Content Information and acquires preservation metadata in accordance with the documented processes. The repository must maintain viewable documentation on how the repository acquires and manages Preservation Description Information (PDI).
	B2.10 Repository has a documented process for testing understandability of the information content and bringing the information content up to the agreed level of understandability.
	B2.11 Repository verifies each AIP for completeness and correctness at the point it is generated.
	B2.12 Repository provides an independent mechanism for audit of the integrity of the repository collection/content.
	B2.13 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (AIP creation).

Standards	Criteria
TRAC	**B3. Preservation planning**
	B3.1 Repository has documented preservation strategies.
	B3.2 Repository has mechanisms in place for monitoring and notification when Representation Information (including formats) approaches obsolescence or is no longer viable.
	B3.3 Repository has mechanisms to change its preservation plans as a result of its monitoring activities.
	B3.4 Repository can provide evidence of the effectiveness of its preservation planning.
	B4. Archival storage & preservation/maintenance of AIPs
	B4.1 Repository employs documented preservation strategies.
	B4.2 Repository implements/responds to strategies for archival object (i.e. AIP) storage and migration.
	B4.3 Repository preserves the Content Information of archival objects (i.e. AIPs).
	B4.4 Repository actively monitors integrity of archival objects (i.e. AIPs).
	B4.5 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Archival Storage).
	B5. Information management
	B5.1 Repository articulates minimum metadata requirements to enable the designated community(ies) to discover and identify material of interest.
	B5.2 Repository captures or creates minimum descriptive metadata and ensures that it is associated with the archived object (i.e. AIP).
	B5.3 Repository can demonstrate that referential integrity is created between all archived objects (i.e. AIPs) and associated descriptive information.
	B5.4 Repository can demonstrate that referential integrity is maintained between all archived objects (i.e. AIPs) and associated descriptive information.
	B6. Access management
	B6.1 Repository documents and communicates to its designated community(ies) what access and delivery options are available.
	B6.2 Repository has implemented a policy for recording all access actions (includes requests, orders etc.) that meet the requirements of the repository and information producers/depositors.
	B6.3 Repository ensures that agreements applicable to access conditions are adhered to.
	B6.4 Repository has documented and implemented access policies (authorization rules, authentication requirements) consistent with deposit agreements for stored objects.
	B6.5 Repository access management system fully implements access policy.
	B6.6 Repository logs all access management failures, and staff review inappropriate "access denial" incidents.
	B6.7 Repository can demonstrate that the process that generates the requested digital object(s) (i.e. DIP) is completed in relation to the request.
	B6.8 Repository can demonstrate that the process that generates the requested digital object(s) (i.e. DIP) is correct in relation to the request.

Standards	Criteria
TRAC	B6.9 Repository demonstrates that all access requests result in a response of acceptance or rejection
	B6.10 Repository enables the dissemination of authentic copies of the original or objects traceable to originals.
	C. Technologies, Technical Infrastructure, & Security
	C1. System Infrastructure
	C1.1 Repository functions on well-supported operating systems and other core infrastructural software.
	C1.2 Repository ensures that it has adequate hardware and software support for backup functionality sufficient for the repository's services and for the data held, e.g. metadata associated with access controls, repository main content.
	C1.3 Repository manages the number and location of copies of all digital objects.
	C1.4 Repository has mechanisms in place to ensure any/multiple copies of digital objects are synchronized.
	C1.5 Repository has effective mechanisms to detect bit corruption or loss.
	C1.6 Repository reports to its administration all incidents of data corruption or loss, and steps taken to repair/replace corrupt or lost data.
	C1.7 Repository has defined processes for storage media and/or hardware change (e.g. refreshing, migration).
	C1.8 Repository has a documented change management process that identifies changes to critical processes that potentially affect the repository's ability to comply with its mandatory responsibilities.
	C1.9 Repository has a process for testing the effect of critical changes to the system.
	C1.10 Repository has a process to react to the availability of new software security updates based on a risk-benefit assessment.
	C2. Appropriate technologies
	C2.1 Repository has hardware technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when hardware technology changes are needed.
	C2.2 Repository has software technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when software technology changes are needed.
	C3. Security.
	C3.1 Repository maintains a systematic analysis of such factors as data, systems, personnel, physical plant, and security needs.
	C3.2 Repository has implemented controls to adequately address each of the defined security needs.
	C3.3 Repository staff have delineated roles, responsibilities, and authorizations related to implementing changes within the system.
	C3.4 Repository has suitable written disaster preparedness and recovery plan(s), including at least one off-site backup of all preserved information together with an off-site copy of the recovery plan(s).

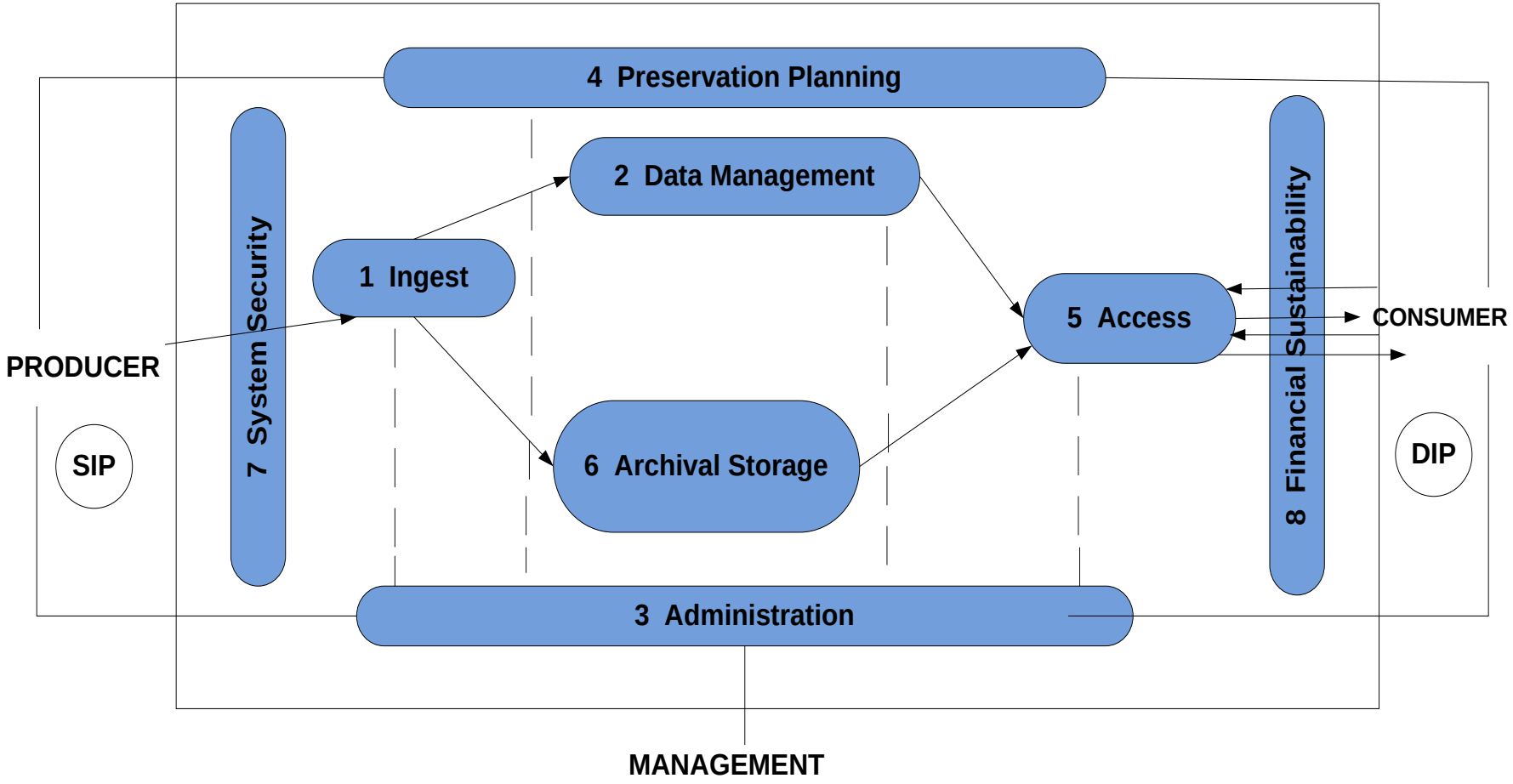
Standards	Criteria
DIN 31644	**Organization**
	C3 Designated communities
	C5 Interpretability
	C8 Funding
	C9 Personnel
	C10 Organisation and processes
	C20 Technical authority
	Management of intellectual entities and their representations
	C1 Selection of information objects and their representations
	C2 Responsibility for preservation
	C4 Access
	C11 Preservation measures
	C12 Crisis/successorship management
	C13 Significant properties
	C14 Integrity: Ingest interface
	C15 Integrity: Functions of the archival storage
	C16 Integrity: User interface
	C17 Authenticity: Ingest
	C18 Authenticity: Preservation measures
	C 19 Authenticity: Use
	C21 Transfer packages
	C 22 Transformation of the transfer packages into archival packages
	C 23 Archival packages
	C 24 Interpretability of the archival packages
	C 25 Transformation of archival packages into access packages
	C26 Access packages
	C27 Identification
	C28 Descriptive metadata
	C29 Structural metadata
	C30 Technical metadata
	C32 Administrative metadata
	C31 Logging the preservation measures

Standards	Criteria
DIN 31644	**Infrastructure and security**
	C6 Legal and contractual basis
	C7 Legal conformity
	C33 IT infrastructure
	C34 Security
	Organizational Infrastructure
ISO 16363	**Governance and Organizational Viability**
	The repository shall have a mission statement that reflects a commitment to the preservation of, long-term retention of, management of, and access to digital information.
	The repository shall have a Preservation Strategic Plan that defines the approach the repository will take in the long-term support of its mission.
	The repository shall have a Collection Policy or other document that specifies the type of information it will preserve, retain, manage, and provide access to.
	Organizational Structure and Staffing
	The repository shall have identified and established the duties that it needs to perform and shall have appointed staff with adequate skills and experience to fulfil these duties.
	Procedural Accountability and Preservation Policy Framework
	The repository shall have defined its designated community and associated knowledge base(s) and shall have these definitions appropriately accessible.
	The repository shall have preservation policies in place to ensure its preservation strategic plan will be met.
	The repository shall have a documented history of the changes to its operations, procedures, software, and hardware.
	The repository shall commit to transparency and accountability in all actions supporting the operation and management of the repository that affect the preservation of digital content over time.
	The repository shall define, collect, track, and appropriately provide its information integrity measurements.
	The repository shall commit to a regular schedule of self-assessment and external certification.
	Financial Sustainability
	The repository shall have short- and long-term business planning processes in place to sustain the repository over time.
	The repository shall have financial practices and procedures which are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements.
	The repository shall have an ongoing commitment to analyse and report on financial risk, benefit, investment, and expenditure (including assets, licenses, and liabilities).

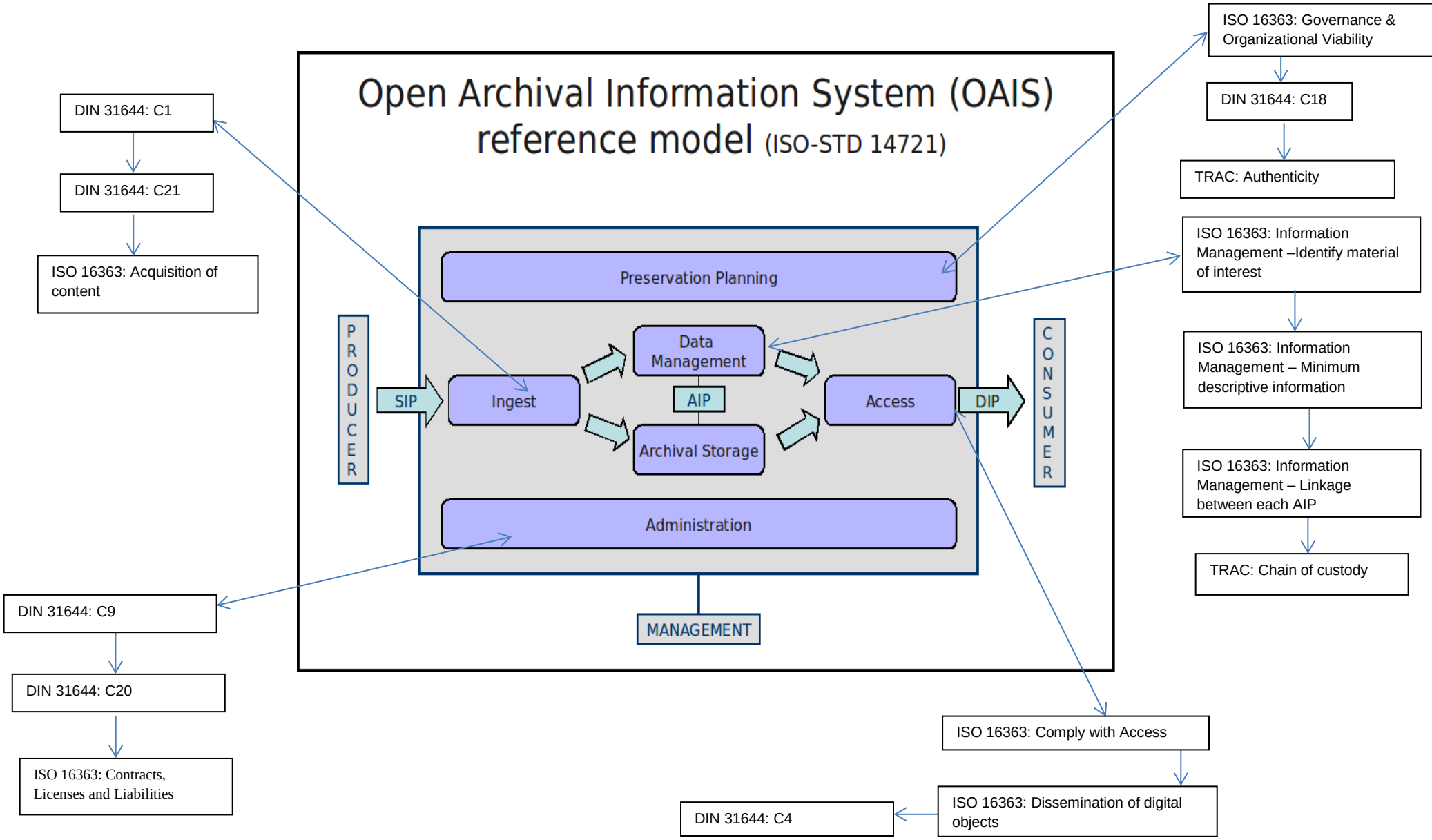
Standards	Criteria
ISO 16363	**Contracts, Licenses, and Liabilities**
	The repository shall have and maintain appropriate contracts or deposit agreements for digital materials that it manages, preserves, and/or to which it provides access.
	The repository shall track and manage intellectual property rights and restrictions on use of repository content as required by deposit agreement, contract, or license.
	Ingest: Acquisition of content
	The repository shall identify the Content Information and the Information Properties that the repository will preserve.
	The repository shall clearly specify the information that needs to be associated with specific Content Information at the time of its deposit.
	The repository shall have adequate specifications enabling recognition and parsing of the SIPs.
	The repository shall have mechanisms to appropriately verify the identity of the Producer of all materials.
	The repository shall have an ingest process which verifies each SIP for completeness and correctness.
	The repository shall obtain sufficient control over the Digital Objects to preserve them.
	The repository shall provide the producer/depositor with appropriate responses at agreed points during the ingest processes.
	The repository shall have contemporaneous records of actions and administration processes that are relevant to content acquisition.
	Ingest: Creation of the Archival Information Package (AIP)
	The repository shall have for each AIP or class of AIPs preserved by the repository an associated definition that is adequate for parsing the AIP and fit for long-term preservation needs.
	The repository shall have a description of how AIPs are constructed from Submission Information Package (SIPs).
	The repository shall document the final disposition of all SIPs.
	The repository shall have and use a convention that generates persistent, unique identifiers for all AIPs.
	The repository shall have access to necessary tools and resources to provide authoritative Representation Information for all of the digital objects it contains.
	The repository shall have documented processes for acquiring Preservation Description Information (PDI) for its associated Content Information and acquire PDI in accordance with the documented processes.
	The repository shall ensure that the Content Information of the AIPs is understandable for their designated community at the time of creation of the AIP.
	The repository shall verify each AIP for completeness and correctness at the point it is created.
	The repository shall provide an independent mechanism for verifying the integrity of the repository collection/content.
	The repository shall have contemporaneous records of actions and administration processes that are relevant to AIP creation.
	Preservation Planning
	The repository shall have documented preservation strategies relevant to its holdings.
	The repository shall have mechanisms in place for monitoring its preservation environment.
	The repository shall have mechanisms to change its preservation plans as a result of its monitoring activities.
	The repository shall provide evidence of the effectiveness of its preservation activities.

Standards	Criteria
ISO 16363	**Archival Information Package (AIP) Preservation**
	The repository shall have specifications for how the AIPs are stored down to the bit level.
	The repository shall have contemporaneous records of actions and administration processes that are relevant to storage and preservation of the AIPs.
	Information Management
	The repository shall specify minimum information requirements to enable the designated community to discover and identify material of interest.
	The repository shall capture or create minimum descriptive information and ensure that it is associated with the AIP.
	The repository shall maintain bi-directional linkage between each AIP and its descriptive information.
	Access Management
	The repository shall comply with Access Policies.
	The repository shall follow policies and procedures that enable the dissemination of digital objects that are traceable to the originals, with evidence supporting their authenticity.
	Technical Infrastructure Risk Management
	The repository shall identify and manage the risks to its preservation operations and goals associated with system infrastructure.
	The repository shall have adequate hardware and software support for backup functionality sufficient for preserving the repository content and tracking repository functions.
	The repository shall have effective mechanisms to detect bit corruption or loss.
	The repository shall have a process to record and react to the availability of new security updates based on a risk-benefit assessment.
	The repository shall have defined processes for storage media and/or hardware change (e.g. refreshing, migration).
	The repository shall have identified and documented critical processes that affect its ability to comply with its mandatory responsibilities.
	The repository shall manage the number and location of copies of all digital objects.
	Security Risk Management
	The repository shall maintain a systematic analysis of security risk factors associated with data, systems, personnel, and physical plant.
	The repository shall have implemented controls to adequately address each of the defined security risks.
	The repository staff shall have delineated roles, responsibilities, and authorizations related to implementing changes within the system.
	The repository shall have suitable written disaster preparedness and recovery plan(s), including at least one off-site backup of all preserved information together with an offsite copy of the recovery plan(s).

Annexure 3: Adapted OAIS-based model



Annexure 4: Synthesis of international repository standards



Annexure 5: Consent form

Informed Consent Form for repository/centre managers

I am Glenn Tshweu, a Master's student at the University of Pretoria.

The concept 'trusted repository' is linked to both an ISO and a DIN standard. There are several other measures of 'trustiness' (for example TRAC and WDS) but they all focus on individual repositories gaining full trusted repository status. I am conducting a research project entitled **An investigation into the extent to which South African repositories comply with international trust standards**. I have developed a framework (model) that I plan to promote for establishing a level of trust for South Africa's repositories. The intension today is to, within the South African context; gain your opinion regarding the workability of a framework that will focus on repositories acquiring trusted status based on international standards.

1. I confirm that I have read and understand the Plain Language Statement for the above study and have had the opportunity to ask questions.
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason.
3. I hereby give permission to being voice recorded and the collected data to be curated for the research on 'An investigation into a model for evaluating South African research data repositories'. I understand that I am participating out of my own will and without being forced in any way to do so. I also understand that I can stop participating at any point should I wish not to continue and that this decision will not in any way affect me negatively. I understand that my participation will remain confidential, and that any research reports or related voice recordings to recipients outside the researcher will not reveal my identity.
4. I agree/do not agree (delete as applicable) to take part in the above study.

_____	_____	_____
Name of participant	*Date*	*Signature*
_____	_____	_____
Researcher	*Date*	*Signature*