The role of academic libraries in implementing research data services: a case study of the University of KwaZulu-Natal Libraries

Zizipho Madibi MDBZIZ001

A minor dissertation submitted in *partial fulfillment* of the requirements for the award of the degree of Master of Philosophy specializing in Digital Curation

Faculty of the Humanities

University of Cape Town

[2021]

COMPULSORY DECLARATION

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signed by candidate

Signature:_____ Date:_____

Dedication

To my mom – this is for you Wawa. Continue resting maSango, you have done well.
To my daughter, Lazola “Matie” Madibi – I love you.

Table of Contents

Acknowledgements

Now unto Him that is able to keep me from falling, and to present me faultless before the presence of His glory with exceeding joy, to the only wise God my Saviour, be glory and majesty, dominion and power, both now and ever. Amen!

To my supervisor, Michelle Kahn, for her patience, guidance, understanding, mentoring, re-assurance and motivation, I will forever be grateful. This study would have never been successful without her. Thank you.

My gratitude also goes to the University of KwaZulu-Natal for granting me permission to collect data as well as UKZN researchers and library staff members for participating in the study.

To my family: my kids Lazola, Lonwabo, Bukhobenkosi, Ntandoyenkosi and Hlelolwenkosi - my siblings, Luleka, Biggy, Banele and Songezo – you have inspired this study in ways you can never understand. I am grateful and appreciative of your love and support. Thank you for being my family.

To my colleagues and friends whom have been my support during the study – I am thankful.

Abstract

This study investigated the role of academic libraries in implementing research data services, UKZN being the case study. The objectives of the study were to identify the need for research data services among UKZN researchers, to identify the major challenges associated with introducing research data services at UKZN, and to determine the possibility of implementing research data services at UKZN Libraries.

The Data Curation Centre Lifecycle model was adopted as a framework for the study as it manages to connect the different stages of research data management. The study took a mixed methods approach of which interviews and a survey were used. A purposive sample was used to select library staff and random sample was drawn from 1341 UKZN academics. From a sample of 1341, 299 was the minimum size recommended by the Raosoft sample size calculator for a 5% margin of error and 95% confidence level.

For quantitative analysis, an online questionnaire was administered using Google Forms. A series of questions were formulated for guidance in obtaining answers to the study objectives. Google Forms was used for the analysis while figures and tables were created using Microsoft Excel. Interviews from the library staff were recorded and data from interviews was transcribed into Microsoft Word.

The study revealed that UKZN Libraries are still struggling with RDM policy development. The findings of the study revealed that researchers who responded to the study showed a lack of RDM awareness while library staff showed a moderate level of awareness. The study revealed that researchers at UKZN work with different types of data and they use different storage options such as removable storage devices, computer hard drives and cloud services. Although a few researchers have developed data management plans at UKZN, they have not done so because they were mandated by the institution - UKZN has not yet developed DMPs and library staff are not aware which funders require DMPs.

The researchers who responded to the study showed interest in different trainings such as, training on data storage, development of DMPs and metadata creation. The library staff were more eager to provide data storage, data archiving and sharing mainly because of the existence of the UKZN data repository (Yabelana).

Study recommendations are based on the analysed data. One of the recommendations was that UKZN Libraries should assume a role of being an advisor and trainer for research data services at UKZN.

Keywords: Research data management, research data services, awareness, challenges, implementation, data.

List of acronyms and abbreviations

OA – Open Access

RDM – Research Data Management

RDS – Research Data Services

UKZN – University of KwaZulu-Natal

UCT – University of Cape Town

NRF – National Research Foundation

DCC – Digital Curation Centre

DOI – Digital Object Identifier

DMP – Data Management Plan

USA – United States of America

UK – United Kingdom

UCL - University College London

NSF - National Science Foundation

HEI - Higher Education Institutions

URAP - University Ranking by Academic Performance

DVC - Deputy Vice Chancellor

ACRL - The Association of College & Research Libraries

IR - Institutional Repositories

EPSRC - Engineering and Physical Sciences Research Council

OECD - Organisation for Economic Co-operation and Development

ARL - Association of Research Libraries

CARDIO - Collaborative Assessment of Research Data Infrastructure and Objectives

DAF - Data Asset Framework

AIDA - Assessing Institutional Digital Assets

VCUL - Virginia Commonwealth University Library

JHU - Johns Hopkins University

List of figures

List of tables

Chapter one: Introduction to the study

1.1. Introduction

Research and knowledge creation are an essential part of university life (Allan, 2010:3). Progress in scientific research is becoming increasingly more intense, and the documentation of academic activity is becoming more extensive (Bryant, Lavoie & Malpas, 2017:5), stretching beyond the mere publishing of research outcomes in monographs, journal articles, and conference proceedings, to include a scope of research outputs produced both during the process of research and afterwards. The primary example of these is research data (Bryant, Lavoie & Malpas, 2017:5).

According to Cox and Pinfield (2014:299), advances in digital technology across disciplines allow an accelerating number of researchers to produce and utilise large datasets as part of the research process. These datasets are valuable. In many disciplines, retention and long-term curation of data sets is becoming part of scholarly practice by means of which to enable verification of research and to promote data reuse for new research (Bryant, Lavoie & Malpas, 2017:5).

Countries such as the United Kingdom (UK), the United States of America (USA), and Australia now require applicants to provide research data management and sharing plans when applying for grants so as to acknowledge the potential value of research data and to provide a return on public investment (Kahn et al., 2014:296). It is believed that research data that are publicly funded are of public interest, and ought to be made accessible to the public with minimum restrictions (University of Cambridge, 2018). South Africa's National Research Foundation (NRF) (National Research Fund [NRF], 2015) declared that the support of research through public funding is essential for growing our knowledge economy, advancing innovation and stimulating development. According to the NRF (NRF, 2015), the publication of NRF-funded research, including data that supports that research, contributes to the knowledge base of our country.

University libraries offer research support services and, with the growth of interest in datasets and their sharing, they are increasingly offering services in support of research data management. This study investigated the role of the library in the implementation of research data services (RDS), using the University of KwaZulu-Natal (UKZN) as a case study. A lifecycle model served as a framework, as it can represent the connection between the stages

of research data management and developing support services (Carlson 2014:66). Higgins (2008) regards a lifecycle model as essential, as it assists in ensuring that all the important stages of data management are clear, understandable, and planned for, before the research project begins, so as to ensure data can be shared and, where needed, preserved in the long-term. For better understanding of the study and for the purposes of introducing the reader to the topic, subtopics will be used to present some specific details about the study.

1.1.1. Data sharing and its benefits

According to the University College London (UCL) (University College London [UCL], 2018), after the life of a research project, research data continues to be useful. Data sharing by everyone – not only those whose funders required it – has the potential to open new research approaches without researchers having to collect the same data.

Tenopir et al. (2011:1) and the University of Bristol agreed that data sharing is associated with the use and reuse of data. The many benefits associated with data sharing include:

- possibility of re-analysis of data that assists with the verification of findings;
- no need for data re-collection when data is available;
- possibility of different interpretations from the same data;
- protection of research from fabrication and falsification;
- improvement of data security and decreased risks associated with data loss (Tenopir et al. 2011:1);
- ensured accountability making data available to the entire public, and showing the accountable use of public funds;
- maximises research efficiency (data can be accessed easily);
- provides benefits such as adhering to funder and regulatory requirements; and
- adherence to funder and regulatory requirements could lead to further funding (University of Bristol, 2012).

Many of these benefits can only be realised if data are managed well. The management and sharing of data benefits researchers, the scientific community, academic institutions, and the public at large (Strasser et al., 2012). For researchers, good data management potentially leads to maximised data visibility, research of a higher quality, and improved citation rates (Jones, Pryor & Whyte, 2013:2). In South Africa, researchers and those supporting research are

progressively becoming aware of the importance of sharing research data (Kahn et al., 2014:296), and therefore of the need to manage data properly.

1.1.2. Research data management

Donnelly (2012:6) defines research data management (RDM) as the "active management and appraisal of data over the lifecycle of scholarly and scientific interest". According to Oxford University (2018), RDM is a general expression used to incorporate the following activities: organising, planning, creating, structuring, as well as storing and caring for the information generated during a research project.

A global agenda around RDM has emerged over the last few decades (Cox et al. 2017:2182). This emergence was motivated by a variety of aspects, including the rise in data intensive science, change in research funders' policies, research institutions recognising a need for an organised approach to data management, and the need to improve access to research outputs (Borgman, 2015 cited in Cox 2017:2182). Tansley and Tolle (2009) affirm that advances in technology now permit collection, storage, analysis, and communication of increasing amounts of scientific data.

According to Halbert (2013:2), discussions around RDM began when the National Science Foundation (NSF) of the USA released two reports that mentioned a need to build up national data management capacity. These discussions were mainly initiated due to challenges in long-term preservation and access to data sets that are publicly funded (Halbert 2013:1). The management of research data has consequently become one of the essential parts of the research process, with activities including the descriptive naming of files for easy access, keeping track of different versions of files, deleting unneeded data, regular data backups, and access control of data (Jisc, 2018). Precisely how data are managed depends on the type of data involved, how that data is generated, and how it is to be used (University of Bristol, 2012).

According to Bryant, Lavoie and Malpas (2017:5), universities benefit from RDM as it can result in the publication and reuse of important data sets that have the possibility of boosting the institution's reputation. Because of this benefit, many universities are currently interested in recording the full scope of research outputs created by faculty members and students.

Other potential benefits of RDM include improved data sharing, a decrease in laboratory or field work, improved research integrity, and more research collaborations (Lewis, 2010 cited in Chiware & Mathe, 2015:1). Data can be re-used in a manner not anticipated by original researchers, thereby encouraging new areas of research (Erway, 2013:6).

A number of university stakeholders, including academic libraries, have started to acknowledge the importance of RDM and are developing policies, services, and infrastructure to implement it (Cox et.al, 2017:2182). In doing so, many higher education institutions (HEIs) have started evaluating the needs of their researchers so that RDM policies, services, and infrastructure are appropriate to their researchers and can therefore improve research projects (Ogier et al., 2014).

According to Akers and Doty (2013), researchers on university campuses are progressively receiving support for the management and dissemination of their research data from academic libraries. Providing support for RDM has been suggested as a role for academic or research libraries in addition to the library's common role in providing research and reference services (Akers & Doty 2013).

1.1.3. Research data services

To achieve data management benefits, therefore, academic institutions ought to establish suitable systems and services to support effective RDM (Jones, Pryor & Whyte 2013:2). Academic libraries are changing their roles to conform to the changes in the research landscape (Tenopir, Hughes & Allard, 2015:17) and are intent on establishing research data services (RDS) to enhance support for their faculty members, researchers and students (Tansley, Stewart & Tolle 2009 in Tenopir, Hughes & Allard 2015). A range of RDS is becoming prominent in academic libraries, recognising the need for RDM and the changing roles of librarians (Tenopir et al. 2015:2).

According to Tenopir, Birch and Allard (2012:7), RDS are services offered by libraries to researchers regarding the management of data. Chiware and Mathe (2015:1) state that libraries in South Africa are beginning to work towards providing RDS with reasonable progress "as policies are being formulated, infrastructure set up, library staff trained, and awareness and advocacy campaigns held with academic staff and researchers".

There are a number of RDS that can be offered by libraries, and they can take many forms. Libraries can offer services from assisting researchers with data management planning or

metadata standards in their disciplines to creating institutional data repositories in which they can store their datasets in the long term (Tenopir et al., 2017:25).

1.2. Study context: University of KwaZulu-Natal

The University of KwaZulu-Natal (UKZN) was formed on 1 January 2004 because of the merger between the University of Durban-Westville and the University of Natal. The vision of the University is "to be the premier university of African scholarship" and its mission is to be a truly South African university that is academically excellent, innovative in research, entrepreneurial, and critically engaged with society (University of KwaZulu-Natal [UKZN], 2019).

The university is organised around four colleges consisting of nineteen schools. The colleges include College of Agriculture, Engineering and Science, College of Health Sciences, College of Humanities, and College of Law and Management Studies. UKZN comprises five campuses and five campus libraries, and a number of branch libraries that aim to provide an environment that is conducive to teaching, learning, and research (UKZN, 2021).

From amongst universities' core activities, the key element that separates them from other educational institutions is the generation and integration of new knowledge (UKZN, 2017). One of the foci for UKZN, therefore, is research. University Ranking by Academic Performance (URAP) has ranked UKZN third out of sixteen South African universities; and out of two thousand universities across the world, URAP placed UKZN in 396th position (University of KwaZulu-Natal Ndaba Online, 2017). The URAP 2016-2017 world ranking methodology is based on academic performance and determined by the quality and quantity of scholarly publications and international research collaboration performance (University of KwaZulu-Natal Ndaba Online, 2017).

UKZN embarked on a research flagship initiative as a major step to advance their 2017-2021 strategic plan (UKZN, 2017:2). UKZN Vice Chancellor (VC), Nana Poku, declared these research flagships as "key to the University's vision of consolidating its existing strengths; furthering its innovative, applied research; enhancing its international reach and standing, and re-configuring the way in which a 21st century university relates to key stakeholders, both in South Africa and throughout the African continent" (UKZN, 2017:2)

The four flagships are:

- social cohesion;
- African health;
- big data and informatics; and
- African cities of the future. (UKZN, 2017:3)

The flagships support and encourage collaborative research to tackle locally responsive problems and those that are internationally relevant (UKZN, 2017: 3). In addition to the VC's message, the Deputy Vice Chancellor (DVC) of Research, Professor Deresh Ramjugernath mentioned that one of UKZN's priorities is reinforcing initiatives around human capacity development, especially training and developing the next generation of researchers, while also expecting high quality research from its researchers (UKZN, 2017: 3). The DVC further mentioned that UKZN is known for its research intensiveness, both locally and globally. In 2017, UKZN invested in an integrated online research management system essential for improving research management processes at UKZN (UKZN, 2017:4).

UKZN has gone to great lengths to enhance the quality of its research, however, an aspect of data management is still missing. As research becomes more collaborative, data-intensive and computational, researchers at UKZN will be faced with a range of data management needs. These needs will influence RDS in the university. To implement effective RDS there must be an understanding of the current situation with regards to institutional practices, capabilities and preparedness to support RDM. Many researchers at UKZN are funded by the NRF – an added incentive for the institution to adopt RDM.

According to Mrs. Lindiwe Khumalo, a librarian involved in RDM at UKZN, the library has not yet formally implemented RDS. Currently, the institution is still grappling with data policy formulation. While there are many RDS with which libraries can get involved, UKZN seems to be focusing on the introduction of an institutional data repository. The institution has launched its data repository called Yabelana (https://yabelana.ukzn.ac.za/), powered by Figshare. According to Khumalo, despite the introduction of the data repository, the library has not gone to great lengths to create awareness of RDM, due to the lack of support services for data management available to researchers.

1.3. Research problem

In 2015, the National Research Foundation (NRF) of South Africa issued an Open Access (OA) statement that encourages research output from NRF-funded research to be deposited in an accredited OA institutional repository (IR). It is also expected that data supporting the publication be deposited in a certified OA repository with a digital object identifier (DOI) (NRF, 2015: online). In response to the NRF statement, as well as to institutions' own research strategies, RDS are needed at South African research institutions; however, implementing them is not a quick and easy job. Institutions need to offer services that are required by their researchers, and which they (the institutions) are capable of offering.

The Association of College & Research Libraries (Association of College & Research Libraries [ACRL], 2017) reported that out of the thirteen top trends and issues affecting academic libraries in higher education in the United States and Canada in 2017, the top three were: RDS, data policies, and data management plans (DMPs); and professional development for librarians providing RDS. Research data has remained on the ACRL 'top trends' list in the ensuing years. Investigating how libraries can respond to these trends is important. This study, therefore, seeks to investigate the role of the library in implementing RDS, using UKZN as a case study.

1.4. Research objectives

The aim of the study is to determine the role of UKZN Libraries in the implementation of Research Data Services at the University of KwaZulu-Natal.

In order to determine the role UKZN Libraries can play in the implementation of RDS at UKZN, the study seeks:

1. to identify the need for research data services among UKZN researchers;
2. to identify the major challenges associated with introducing research data services at UKZN; and
3. to determine the possibility of implementing research data services at UKZN Libraries.

1.5. Research questions

Research questions guiding the study are as follows:

1. To what extent are UKZN researchers in need of research data services?

2. What are the challenges associated with implementing research data services at UKZN?
3. In response to the identified needs and challenges, what are the research data services that could be offered by UKZN Libraries and what are the possibilities of their implementation?

1.6. Significance of the study

The findings from the study may inform the library and university management about the usefulness and importance of RDS in achieving university goals concerning research. The study motivates for further exploration of RDM at UKZN and contributes to the body of knowledge on the implementation of research data services in libraries and in the South African context in particular. The study will provide a valuable case study and encourage thoughts of research data services for other libraries that might require implementation assistance. The study might help other libraries in the same predicament to effectively manage the research data on behalf of their institutions.

1.7. Research methodology

The study utilised a mixed methods research, viz. the combination of both qualitative and quantitative research approaches, with a case study design. The researcher decided to employ convergent mixed methods and used interviews and questionnaires as methods of data collection. The research population for the study was limited to researchers who are actively generating data at UKZN and professional library staff, sampling 299 researchers for participation in the study. Data were gathered from this sample by means of a questionnaire using an online survey. Four library staff members, purposively sampled, were interviewed.

All ethical issues were addressed. The researcher obtained clearance from the University of Cape Town and the University of KwaZulu-Natal, and sought the consent of the participants before collecting data.

1.8. Delimitations of the study

The study focused on UKZN. Universities have different policy frameworks, as well as different governance and technical capabilities, which shape how research data management is supported in different universities. Because of that, findings from this study cannot be

generalised to all higher education institutions. Limitations of the study are discussed further in Chapter 5.

Participants in the study were confined to professional library staff and researchers actively generating data at UKZN. These parties are believed to be major stakeholders in the development and implementation of RDM in universities. RDM involves researchers as they are the ones collecting data and conducting studies at institution level. Having them participate in the study is essential to get some input from data originators as a way to discover awareness and introducing RDM concept. Other RDM stakeholders such as university leadership and other research units were not part of the study.

1.9. Study structure

The study is divided into five chapters. Chapter One entails the introduction, the background to the study, research problem, research objectives, research questions, the significance of the study and the research methodology.

Chapter Two encompasses the literature review. Literature review is about issues relevant to the study as presented by other scholars in the same field.

Chapter Three presents the research design and research methodology that discusses the overall design of the study, methods and instruments used for data collection and analysis.

Chapter Four presents findings and summary of data collected in tables, figures, and texts so as to assist in the understanding of data.

The last chapter discusses the findings of study. The researcher suggests some recommendations for further studies and summarises the dissertation.

1.10. Summary

The chapter introduced the study, explained the background to the study, discussed RDM and libraries, identified the research problem, and indicated the research objectives and research questions. The chapter further described the significance of the study, gave a brief description of the research methology, outlined the study limitations and reported on the structure of the study.

Chapter Two: Literature Review

2.1. Introduction

This chapter reviews the growing literature on the concept of RDM, focusing on the academic library and research support and RDS relating to Digital Curation Centre lifecycle model. These are examined in order to help identification of what RDS might be needed, specific libraries' RDS, RDM stakeholders, possible academic library RDS, assessing RDS needs of researchers, as well as RDS challenges and related studies.

2.2. The academic library and research support

As a means to comprehend the necessity for RDS and its successful implementation in libraries, it is essential to establish the role of academic libraries in academic institutions. Currently, the library is portrayed as an entrance (often virtual) through which students and faculty members access information, and no longer just a place where information is kept, as focus moves from the possession of information to its access and management (Flores et al., 2015).

According to Carey, Justh and Williams (2003), an academic library is an entity in an academic institution that provides:

- an organised collection of both printed and other reading materials;
- well-trained staff to provide and interpret educational materials as required to meet the needs of the institutions' patrons;
- a confirmed programme where staff can service the patrons; and
- the physical required facilities to support the aforementioned collection, staff, and programmes.

According to an observation by MacWhinnie (2003), in recent decades, academic libraries have gone through substantial change. There has been an increase in technology use by means of which to organise and circulate information, and computers have become a significant tool by means of which to access information. Libraries have modified their physical space so as to cater for the technology necessary for library patrons to satisfy their information needs. MacWhinnie (2003) argues that a greater need for research assistance is now expected from librarians, with patrons looking for information in an increasingly greater variety of formats,

such as print, multimedia and electronic media. MacWhinnie (2003) adds that to use the resources available successfully, patrons need the expertise of a librarian; this makes the role of a librarian important as students look for guidance in searching through different electronic information resources.

MacWhinnie (2003:241) further states that, keeping in mind the end goal of meeting novel demands, libraries are implementing new access models for electronic resources and digital information, and helping patrons in gathering data pertaining to their scholarly interests. These models include providing relevant technology for library users to use Online Public Access Catalogs (OPACs), providing access to electronic databases and providing access to information that can be accessible in different formats.

Flores et al. (2015) state that libraries continue to evolve as "multiuse information commons", with computer laboratories, writing centres, and group study rooms to enable experiential learning, independent research, and professional development. It has been noted by several scholars in recent years that academic libraries have taken more responsibility in contributing to the scholarly record. Institutional repositories (IRs) and digital libraries are some of the ways academic libraries are developing services to contribute to the research output of the university (Flores et al., 2015:88). Research data services are an additional way in which libraries can respond to the research mission of a university. As it was previously noted, research outputs have become important to improving the research reputation of the institution.

2.3. Research data and its management

Data are defined as elements of information observed, collected, or created during the course of research (Erway, 2013:5).

There are different types of research data, which are diverse, and vary even within a single discipline: "from sensory data collected in the field; secondary data created dynamically in simulations; to interview recordings or image databases" (Verbaan & Cox, 2014:211). The manner in which data are generated can cause data to be complex in nature, necessitating sufficient recorded metadata about how they were created and used, so that they can be found and accessible to other researchers (Verbaan & Cox, 2014:211).

With the rise in digitisation, the amount of data being produced and the growing requirements to store and preserve data, the awareness of the significance of preserving research data has

grown. In almost every field, the manner in which research is practiced is changing. Data can be created in large quantities, and are expensive to produce, yet invaluable if they can be reused (Verbaan & Cox 2014). Researchers use other researchers' data and conclusions to expand their own ideas, making broader research a collective effort (Institute of Medicine et al., 2009)

Preferably, data should be managed in a manner that any researcher can access and use them even after a long time (Strasser et al., 2012:1). Strasser et al. (2012) further explains that it is important for researchers to establish how to collect, document, organise, manage, and preserve data at the beginning of research so that data management becomes part of their research practice. This benefits the researcher, collaborators, and future users. Data will be easier to find, use, and analyse, and it will be easier for collaborators to understand if it is managed properly. Over a long period of time, practicing good data management means that researchers not involved with the project can find, understand, and use the data in the future (Strasser et al., 2012:1). According to Lewis (2010:145, cited in Corrall 2012:3), data management ought to extend naturally into the existing role of the library to provide access to the published part of that knowledge base.

Figure 1 depicts a pyramid of requirements for "successful data" (De Waard, 2016:50), that is, data that are saved, shared, and trusted. As can be seen on the pyramid, in order for data to be saved effectively, they must be stored and preserved. In order for data to the shared successfully, they must be accessible, discoverable and citable. Lastly, in order for data to be trusted, they must be comprehensible, reviewed, reproducible, and reusable. The figure below suggests a strategy for data management. De Waard (2016:50), comparing this hierarchy to Maslow's hierarchy of human needs, noted that "each builds on, adds value to, and in many cases requires the aspects preceding them."

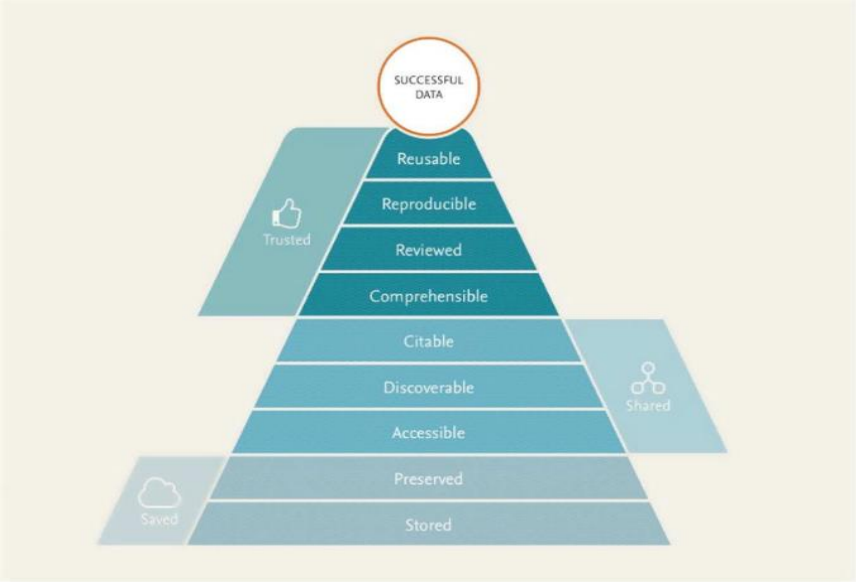

Figure 1 (De Waard, 2016)

Understanding researchers' needs and regularly involving them in institutional RDM initiatives is essential for the development and implementation of RDS in academic institutions (Wilson et al., 2011). With the aim of assisting researchers to create, collect, analyse, store, and preserve datasets, institutions are required to develop policies, infrastructures and services to manage data (Pinfield, Cox & Smith, 2014:2). "Data management can be conceived as a sequence of steps, each of which needs to be adequately completed in order for data to progress to the next step, with the intention that they can ultimately be re-used, thereby maximising their value" (Wilson, Martinez-Uribe, Fraser & Jefreys, 2011:276).

2.3.1. Data curation

According to DCC (2010), digital curation includes maintaining, preserving and adding value to digital objects throughout their lifecycle, from the data planning process where systems are developed and other data creation standards are established through the continuous storing of relevant information housed in institutional repositories (Lee & Tibbo, 2007).

Lee and Tibbo (2007) add that digital curation is the stewardship that allows authentic digital data and other digital assets to be reproduced and reused. One of the important aspects of data curation is the assigning of metadata to data sets. This aspect is essential for successful data discovery and retrieval (MacMillan, 2014:543).

RDM is about the "organisation of data from its entry to the research cycle through to the dissemination and archiving of valuable results" (Whyte & Tedds 2011:1), where it encompasses the activities that lead to data curation.

2.3.2. The research data lifecycle

According to Higgins (2008:137), the DCC Curation Lifecycle Model lays out a visual overview of the stages needed for efficient data curation and preservation from conceptualisation. The multiple activities capturing the RDM process are best illustrated with a data lifecycle model (Higgins, 2008:135). The DCC Curation Lifecycle model offers a generic structure, which identifies curation actions that can be sequentially undertaken for efficient curation and those occasionally undertaken when needed (Higgins, 2008:135). The model can be used by libraries to plan for support services (Carlson, 2014:63)

Pennock (2007:2) discusses the necessity of the lifecycle approach to guarantee the continuity of digital materials, stating that digital materials are vulnerable to change from their creation and that RDM activities undertaken at each or ignored at any stage can influence the success of their curation. A lifecycle approach establishes the necessary stages of data curation and can help to ensure that they are planned for, and that essential actions are implemented in a correct sequence. Managing data using the lifecycle approach enables continuity that also validates the provenance of digital data regardless of technological changes (Pennock 2007:2). Figure 2 depicts the DCC Curation lifecycle model that includes full lifecycle actions, sequential actions, and occasional actions.

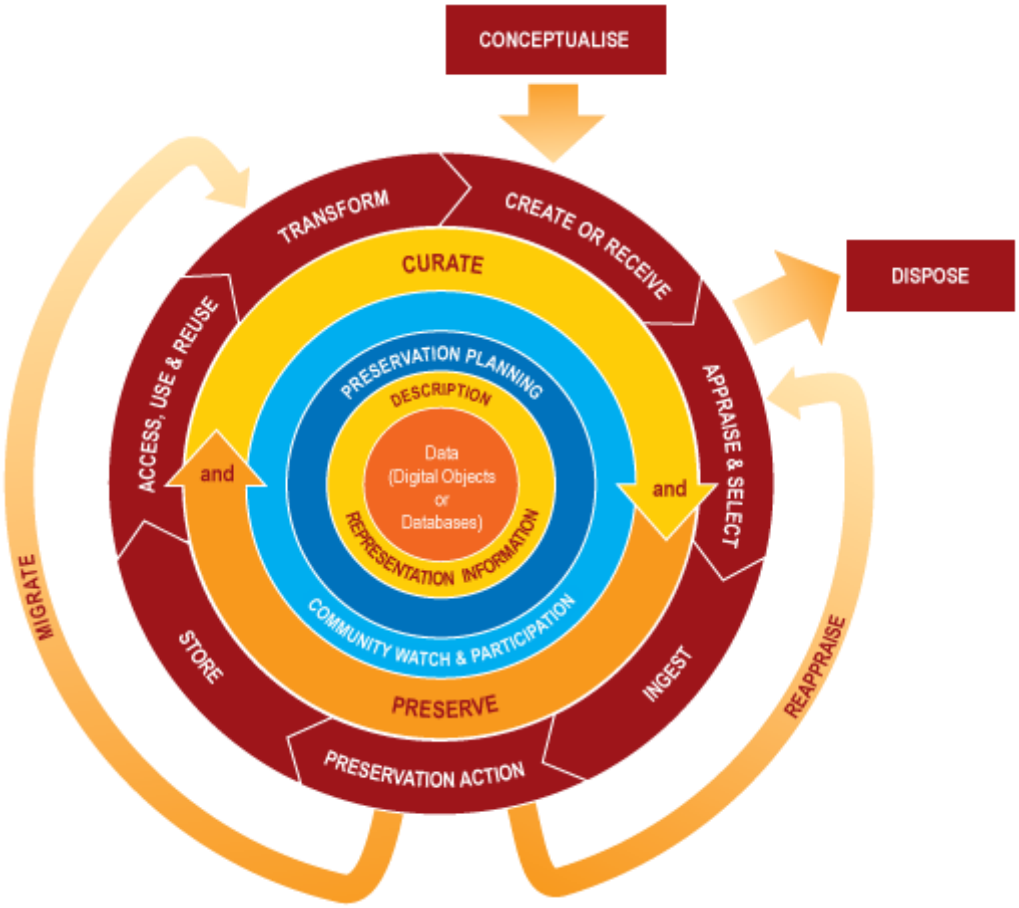

Figure 2. The DCC Curation Lifecycle Model

(Source: Higgins, 2008)

2.3.3. Full Lifecycle actions

The full lifecycle activities include a group of actions that need to be implemented throughout the data lifecycle. The DCC model comprises of four actions that are portrayed as the four central cycles that encompass the data in Figure 2.

The activities shown inside the ring of Figure 2 happen throughout the lifecycle and are sustained in all stages of the lifecycle. These activities are as shown below:

- Description and representation information: includes the administrative, descriptive, technical, structural and preservation metadata that are necessary to describe a dataset in the long term (Higgins, 2008:137).
- Preservation planning: involves the necessary administrative and management plans for all curation lifecycle actions (Higgins, 2008:137).
- Community watch and participation: involves keeping a watch on relevant community activities and taking part in the development of shared standards and tools (Higgins, 2008:137).
- Curate and preserve: involves the need for awareness and undertaking of all the management and administrative actions planned to promote curation and preservation throughout the curation lifecycle (Higgins, 2008:137).

2.3.4. Sequential actions

Sequential lifecycle actions stipulate a collection of activities that must be carried out in a particular order to enable data curation (Constantopoulos et al., 2009).

As explained by Higgins (2008), these actions include the following stages of curation, which can be seen in Figure 2:

- Conceptualise: involves the conceiving and planning of research projects and includes planning for data management;
- Create or receive: data are created or received from data creators, archives or data repositories, and activities include adding administrative, descriptive, structural and technical metadata to the data. At the time of creation, preservation metadata may also be added.
- Appraise and select: occurs once the research project is largely over. At this stage data is evaluated and selected for long-term curation and preservation. Higgins (2008) suggests compliance with recorded policies and guidelines when performing this activity;
- Ingest: the transferring of data to suitable data repositories complying with documented guidelines, policies and legal requirements;
- Preservation action: the undertaking of all required actions for data preservation, such as, data cleaning and validation, producing preservation metadata and ensuring acceptable data structures or file formats;

- Store: involves reliable and secure data storage as outlined by appropriate standards;
- Access, use and re-use: ensures the availability of data to the ones using and reusing them; and
- Transform: the creation of a new dataset from original data by migration to a new format or the creation of a new subset of data from the original.

2.3.5. Occasional activities

Occasional actions do not need to be performed as frequently as the sequential activities. As explained by Higgins (2008:138), they are:

- Dispose: data that has not been selected for long-term curation and preservation according to the recorded guidelines is disposed;
- Reappraise: data that does not meet the validation procedures is returned for further appraisal and reselection; and
- Migrate: data migration to a different format is done in accordance with the storage conditions or to guarantee that data is immune to software and hardware obsolescence (Higgins, 2008:138).

According to Pennock (2007:2), the lifecycle approach is beneficial because:

- Digital materials are fickle and likely to change from technological enhancements throughout their lifecycle;
- Activities or inactivity in every stage of the lifecycle directly affects the ability to manage and preserve digital materials in the following stages; and
- Reliable reuse of digital materials is only possible if materials are curated to retain their authenticity and integrity.

A lifecycle approach shows an illustrative overview of the actions required for effective curation and preservation of data from the beginning stages of conceptualisation of the research and receipt of the data. The model can be used to plan for research data support services and ensure that all essential stages are managed in the correct sequence. It can map activities for support, identify gaps in current provision and prioritise where investment should go. A lifecycle can assist with recognising more actions that may be required, or that are not required by certain disciplines (Higgins, 2008:137).

2.4. RDM stakeholders

A number of stakeholders need to be involved for an institution to provide RDM support throughout the data lifecycle. All stakeholders interested and taking part in RDM should care about research data and ensure that research data are managed in a manner that makes them usable in the long term (De Waard, 2016). According to Tenopir et al. (2015:2) academic libraries are playing a significant part in building a culture and infrastructure for RDS and are very important stakeholders. They added that the collaboration of libraries with other departments may provide a full suite of RDS to their respective institutions (Tenopir et al. 2015:2). Librarians are brought into collaboration to provide needed skills in data management and preservation (Tenopir et al. 2015:2).

RDM development at institutional, national, and international level relies on the joint efforts and coordinated works of the stakeholders involved (Flores et al. 2015:84). According to Flores et al. (2015), considering the role of academic libraries in activities at any of these levels requires a general consideration of the current RDM landscape. They continue to explain that setting up the different partners engaged with RDM activities and characterising their interests, roles, and responsibilities makes it possible to recognise activities where the library is well located to support and facilitate RDM development (Flores et al. 2015:84). According to Flores et al. (2015:85), university leadership and librarians are interested in ensuring researcher compliance with funder requirements, improving the creation and preservation of knowledge, and being aware of the institution's research output and building the university's reputation.

Flores et al. (2015) declared that the spectrum of RDM stakeholders has been variously categorised in the literature, however they can be assembled into four main categories namely: university leadership; government & funders; researchers; and research support units. The four categories of RDM stakeholders are illustrated in Figure 3 from Flores et al. (2015).

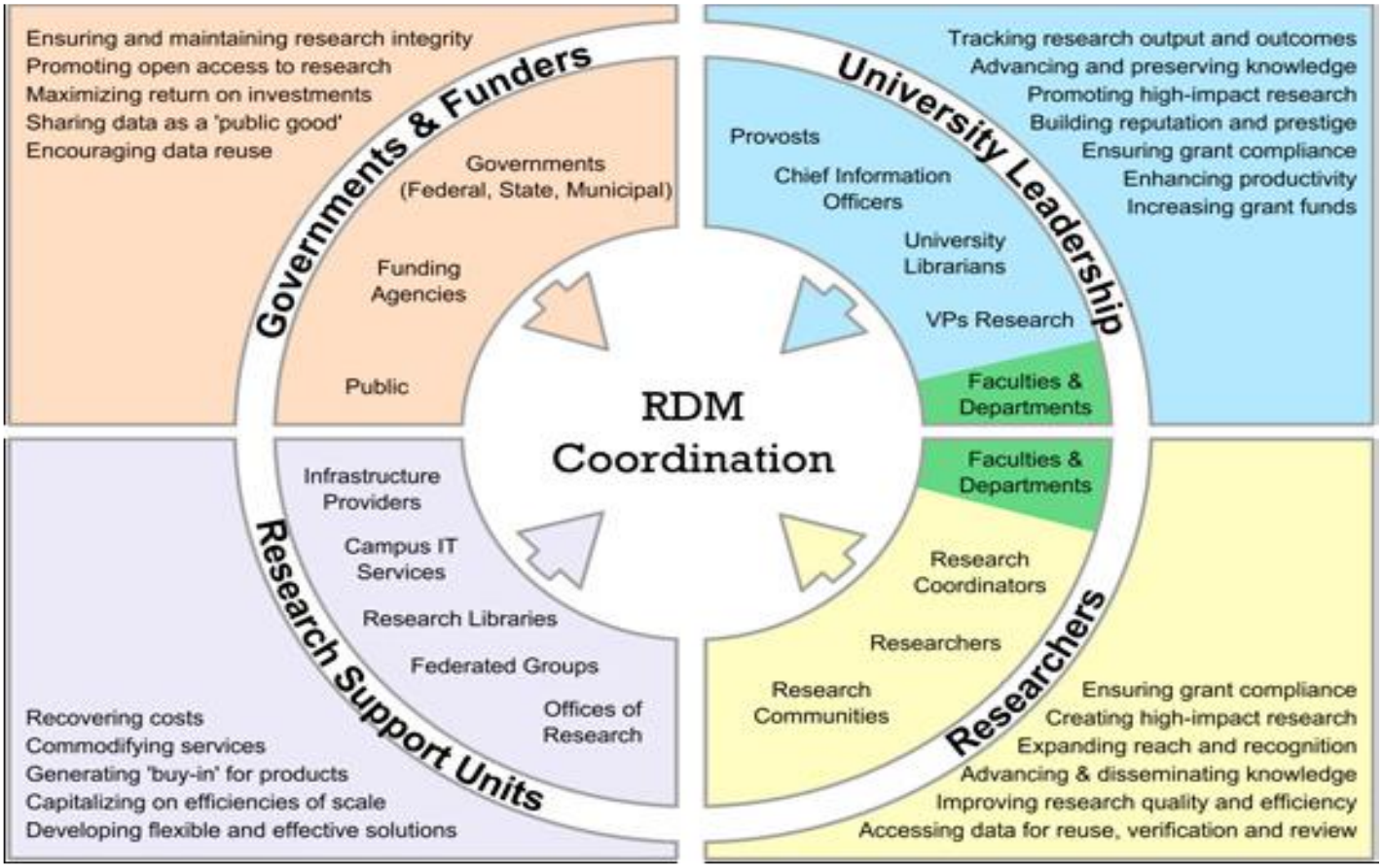

Figure 3: Four categories of RDM stakeholders

(Source: Flores et al. 2015)

Flores et al. (2015) state that the diagramme (Figure 3) is not meant to separate the groups of stakeholders, but to arrange them according to similar interests, roles, and responsibilities in RDM. In Figure 3, partners in RDM development are grouped into four general stakeholders. The focal ring distinguishes each stakeholder unit with general interests recorded in the boxes.

2.4.1. The university leadership

According to Jones, Pryor and White (2013:3), the main role of the university leadership in the introduction of RDS would be to ascertain that the RDS are beneficial, attainable and sustainable – and thereafter, provide understandable, well informed and unambiguous support for them. Consequently, the university leadership ought to provide an RDS advocate at a level of Deputy Vice Chancellor to advocate and to steer the institution's RDS efforts. This person should be a well-balanced representative suitable to represent the interests of different stakeholders:

- they have to approve proposals, plans and strategies for human and technological infrastructure as well as endorse budgets and organisational restructuring;
- they must analyse policy requirements at a national, institutional and funder level; and
- they must also establish the quality of the institution's data assets and data management practices (Jones, Pryor & White, 2013:3).

2.4.2. Government and funders

Government and funding agencies are the principal funders of academic research with an interest in maximising the return on their investments (Flores et. al, 2015:84). Funders encourage RDM activities to ensure that suitable, for example, that data ownership and sharing are included in the research process through requirements for submitting data management plans (DMPs) with grant proposals (Flores et al., 2015:85).

Funders expect researchers to have an effective system to securely store, access and share data (Jones, Pryor & Whyte, 2013:3). The statement from the NRF encouraging researchers to deposit data supporting their publications to an accredited OA repository (NRF,2015) has led institutions to develop RDS to enable their researchers to respond to the encouragement. According to Jisc (2015), the Engineering and Physical Sciences Research Council (EPSRC) Policy Framework on research data has been instrumental in the development of RDS across institutions of higher learning in the UK. Sufficient financial means are needed for RDM development. Jones (2011) states that several funders have indicated that at least ten percent of a research grant can be used for RDM development.

2.4.3. Researchers

Researchers' engagement is critical in the development of RDS as they are the creators and users of research data (Jones, Pryor & Whyte 2015:3). As the creators of research data that requires to be managed and preserved, researchers are the central stakeholders (Erway, 2013:10). Erway further states that researchers have an "intimate relationship" with their data, therefore, trust is essential if institutional RDM services are to meet their needs. Jones, Pryor and Whyte (2015:3) assert that, without the researchers' involvement and support, RDS are guaranteed to have restricted success. While other stakeholders will be responsible for their own roles, researchers are responsible for:

- ensuring that their opinions are represented by being part of the institution's RDM working groups;
- being part of the requirements' gathering and the testing of methods and solutions;
- clearly sharing their disciplines' requirements, opportunities and obstacles encountered with regards to data creation, use and management; and
- promoting the adopted services to their respective disciplines (Jones, Pryor & Whyte, 2013:3).

Researchers are responsible for producing data and for the dissemination of knowledge. They regard RDM as a means of adhering to funder requirements, maximising the effectiveness and quality of their research, and promoting knowledge in their study fields (Flores et al., 2015:85). To avoid researchers' unwillingness to take on new administrative work, their representatives ought to be involved in policy discussions so that all researchers are aware of the decisions and procedures taken in the discussions (Erway, 2013:10).

2.4.4. Research support units

It takes combined efforts from a number of research support units to address the diverse interests and requirements of RDM stakeholders (Flores et al., 2015:85). A university's research office deals with the administration of sponsored research, and related policies and services. The office supports researchers with funder requirements and ought to be in a position to offer well-informed guidance about data management services both in the institution and outside the institution (Erway, 2013:7). To keep good relationships with funders, which includes ensuring responsible data management, the research office is often the unit responsible for the governance of data management programmes (Erway, 2013:8).

As stated by Flores et al. (2015:82) academic libraries have always provided research services, and that due to the prevalence of digital data and their need to be managed, many libraries now include RDM services as part of the services offered. Given its substantial involvement with metadata, institutional repositories, preservation and curation, the library is in a great position to be the main player in data management (Flores et al. 2015:88). The library is believed to be the most suitable place on campus to guarantee safety, sustainability, and trusted stewardship of research data (Erway, 2013:10).

Erway (2013:11) states that the library does not have to take ownership of all the RDM responsibilities and functions, however, by starting a conversation university-wide, it can be part of it, and contribute its expertise.

2.4.4.1. The academic library

According to Lewis (2010), one of the drivers of RDM is the argument of public access, which is used in connection with the open access of published research papers. The argument is that society ought to benefit from access to the rewards of research that has been funded by the public. This is a viewpoint also shared by the Organisation for Economic Co-operation and Development (OECD)'s Principles and guidelines for access to research data from public funding. The guidelines state that "Sharing and open access to publicly funded research data not only helps to maximise the research potential of new digital technologies and networks, but provides greater return from the public investment in research" (Lewis, 2010). The need to stay relevant in a changing digital world, as well as demands from funding agencies and university administration, have resulted in academic libraries incorporating RDM into the library services they offer (Flores et al., 2015).

The global community has shown burgeoning interest in RDM, with the recent example being the Bill & Melinda Gates Foundation enforcing open data policy in early 2017 (Perrier et al., 2017). According to Perrier et al. (2017), Zika virus researchers are working on publishing all epidemiological and clinical data once these are collected and analysed.

Librarians in the United States are working to influence RDM national development through the Association of Research Libraries (ARL). For example, the NSF did not only fund the workshop organised by the ARL on Long-term Stewardship of Digital Data Sets (Friedlander and Adler, 2006:11), however, has also funded data curation projects, acting on the workshop's key recommendations. The projects involved libraries working together with other stakeholders and the inclusion of DMPs in grant proposals. Academic libraries are effectively represented on the Government's Research Data Strategy Working Group in Canada (Canada, 2011) and similarly in Australia, senior members of the library community have been appointed to the Steering Committee of the Australian National Data Service (ANDS, 2011). As a means of improving data quality and enabling re-use, research funders in the UK and USA require the recipients of public funding to undertake good RDM (Kahn et al., 2014).

Flores et al. (2015:88) state that librarians' expertise in research methodology and knowledge management puts them in a good position to offer relevant leadership in RDM within their institutions. Flores et al. (2015) further note that because of the history of academic libraries in supplying data for research use, librarians are familiar with the reuse requirements relating to research data. Chiware and Mathe (2015:1) observe that some South African libraries are at the beginning of providing RDM frameworks, formulating policies, setting up infrastructure, and making their communities aware of RDM services.

2.6. Academic library research data services

Research data services may include consultations on DMPs, assistance with metadata standards, support in finding data sets, and technical services, such as offering support for data repositories, dataset preparation for a repository, and assigning metadata to datasets.

Jones, Pryor and Whyte (2013:2) present some of the RDS that could be implemented by libraries at different stages of research:

- At pre-project: libraries may assist with guidance on professional use of online tools, costing data management activities and assistance with DMPs. These services relate to one of the sequential actions in the data lifecycle. The action is called "conceptualise", this is the stage where conception of the research project and planning of data management occurs.
- Through the research project: guidance on data recording, formats and standards to facilitate data reuse, advice on storage, data management and data analysis to accomplish best practices. These services primarily relate to one of the sequential actions in the lifecycle called "create and receive".
- Post project: guidance on long-term data selection, assisting researchers on how to archive data and support to make data accessible to particular users. The post-project stages are from appraise and select, up until a new project is started through use/re-use. Activities can involve the cleaning of data and the assigning of preservation metadata.

Possible library services are presented in more detail below.

2.6.1. Formulating RDM policy and strategy

According to Jones, Pryor and Whyte (2013: 5), designing a comprehensive RDM strategy is a starting point by means of which to ensure that RDM services develop in a well-organised manner. According to Pryor (2013), RDM strategy ought to be a vision statement providing details of an institutions' long-term goals, objectives and courses of action essential to accomplish them.

An RDM policy reifies the principles that an organisation has agreed will direct the decisions and actions important to accomplish what is desired by the organisation (Jones, Pryor & Whyte 2013: 5).

Jones, Pryor and Whyte (2013) warn that the policy is not to be mistaken for either a strategy or a plan, since it does not clarify what will be done to accomplish those results. RDM policy is beneficial as a public statement of intent and expression of management and senior role players (Pryor, 2013). Even though there are typical subjects that RDM policies appear to cover, there is great difference between the policies of different institutions, where institutions have adopted different ways to define their policies (van Zeeland & Ringersma, 2017).

Policy development is a process that involves the whole institution. However, academic libraries can ensure their contribution in the discussion of policy development by initiating the conversation (Erway, 2013:1). Libraries are the central and neutral place in universities and with their connections throughout the university, they are at an advantage to assist with the management and representation of stakeholder interests to a policy development committee (Flores et al., 2015).

2.6.2. Guidance, training and support

There are several ways in which libraries can offer RDM support and training to their researchers. Jones, Pryor and Whyte (2013:8) state that fundamental guidance is needed in all stages of data management. Therefore, it is helpful for academic libraries to put together websites that advise on RDM best practice and direct researchers to further institutional support that covers the "entire research lifecycle from applying for funding, through creating and managing data, to long-term preservation and reuse" (Jones, Pryor & Whyte 2013:8). The University of Pretoria (UP) seems to have heeded this call, as they have developed a subject

guide, which provides a link to the institution's RDM policy, appropriate manuals to the best practices, tools, and templates, and a free online training course (Van Wyk, 2016).

Some researchers may require extra hands-on training and customised support, where detailed technical support, such as designing a database for their research, may be needed in the course of a project (Jones, Pryor & Whyte, 2013:10).

Jones, Pryor and Whyte (2013:8) explain that some studies have recognised a need for the introduction of support staff in research training courses to raise awareness of the help they can provide researchers in terms of data management.

2.6.3. Data repositories

Data repositories are large databases assigned to manage, share, access and archive researchers' data sets (Uzwyshyn 2016:18). Data repositories may be disciplinary, or can collect data that are more general (Uzwyshyn, 2016:18). Library data repositories are usually more general, and enable long-term preservation of the range of data generated by an institution (Uzwyshyn 2016:18).

According to MacMillan (2014:545), as a requirement of publication, scientific journal publishers such as Nature, Elsevier, PlosOne, Cell and Springer ask for the data supporting publications to be made accessible to the reviewer and/or reader. This is a call for researchers to submit data related to a publication to a suitable data repository so that there is a constant connection between articles and related data sets (MacMillan, 2014). Pampel et al. (2013) declare that the open access to publicly funded research is guaranteed to fail without trustworthy and sustainable data repositories that support researchers to share their research data.

Data repositories do not accept every kind of data, because some kinds require specialised software, or because datasets may be too large, and some repositories may not be in a position to maintain this data for the long-term (Van den Eynden et al., 2011). This point is corroborated by Cox and Pinfield (2014) where they point out the complexity and scale of RDM issues in universities. They state that good data management and the way in which data are generated will widely differ within an individual institution because of the fluidity and fragmentation of academic disciplines. In a survey conducted by Tenopir et al. (2015), interviewees raise

concerns about data storage and that the library would not have resources to sustain the demands for space in its repository.

2.6.4 Data management plans

A data management plan (DMP) describes the type of data that will be generated; the policies that apply to the data (e.g. funding and institutional policies); the owner and who will have access to the data, data management practices to be utilised (e.g. backups, access control and archiving); required equipment and facilities (e.g. back-up server and repository); and the assigning of roles and responsibilities for each aspect of the plan (Australian National University, 2012, cited in Si et al., 2015).

Funders insist on DMPs and researchers must comply. By not having DMPs in place, researchers are risking their chances of obtaining funding (Uzwyshyn, 2016:2). Uzwyshyn, (2016:2) notes that a great number of funded researchers currently do not share their data, however, with funders insisting on DMPs, this situation is quickly changing. DMPs provide a chance for researchers to indicate that they are aware of good practice and guarantee that their proposals are aligned with funders' data policies (Jones, 2011).

Jones, Pryor and Whyte, (2013:11) explain that, when creating good DMPs, it is crucial that researchers are provided with guidance, tools, and support services. Libraries can support the creation of DMPs by:

- creating templates and guidelines on what to include in the DMP template;
- providing customised DMP online; and
- training and advisory support such as consultancy services.

According to Jones, Pryor and Whyte (2013:11), libraries should offer guidance with regards to what should be encompassed by DMPs. They further pointed out that it is important for institutions to provide guidance on suitable methods and determine a few recommended approaches for storage and backup. Jones, Pryor and Whyte (2013:11) further state that more training on data management planning and is needed and this is something suitable for the library to do.

2.6.5. RDM Consultancy services

Jones, Pryor and Whyte (2013:8) state that one-on-one consultancy sessions enable researchers to ask precise questions so they can embrace appropriate ways to create, manage, and share their data. It is important to think on a range of alternatives to engage different groups of researchers as they might not ask similar questions. For example, one may reach out to emerging and prolific researchers separately.

2.7. Challenges associated with RDS provision

Research institutions are confronted by many challenges in the process of RDM. These challenges include the long-term preservation of large amounts of data, constant assigning of metadata, keeping abreast of data standards, and sharing data successfully while observing possible restrictions (Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, 2010:366)

RDM is associated with challenges such as:

Cox, Verbaan and Sen (2012) explained that librarians are in an ideal position to offer RDS, however, that this does not come without certain challenges. They mentioned that existing roles for librarians are demanding, and that to offer RDM services, these services would need to be de-prioritised. According to a survey conducted by Cox and Pinfield (2014:4), leadership in RDM has been provided by libraries, especially in advocacy and policy development, however service development focusing on advisory and consultancy services is still restricted. The survey revealed that other crucial challenges included gaining RDM support from researchers, senior managers working with other support services, and resourcing RDS. Cox & Pinfield (2014) further asserted that librarians often lack the individual research experiences needed to be key players in the field of RDM and so may lack the extensive understanding of the needs and practices of researchers. However, Carlson and Garritano (2010) cited in Cox & Pinfield (2014) highlight the need for risk taking, open-mindedness, and creativity in assuming such new roles.

Establishing and supporting RDM infrastructure is a new arena for librarians. Librarians engage more in consulting and training-oriented activities than technical RDS (Tenopir et al., 2014); (Cox & Pinfield, 2014). This highlights some of the challenges libraries might face in implementing RDS. Chiware and Mathe (2015:8) mentioned the restricted skillsets of

librarians and, in some cases, lack of understanding of different research data and their practices in different disciplines. As a solution, they suggest assigning existing staff members to new RDS roles with relevant training. There are other challenges listed by Cox, Verbaan and Sen (2012). These include the lack of domain-specific knowledge by librarians, the complicated task of translating the existing library practices, such as metadata creation, to research data contexts. Libraries' focus in recent years has been on supporting teaching, this makes it difficult for them to engage researchers in RDS (Cox, Verbaan & Sen, 2012).

Patterton, Bothma and Van Deventer (2018:15) observed the lack of RDM implementation guidelines for libraries with insufficient resources. Misinformed management not prepared to invest in RDM infrastructure or attending to matters that are more urgent could cause this insufficiency (Patterton, Bothma & van Deventer 2018:15). Other challenges were reported by Raboin, Reznik-Zellen & Salo (2013) when reporting on the experiences of three institutions in the USA (Tufts University, University of Massachusetts, University of Wiscosin). The authors concluded that in developing organisational capacity for RDS, gaining support from management, capability of staff to manage RDS, and the ability to meet researchers' evolving needs were the common challenges.

A study by Perrier et al. (2018:178) identified a lack of resources and funding as a challenge that hinders the availability of infrastructure such as data storage and data quality. Convincing researchers to be part of the RDM initiative, skills gap, and lack of resources were some of the challenges mentioned by Perrier et al.

2.8. Assessing RDS needs of researchers

Researchers are the main stakeholders in RDM and involving them in the development of RDS is of utmost importance (Jones, Pryor & Whyte 2015:3; Erway, 2013:10). Before implementing RDS, it is essential for the library to perform needs assessments to establish the relevant audience for the services (Read, 2007:72). As stated by Flores et al. (2015:94), needs assessments undertaken through surveys and interviews with university researchers assist librarians in the discovery of how different campus stakeholders manage research data, and how RDS might assist in RDM.

Due to the increasing significance of data management in academic institutions, it has been noted that some higher education institutions have opted for their own RDM strategies. The

evidence of this has been the undertaking of requirements-gathering exercises with the assistance of tools such as Data Asset Framework (DAF), Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO), and Assessing Institutional Digital Assets (AIDA) (Rans & Jones, 2013:1). These tools support institutions conduct surveys on researchers' RDM practices and available infrastructure to assess RDM needs. The DAF is one of the several tools that provide institutions with the means to identify, locate, describe, and assess the management of their research data (Data Asset, 2017). Many research institutions have developed policies based on the findings of this tool. According to Ekmekcioglu and Rice (2009), in May 2011, the University of Edinburgh released a first RDM policy based on DAF findings.

CARDIO is a benchmarking tool that is applied at either departmental or research group level to develop a data management strategy (DCC, 2017). This tool creates an agreement between data creators, information managers and services providers, identifies practical RDM goals for improvement, identifies inefficiencies in operation, and most importantly, it supports the collaborative assessment of RDM needs (CARDIO, 2017). Every institution that engages in RDM must have the requisite skilled staff, infrastructure, resources and upper management support to ensure effective data management for purposes of evidence, reuse, and validation.

Case studies of research groups' data practices are significant in understanding their RDM needs. This involves observing the work produced by researchers, interviews, workshops, or focus groups (Whyte & Allard, 2014). As suggested by Flores (2015:94), this study made use of an online questionnaire to assess researchers' needs. Interviews and focus groups could also have been a suitable way to collect relevant data from researchers before setting up RDS.

Regarding any new and advanced library services, it is essential to promote them and raise awareness among possible users. In developing the service it is of utmost importance to identify current and potential users, understand their needs, and ensuring that they are aware that the service is accessible to them (Read, 2007:61).

2.9. Related studies

This section looks into other studies that have investigated the implementation of RDS in academic libraries in different countries. According to Henderson & Knott (2015), literature does not recommend a single implementation model for RDS.

Henderson and Knott (2015:49), giving an account of RDM implementation experience at the Virginia Commonwealth University Library (VCUL), stated that "by learning from the successes and mistakes of data management pioneers, some missteps could be avoided, some steps could be condensed or eliminated, and the time to set up a working service could be reduced". The VCUL implementation plan started with easy support services such as helping with the DMPTool through "consultations, group presentations, and teaching opportunities" with those seeking assistance with DMPs. Consultations with researchers were encouraged by the promotion of a DMPTool in different subject areas (Henderson & Knott 2015:49). They also focused on strategic partnerships and creating new relationships with researchers. The plan going forward indicated that services provided include:

> Assistance with data collection, data processing, analysis of data, saving data for the long-term, data curation, sharing data, finding data to reuse, depositing data in repositories, using GIS services, presenting data, designing research, collecting metrics to show the impact of shared data, writing DMPs for grants, and teaching best practices for data collection and use (Henderson & Knott, 2015:49).

A study that was conducted at the Johns Hopkins University (JHU) revealed that the services required included storage, archiving, preservation and data curation (Shen & Varvel, 2013).

The JHU focussed on environmental responsiveness (JHU internal needs), socio-technical readiness (human and technical sustainability), and marketing and collaboration strategies (Shen and Varvel, 2013). The JHU data management services showed "environmental responsiveness" from the beginning by developing a service specifically to meet the needs of the faculty members. To create JHU data management services, several exercises were performed including a meeting with researchers to understand needed services. Shen & Varvel's (2013) study revealed that JHU was successful in the development of DMP consultations, adoption of services and acceptance of the JHU data archive. A challenge facing services at JHU included the difficulty to convince researchers that their proposals can be improved by the use of DMPs.

Delasalle (2013) notes in reporting on the case of the University of Warwick that the university based its approach towards RDM on "setting requirements and providing support". The initial step in setting requirements was the approval of the University of Warwick Research Data Management Policy by the University's Steering Committee (Delasalle, 2013). Delasalle

(2013) asserted that it was important for the policy to define research data, its compatibility with funders' requirements and its suitability with the needs of the university to set direction for the future. From the interviews and data management events that were carried out at the University of Warwick, a number of challenges were revealed, including:

- researchers' uncertainty about the value of data sharing;
- the cost of keeping data compared to reproducing them when needed;
- feeling intimidated by new requirements that threaten to hinder the progress of research;
- metadata not being a priority to some physicists;
- ways to make personal data anonymous; and
- non-public access to data for researchers only.

2.10. Summary

This chapter presented a review of academic literature relevant to the subject of RDS. A few stages for the implementation of RDS were distinguished. These include RDS needs assessment and awareness, challenges associated with RDS provision and RDM stakeholders as different stakeholders are significant for RDS implementation. They support and complement each other in the implementation of the services. This will ultimately ensure data are easier to find, more visible and generate more impact and improve data sharing between researchers and other stakeholders alike. All research institutions ought to incorporate RDS in their research activities.

Chapter Three: Research methodology

3.1. Introduction

This chapter deals with the methodology used for this study. It describes the processes of collecting, analysing and interpreting data that were utilised to answer the research questions that guide this study. This chapter provides a detailed description of the methods and processes used in the study as well as the reasons for using those procedures.

3.2. Research approach

According to Creswell (2014:3), there are two approaches that are commonly used in research, namely qualitative and quantitative. When these research approaches are combined, it is known as a mixed methods approach.

3.2.1. Quantitative research

Quantitative research examines social phenomena by measuring different variables regarding numerical analysis. In quantitative research, data is displayed through tables, numbers and rates (Babbie, 2009:35). Hopkins (2000:1) stated that researchers measure variables on a sample of subjects, which can be tissues, cells, animals, or humans. Researchers reveal the relationship between variables using effect statistics, such as correlations, relative frequencies, or differences between means (Hopkins (2000:1).

3.2.2. Qualitative research

According to Leedy and Ormrod (2013:140), qualitative research aims at understanding the meaning human beings allocate to social problems. Qualitative research usually investigates issues conducted in a natural environment so as to fully understand events and be able to answer questions using detailed studies of small groups of people. Aurini, Heath and Howells (2016:4) add that qualitative researchers are interested in understanding how people make sense of the world around them as well as describing the context and meaning of individual and group life.

3.2.3. Mixed methods research

This research study takes a mixed methods approach. The research approach to be utilised in a study is determined by the nature of the research problem. According to Johnson et al. (2007:123, cited in Harrison and Reilly, 2011:8), mixed methods research sees a researcher combining elements of qualitative and quantitative research approaches for the purposes of breadth as well as depth of understanding.

According to Plano, Clark and Creswell (2011:8), qualitative and quantitative data provide different perspectives in research and each have limitations. The qualitative understanding emerges after investigating a few peoples' perspectives. Quantitative understanding emerges after evaluating the responses of a large number of people according to a few variables (Plano Clark & Creswell, 2011:8). Plano Clark and Creswell (2011) add that when researchers qualitatively study a few individuals, they lose the ability to generalise the results, such that when they quantitatively investigate a large number of people, the understanding of individual perspectives is lost.

Ngulube (2012:118) explains that researchers realised qualitative and quantitative methods are not a complete opposite of one another, but that instead, they complement and support each other with the potential of acquiring the complete picture of a given social phenomenon. There is a belief that using both methods in mixed methods research can lead to a better understanding of research problems (Creswell & Garrett 2008, cited in Ngulube, 2012: 118).

In this case, the researcher chose to employ a convergent mixed methods approach. This approach merges qualitative and quantitative data in order to provide a comprehensive examination of the research problem. In convergent mixed methods research, the researcher collects both qualitative and quantitative data almost at the same time and then combines the information in the interpretation of results.

3.3. Research design

Research designs provide guidance as to how to conduct research (Sarantakos, 2005:106). They describe the resources needed to complete the research, they are important in directing research, and they provide a plan on how to acquire information from research participants and on how participants will be selected (Welman, Kruger & Mitchell, 2005:52). This study made use of a case study design to investigate the role UKZN Libraries can play in the implementation of

research data services. Leedy and Ormrod (2013: 141) define a case study as "a particular individual, program [sic], or event studied in-depth for a defined period of time." They further mentioned that a case study may be suitable for investigating how a situation changes over time because of certain interventions as well as learning more about a lesser-known situation. According to Sarantakos (2005:212), researchers emphasise that research findings generated through a case study cannot be generalised because they are restricted to a particular community at a particular time. Case studies can be associated with descriptive, exploratory or explanatory characteristics and they support in-depth analysis because even though the sample size might be small, the case is investigated in detail (Leedy & Ormrod, 2013:141).

A descriptive case study, also known as intrinsic case study, aims to describe, interpret and analyse a specific event (Yin, 2003 quoted in De Vos et al., 2011). The aim of the descriptive case study is not to understand a wide social issue, but only to describe the case at hand (Mark, 1996 quoted in De Vos et al., 2011).

An explanatory case study, also referred to as an instrumental case study, strives to test and build theory focusing on advancing the researcher's understanding about a particular phenomenon (De Vos et al., 2011).

An exploratory case study strives to investigate something that is not well researched (Hesse-Biber & Leavy, 2011:10). This particular study focuses on RDM, an emerging area of scientific research and a new concept in universities, and therefore employs an exploratory case study design. This study investigates a single case of the University of KwaZulu-Natal, collecting data from different sources. These sources have been identified as important role players in the implementation and development of RDM at institutional level.

3.4. Data collection methods

According to Iarossi (2006), the validity of any scientific research depends on the quality of instruments used to collect data for the study. Appropriate data collection is the foundation upon which studies are built, and it must be done in a manner that ensures that results generated from the data are reliable, universal and repeatable, in addressing specific research questions (Iarossi, 2006).

Primary data was collected for this study, because data was collected to address the research problem at hand. Primary data refers to original data collected by the researcher from the field by means of engaging with the study population (Gelo, Braakmann & Benetka, 2008:275).

Different data collection methods can be used in mixed research methods. These include observations, textual and visual analysis, interview, focus groups, and questionnaires (Gill et al. 2008:291). This study used interviews and questionnaires as methods of data collection.

3.4.1. Interviews

According to Gill et al. (2008:291), there are three basic types of research interviews, namely, structured, semi-structured, and unstructured interviews. Structured interviews consist of a list of fixed questions. They do not give room to follow up questions for answers that may require further explanation. They are relatively quick and easy to conduct, however, they offer a limited number of participants' responses and therefore are not suitable when depth is needed (Gill et al 2008:291).

According to Punch and Oancea (2014:182), unstructured interviews are comprehensive explorations of experiences and interpretations of participants in their own terms. Unstructured interviewing are known for in-depth interviews, being non-standardised, and having open-ended questions, and are sometimes called ethnographic interviews (Punch & Oancea, 2014:182).

Semi-structured interviews rely on a particular list of questions and they strive for the discussion to remain about those questions, however, they give participants some liberty to discuss what is interesting and important to them (Hesse-Biber & Leavy, 2011:102). A researcher utilising semi-structured interviews is likely to permit the discussion to develop, discovering new topics that are relevant for the participant. Semi-structured interviews are standardised, but have an open-ended interview schedule. The same set of questions are asked from participants, however, they are allowed to approach the question and respond to it in any way they want (Aurini, Heath & Howells, 2016: 82). In this study, semi-structured interviews constituted one data collection method used, for which a list of questions was developed to ensure qualitative efficacy.

Semi-structured interviews were suitable for this study because:

- the researcher could provide clarity to questions that could be misunderstood or misinterpreted;
- both closed and open-ended questions could be asked; and
- in-depth collection of data was possible, since the researcher can ask for more information where needed (Kumar, 2011:150, Kombo & Tromp, 2006:93-94).

3.4.2. Questionnaires

As stated by Bowling (2002) in Jones et al. (2008:16), questionnaires are known for having particular advantages compared to other data collection methods such as interviews. The low cost of data collection and processing, and little training needed to administer them are recorded as the main advantages of questionnaires (Jones et al. 2008:16). More advantages include the larger population that can be reached by questionnaires compared to interviews, and the fact that they can be delivered in different ways, whether on paper or electronically (Jones et al. 2008:16).

For this study, questionnaires were the second data collection method utilised, and were administered online using Google Forms. An online questionnaire was deemed relevant for this study because of the user-friendliness of online questionnaire tools. Another advantage of online questionnaires mentioned by Jones et al. (2008:17) is that they can be designed to be "visually attractive and data entry can be controlled at every point by real-time error checking and correction" to lead the respondent through the process.

3.5. Research population

Pickard (2013:60) describes a research population as a complete set of individuals about which inferences will be made. The research populations for this study were researchers (1341), who are actively generating data at UKZN (DHET, 2018), and professional library staff.

3.6. Sampling and sampling technique

According to Flick (2011:77), sampling refers to techniques for ascertaining that researchers have the "right" cases in their study. In quantitative studies, these cases should enable generalisation from the sample to the population, because the sample represents the population. Kombo & Tromp (2006:77) define sampling as a process of selecting a specific number of

individuals or objects from the study population whereby the selected individual or objects have similar characteristics to that of the entire population. There are two main methods of sampling: namely, probability and non-probability sampling. The difference between the two sampling methods is that in a probability sample, the chances of members of the larger population being selected for the sample are known, while in a non-probability sample, the chances of the larger population being selected for the sample are unknown (Cohen, Manion & Morrison 2000:99).

The researcher utilised purposive sampling for the qualitative part of this study. A purposive sampling technique was utilised to identify library staff who are involved in RDM at UKZN. The aim of purposive sampling is to sample participants strategically to ensure that those sampled are suitable to the research questions (Bryman and Bell 2011:187). Cohen, Manion & Morrison (2000:103) state that in purposive sampling, researchers are at liberty to choose cases to incorporate in the sample based on their discretion. This gives them freedom to gather a sample that is sufficient to their particular needs. In this study, I purposively sampled four library staff with knowledge of RDM.

Random sampling was used to sample researchers for participation in the study. "Random sampling is a method of drawing a sample of a population so that each member of the population has an equal chance of being selected for participation" (De Vos et al., 2011:226). (5% margin of error and 95% confidence interval) was determined using the Raosoft calculation (Raosoft Inc., 2004).

3.7. Ethical considerations

Prior to data collection, the study observed specific ethical considerations. This is of utmost importance, because data were collected from human subjects. Ethical considerations were developed to protect research participants, researchers, and the organisations they represent (Devlin 2018:103). The researcher asked for and obtained clearance from the Ethics Review Committee of the Department of Knowledge and Information Stewardship in order to comply with the research ethics and to authorise the data collection process before the approval of data collection instruments (see Appendix A). A researcher requires a research permit prior to data collection (Odiya, 2009:98).

After then applying for and receiving clearance from UKZN to collect data from research and library staff, The researcher obtained consent from the participants in the form of an informed consent form (See Appendix B). Interview participants (library staff) signed this form so as to ascertain that the researcher is permitted to collect the relevant data. For online participants, it was indicated that by continuing with the questionnaire, they confirmed their consent. According to Kombo and Tromp (2006:82) and Flick (2009:37), consent forms ensure that interviewees are willing participants in the research activity. During the presentation of research findings, identity of participants was not be revealed, preserving the anonymity and confidentiality of the participants. Furthermore, a DMP which provides clarity on the issues concerning data access is included (see Appendix E).

3.8. Data collection

Data collection began after receiving ethical clearance from UKZN. Emails with consent forms were sent to the four library staff asking them to participate in the study. Library staff were interviewed (See Appendix D) in December 2019.

On 29 January 2020, an email with a link to the Google Forms questionnaire was sent to UKZN to be placed in a notice system. The UKZN notice system is a web-based system for reading existing notices and placing new notices. After notices have been moderated, the system generates emails and send to targeted audiences in relevant campuses (UKZN, 2017). The questionnaire did not get adequate responses the first time. More responses were solicited by re-sending the questionnaire on 31 March and it ran until 30 April. The survey ran for a total number of three months, from 29 January to 30 April 2020.

3.9. Data analysis and interpretation

According to Nueman (2012:352), data analysis involves sorting out data in an appropriate manner to simplify understanding of the research findings. It is at the analysis stage that a researcher distinguishes patterns and formulates generalisation. Data analysis assists researchers make judgments based on evidence and reach conclusions so that research questions can be answered (Nueman, 2012:352). Wallace and Van Fleet (2012:265) state that data analysis serves two purposes, namely, preparing data for use, and making sense of the data. The researcher in this study firstly analysed research data obtained from interviews. Data

from interviews was transcribed into Microsoft Word. Themes were assigned to qualitative data for example, answers to open-ended questions. Data from the questionnaire was entered into Microsoft Excel because tabular representation was required at a particular point. Qualitative and quantitative data were considered together at the interpretation stage of the study.

3.10. Reliability and Validity

Validity refers to the degree to which a concept is accurately measured in a study (Heale, 2015:66). Different data collection methods that were utilised in this study are means of ensuring some validity as data can be compared.

Reliability refers to the degree of consistency of the measure (Heale, 2015:66). The pilot study to test validity was not done due to time constraints. However, questions were constructed in a manner that would not bring confusion to the respondents and the same questions were asked of all participants. According to Creswell and Creswell (2018:199), reliability and validity play an important role in improving the accuracy of the assessment and evaluation of research work.

3.11. Research bias

Smith and Noble (2014:100) refer to research bias as an expression of the subjectivity of the researcher. Bias occurs in all research, across research designs, and is often difficult to eliminate. They mention that it can happen in all research stages and it affects validity and reliability of study findings. The researcher ought to endeavour to restrict bias and ensure that the study findings are based on the collected data (Pickard, 2008:21). This approach was taken in the current study. Although raw data (audio recordings and transcribed interviews) from this study will not be attached to the final work, they will be available for verification purposes as relevant. This will be done in a manner that does not compromise the confidentiality and anonymity of participants.

3.12. Summary

This is a case study of the role of UKZN Libraries in the implementation of RDS. Four members of the university library staff and 299 researchers were selected. Information was obtained from them using interviews and questionnaires. This generated qualitative and quantitative data that was analysed using Microsoft Excel software and expressed in words,

respectively. Ethical approval was obtained from the University of Cape Town Research Ethics Committee and the gatekeeper permission was obtained from the University of KwaZulu-Natal. Consent was obtained from all participants. This chapter discussed the methodology, the design, data collection methods and instruments, ethical issues and research subjectivity of the study. Chapter Four presents findings gathered from collected data.

Chapter Four: Data analysis and interpretation

4.1. Introduction

This chapter presents the analysis of the qualitative data obtained from interviews with librarians and quantitative data collected from questionnaires with researchers. The chapter presents the data to retrieve answers to the research questions that guide this study.

4.2. Response rates

Of the 299 researchers invited to respond to the questionnaire, only 33 participated, representing a response rate of 11 percent. According to Massey & Tourangeau (2013:226), there are three categories for non-response:

a) non-contact (not able to communicate with potential respondents);

b) refusal (contact was made but the respondent refused to participate); and

c) 'other' reasons.

From the non-response categories, this study is presumed to have experienced a combination of the above. The researcher has identified the following reasons for the low response rate, which could have resulted in non-contact, refusal, or other reasons for researchers not participating:

- UKZN embarked on a very long student protest in February 2020 and the university had to be non-operational for about two months; and
- shortly after classes resumed, South Africa began a nationwide coronavirus lockdown.

This meant that the targeted respondents were not able to fully focus on the questionnaire as they were not at the university, some probably without email access, and were then faced with the pandemic and subsequent lockdown.

4.3. Quantitative research analysis and findings

4.3.1. Questionnaire responses: Overview

Tables and graphs were used for the presentation of data. Because of the low response rate, the percentages are not shown where numbers are very small. The researcher has also not shown percentages for single responses, but has included them where useful. The researcher has used decimal points where needed so that percentages add up to 100 percent.

4.3.2. Preparation and data checking

Data collected was assessed for completeness and consistency. All thirty three respondents completed all the questions. If the respondents left any questions unanswered, it was because they were follow-up questions that they were not directed to them (for example, Question 7).

4.3.3. Analysis of collected data

The questionnaire was divided into two sections. The purpose of the first section was to collect demographic information about which UKZN colleges and schools the respondents were from and their highest level of education. The second section was developed to gather information about the knowledge and behaviours of UKZN researchers concerning RDM to inform their RDS needs and identify potential challenges of implementing RDS. The data analysis that follows maintains the structure of the questionnaire (See Appendix C), addressed under the following headings.

- demographic aspects (4.4.1);
- the types of data researchers are collecting (4.4.2);
- RDM - how researchers manage their data (4.4.3);
- different RDS that could be offered by UKZN Libraries (4.4.4); and
- general views about the implementation of RDS at UKZN (4.4.5).

4.3.3.1. Demographic aspects

The researcher collected demographic information to find out more about the respondents as well as to see if responses could be generalised across colleges or schools. Table 1 shows how many respondents came from each college and from schools within the colleges.

It shows that the majority (14; 42, 4%) of respondents came from the College of Humanities.

The College of Agriculture, Engineering and Science and the College of Law and Management Studies had the same number of respondents (7; 21, 2%). The College of Health Sciences had the least number of respondents (5; 15, 2%).

Table 1: ***Colleges and corresponding schools to which respondents belong (N = 33)***

College	**School**	**Number**
Agriculture, Engineering and Science (n=7)	Chemistry and Physics	1
	Agricultural, Earth and Environmental Sciences	4
	Engineering	2
Health Sciences (n=5)	Health Sciences	2
	Laboratory Medicine & Medical Sciences	1
	Nursing & Public Health	2
Humanities (n=14)	Applied Human Sciences	3
	Arts	1
	Education	2
	Social sciences	7
Law and Management Studies (n=7)	Business and Leadership	1
	Law	2
	Management, IT and Governance	4

The respondents to indicate their highest level of education. Figure 4 shows that 11 respondents (34 %) hold a Masters' degree, followed by those with honours degree (8; 24%), doctorate (7; 21%), then bachelor's degree (4; 12%). Two respondents indicated that they were professors (6%), while one respondent selected other and specified that they hold a National Senior Certificate with a Bachelors' pass. The invitation was sent to researchers only; there is no clear understanding as to how someone without a degree was included in the sample. It is assumed that those with merely a bachelor's degree fell within the research sample, because they are currently doing their Masters' degrees; therefore, they considered bachelor's degree as the relevant level to select.

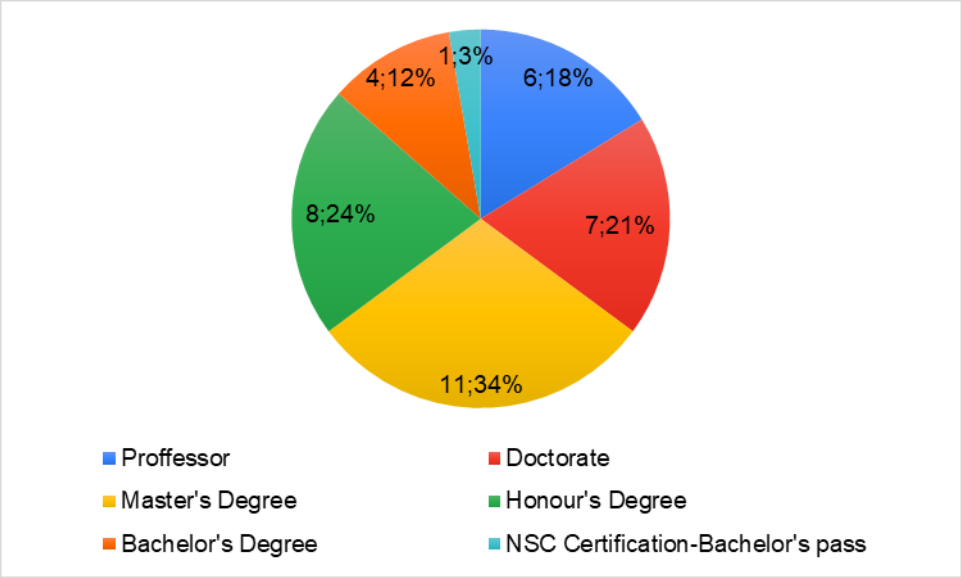

Figure 4. ***Highest level of education N=33***

4.3.3.2. The types of data researchers are collecting

The researcher aimed to establish the type of data that the participants are collecting to help to identify which RDS may be implemented in terms of storage.

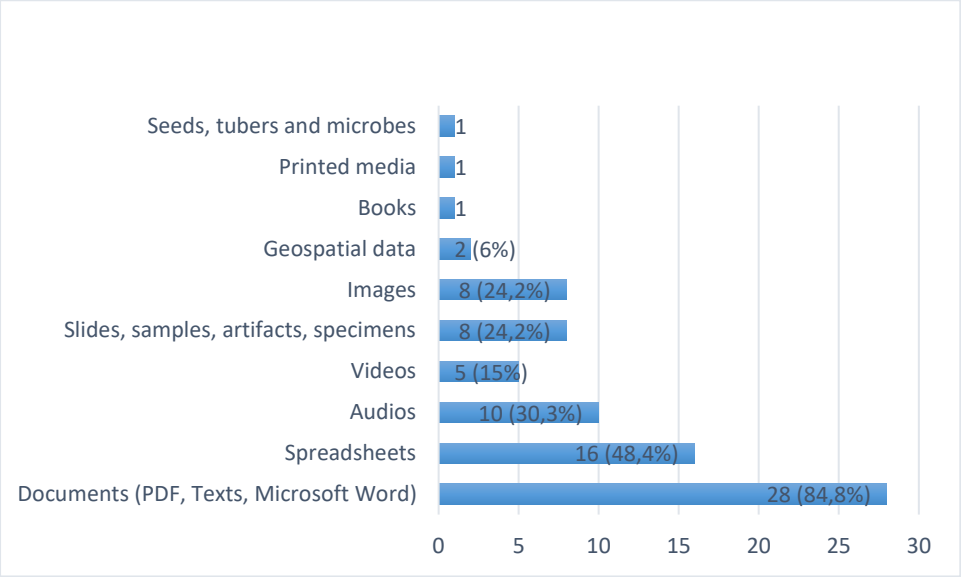

Figure 5. ***Types of research data created/used by researchers N=33***

The researcher provided the respondents with a list of data types from which to select. The list included types of data as well as file types on which data were saved and respondents could select as many as applied. Figure 5 shows the number of researchers who used different data types. Most researchers (28; 84.8%) work with text documents such as PDFs and Microsoft Word files, followed by those who work with spreadsheets (16; 48.4%); ten work with audio files; while eight researchers use images and slides, respectively. Five respondents use videos, while two work with geospatial data. Two researchers selected other; one indicated that they work with printed media and the other with living materials like seeds, tubers, and microbes.

In the next question (Question 5), in order to establish whether researchers were aware of data sharing and its barriers, such as file format obsolescence, they were asked as to whether the formats and software they use enable data sharing and long-term access to data. Figure 6 shows that the majority (26; 78%) indicated that their formats and software do indeed enable data sharing and long-term access. Three researchers chose 'no'. Four researchers were not sure, and one researcher indicated that they do not know whether their formats and software enable data sharing and long-term access.

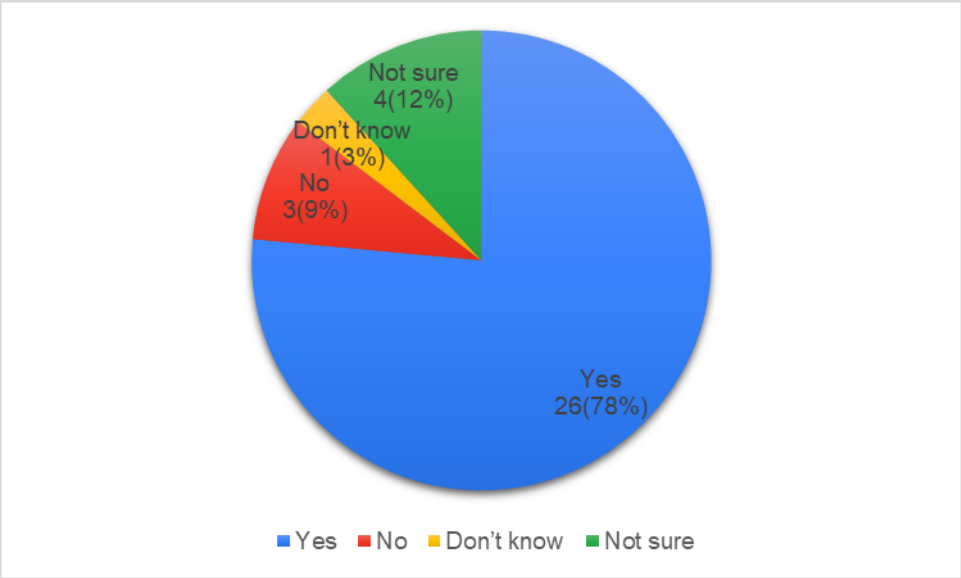

Figure 6. ***Do your chosen formats and software enable sharing and long-term access to the data? N=33***

Question 6 was asked to find out if the participants have developed a DMP for their research and if they are aware of what a DMP is. Participants were given three options to choose from, viz. yes, no, and don't know.

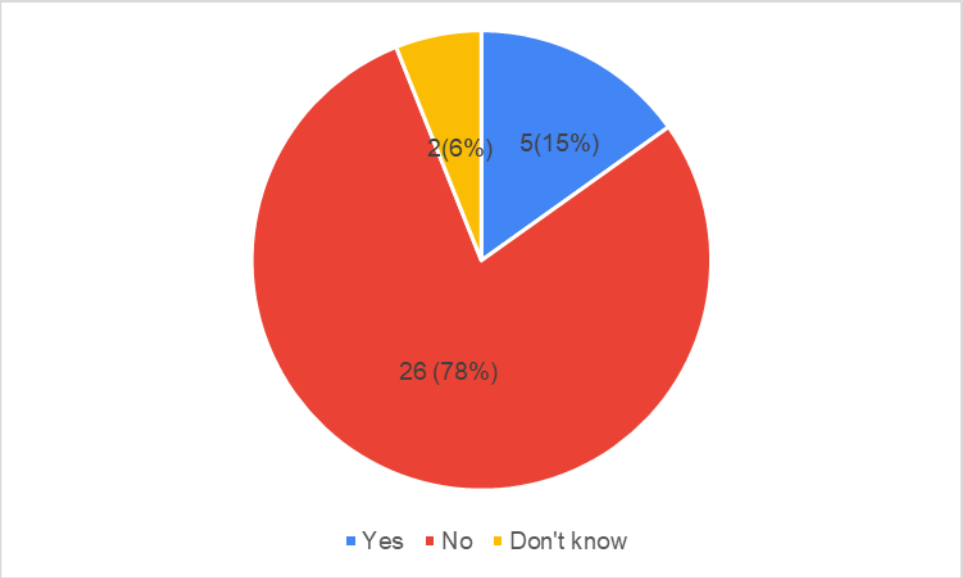

Figure 7. ***Have you ever developed a data management plan (DMP) for your research? N=33***

From 33 participants, 26 researchers (78%) indicated that they have never developed a DMP for their research. Five researchers said they had developed a DMP for their research, and two researchers said they did not know whether they had developed a DMP or not.

Question 7 was a follow-up question from Question 6. The researchers who answered yes to the previous question were asked if they experienced challenges when developing their DMPs. Five researchers participated on the follow-up question which gave them a text field to input their answer. One researcher indicated that there was a problem retrieving the DMP. The other indicated that there were "no resources", although the researcher did not explain: it is assumed that they are referring to resources to develop DMPs. Two researchers indicated that they did not experience any challenges while developing their DMPs. One researcher wrote N/A, presumably meaning that they meant they experienced no challenges. Only a small number answered this question, but from these answers, one can see that the development of DMPs is not compulsory at UKZN because if it were, all participants would have proceeded to this question.

4.3.3.3. RDM - how researchers manage their data

In order to establish where researchers store their research data, respondents were asked to choose from a list of storage options. They could choose multiple options. Figure 8 shows that most respondents (29; 87%) store their research data on a computer hard drive or on USB devices (24; 72%). Just over half of the respondents store their data in emails (17; 51%), followed by researchers who store on the cloud (10; 30%), then those who store on paper (9; 27%).

Six respondents (18%) revealed that they store their research data on a campus server, followed by two who store them on a shared drive, then one who stores in a data repository. This question had an option for "other" and one researcher indicated that he stores living materials in seed stores, in 80°C refrigerators. One specified that they utilise cloud storage with Google Drive.

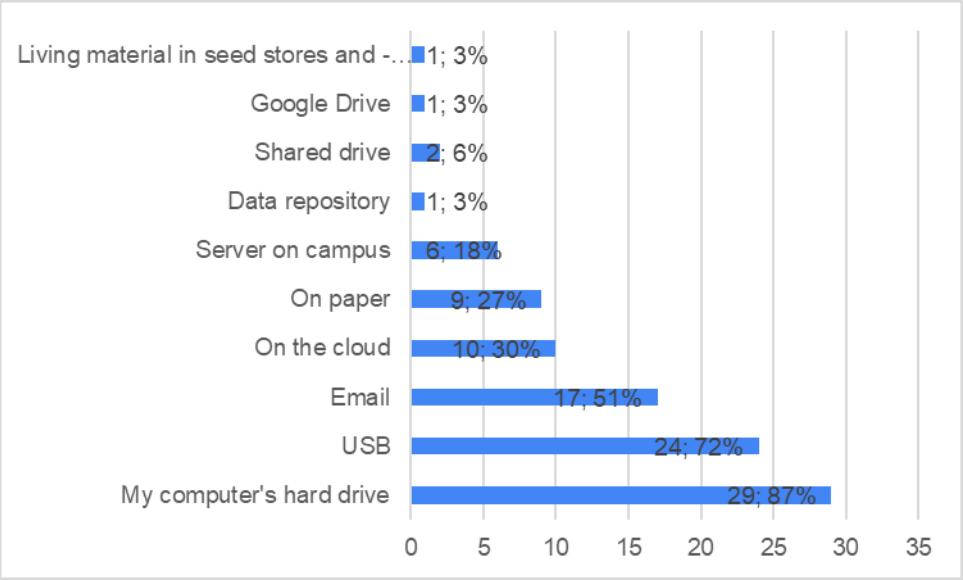

Figure 8. Research data storage N=33

To establish the frequency at which research data are backed up, the researcher asked how often the respondents back up their data.

Figure 9 shows that most (ten; 30%) respondents back up their data daily, followed by seven (21%) who back up weekly, five respondents who back up only when necessary (ad-hoc). Three respondents indicated that they back up their data hourly, while two back up annually. Lastly, there was one researcher who never backs up, and two who did not know whether they back up or not. On the previous question (Question 8), some respondents indicated that they store their data on the campus server or Google Drive. Perhaps these respondents "don't know" whether or not they back up because they expect their data to be backed up by campus IT or Google or other cloud services.

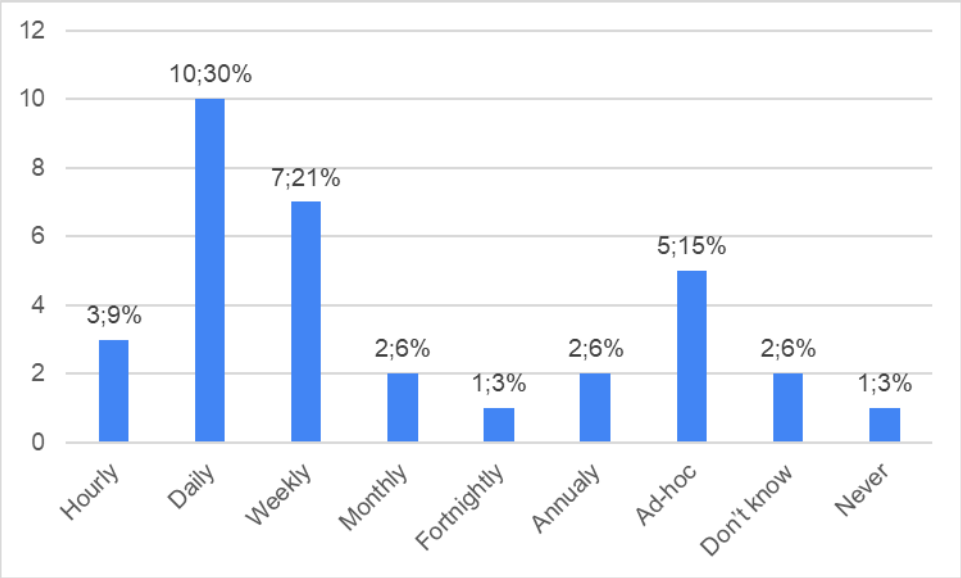

Figure 9. ***Frequency that data is backed up (N=33)***

Question 10 was a follow up question from Question 9, to determine where researchers backup their data. The question allowed researchers to select more than one answer.

Figure 10 shows that a little over half of respondents backup their data on external hard drives (18; 54.5%), some utilise USB devices (15; 45.5%) for backup. Furthermore (13; 39.4%) respondents backup their data on the cloud, followed by five (15.2%) who utilise a server on campus. Lastly, three do not backup their data.

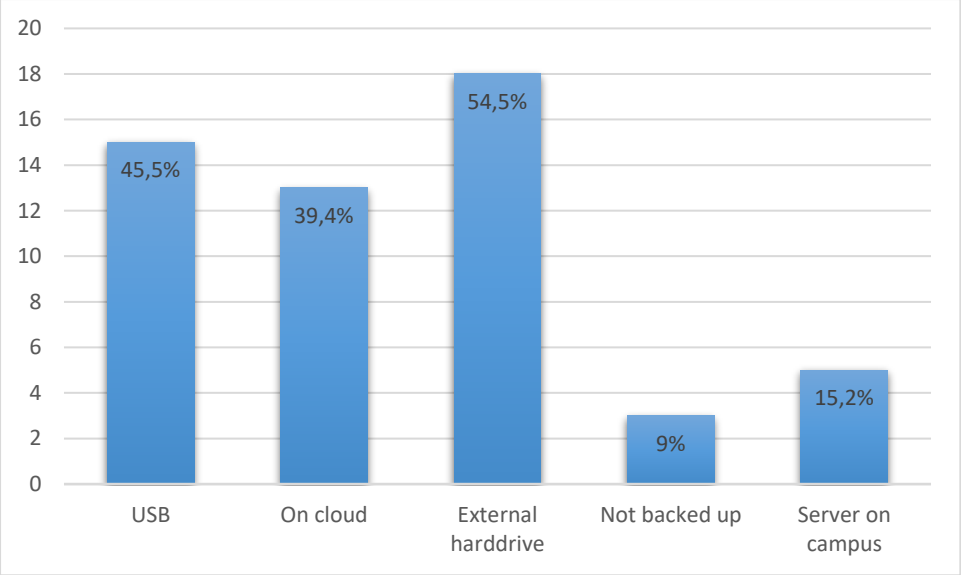

Figure 10. ***Where data is backed up N=33***

To determine how researchers manage their data, they were asked whether they had a plan of recovery in such case that they lost their data. Figure 11 indicates that 15 researchers (45%) reported having no plan on how they would recover their data if they were to lose them, though thirteen researchers (39%) indicated that they did have a plan. Six (18%) indicated that they were not sure.

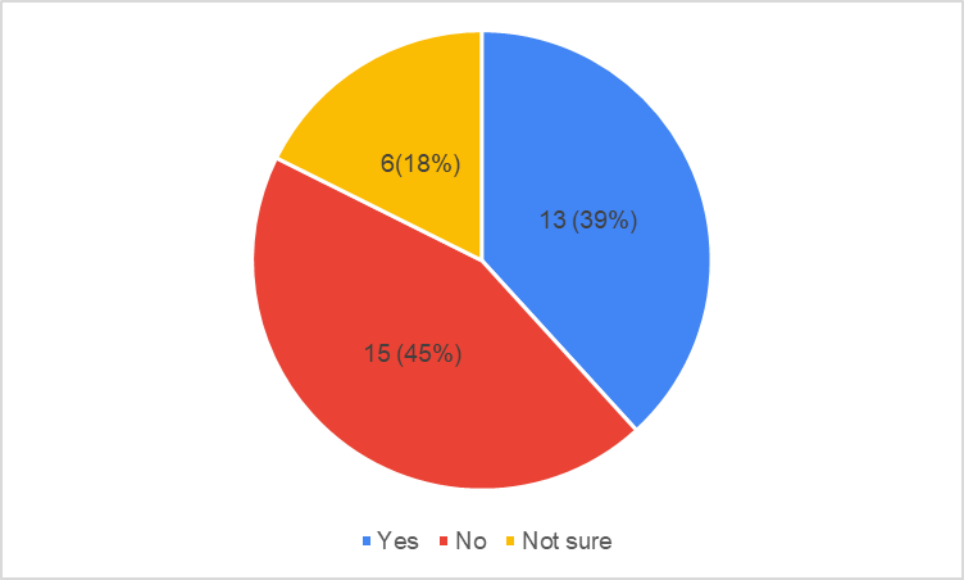

Figure 11. ***Do you have a plan for how your data can be recovered in case you lose it? N=33***

Question 12 concerned metadata. Acknowledging that not everyone would be familiar with the term, an explanation was provided. Figure 12 shows that most researchers (18; 54%) did not assign metadata to their data. Six researchers (18%) indicated that they did assign metadata, while six (18%) indicated that they assign it sometimes. Three researchers revealed that they did not know whether or not they assign, showing, perhaps they are not aware of what it is, or maybe they work in teams and do not know if someone else takes care of it.

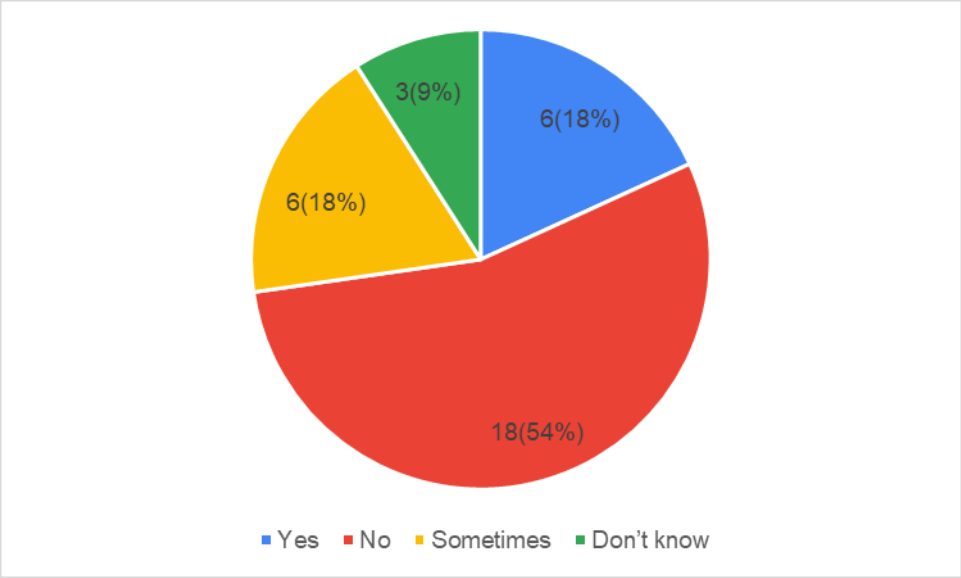

Figure 12. ***Do you assign metadata to your data? N=33***

Question 13 was a follow-up question to (Question 12) to establish whether researchers follow policies or guidelines when assigning metadata to their data. According to responses to Question 12, only twelve researchers should have answered question 13 (those who indicated they assigned metadata, if only sometimes). However, 13 researchers indicated that they do not follow policies or guidelines when assigning metadata. It is not clear how an additional researcher answered "no" as only 12 should have answered this question. This question however, had an option for researchers to choose "not applicable, metadata not assigned". Two researchers indicated that they follow policies, three that they follow policies/guidelines sometimes and 11 researchers selected "not applicable" because they do not assign metadata. Two researchers do not know whether they follow policies or not.

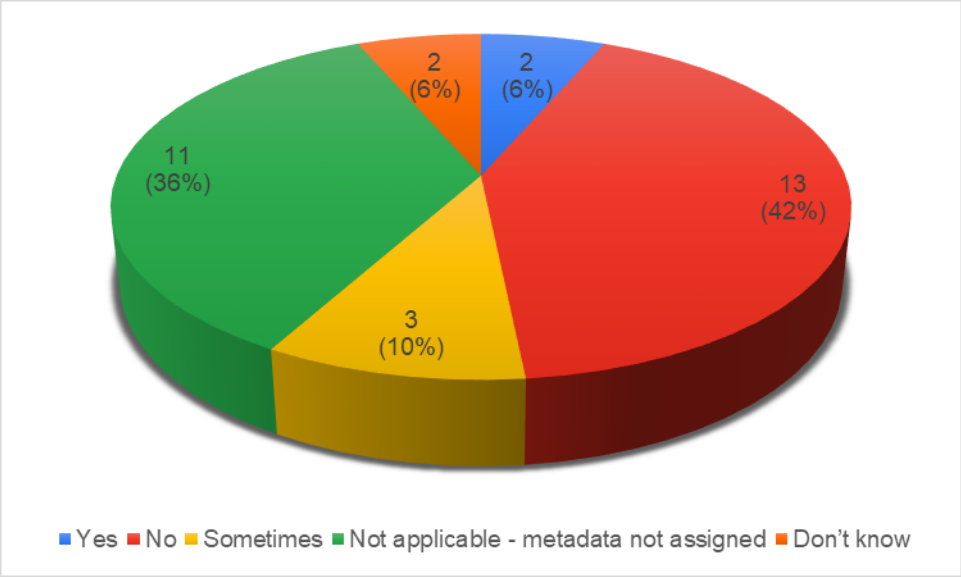

Figure 13. ***Policy/guidelines followed when assigning metadata N=31***

Question 14 served as a follow up question to Question 13, asking those who said they did follow metadata policies or guidelines to share which ones they had followed. Four researchers responded to this follow-up question. A text field was provided for researchers to share the policies followed. Three respondents wrote N/A presumably because they did not assign metadata, and did not follow any policies or guidelines. One researcher indicated that they followed the policies of the institution.

Question 15 served to establish whether researchers shared their data, and with whom. Before posing the question, an explanation was provided about what data sharing means and where/how data could be shared. Respondents were provided with a list to select from and they could select more than one option. The question also had "Other" as an option, so as to determine more about with whom they share their data. Figure 14 shows that 12 researchers (36%) share their data with their study supervisors, followed by 11 (33%) who share their data with no one. seven researchers (21%) share their data with their fellow researchers, while six researchers (18%) share with colleagues on campus and others in the research unit. Figure 11 further shows that two researchers share their data with funders and lastly one researcher revealed that they share their data with the general public. Only one researcher selected "other", but did not provide specifics. Bivariate analysis was performed to determine which college shares their data with public. Cross-tabulation revealed that a researcher from the College of Health Sciences in the School of Nursing and Public Health shared their data with the public.

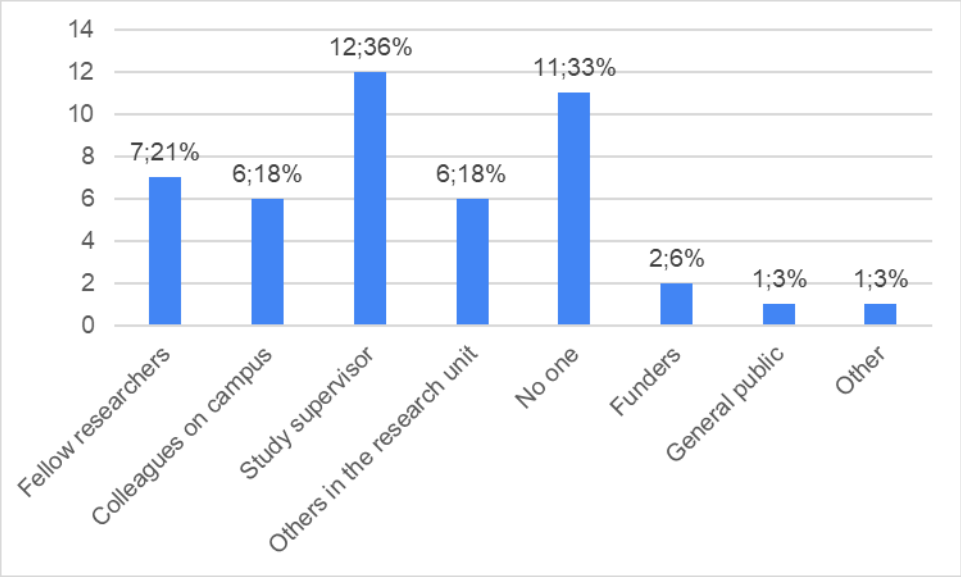

Figure 14. ***Data sharing N=33***

Figure 15 shows the proportion of researchers who had difficulty sharing their data: 18 researchers (55%) had never had trouble when it comes to data sharing, while 14 (42%) researchers indicated that they had trouble sharing their data half the time. Only one researcher indicated that they always have trouble sharing their data. From this question, it is impossible to determine whether respondents were simply not able to share their data from this question, or whether they have never needed to share their data. The question is a double-barrelled question and, in retrospect, should have been phrased differently.

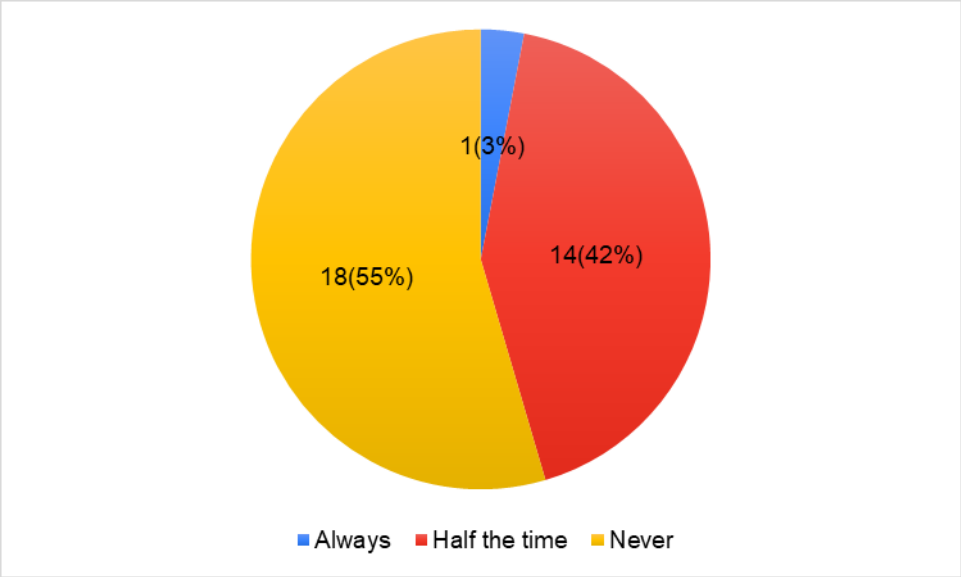

Figure 15. ***How often have you needed to share your data and not been able to? N=33***

In order to establish the ways in which researchers share their data, respondents were asked to select as many as were relevant from different data sharing options.

Figure 16 shows that the majority (28; 85%) of respondents used email to share their data. Fifteen (45%) respondents indicated that they used USBs, 5 (15%) use CDs or DVDs, while very few (ranging from one to three) utilise a data repository, cloud shared drive, file transfer protocol, a web portal for download access, and web transfer. Figure 16 further reveals that one respondent shared data by means of verbal discussion and another one indicated that they have not experienced the need for sharing.

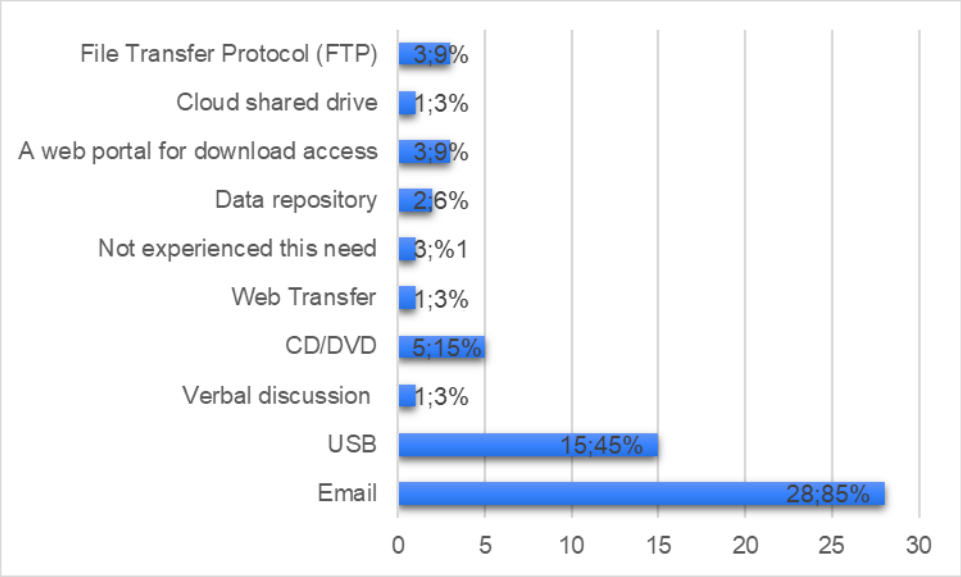

Figure 16. ***Methods used to share data N=33***

Question 17 posed two follow-up questions in order to determine more about the repositories UKZN researchers utilise, and whether they experience any challenges when using them. The questions were aimed at those using repositories; from (Question17) responses, it was therefore expected that only two would answer them. Respondents provided their answers in text fields. One researcher indicated that they utilise the Howard College data repository (Yabelana, the UKZN data repository that was launched in 2019) used by students and staff. In addressing the question about challenges experienced when using the repository, one researcher indicated that access to the repository is restricted at times, while the other researcher complained about access being slow. In total, just seven people responded to Question 18 and nine to Question 19. Their responses were “none” and “n/a”, indicating that they did not use the repository.

The next question was asked in order to determine RDM awareness at UKZN: researchers were asked as to whether they had received any RDM training. Out of 33 respondents, 25 (76%) indicated that they have not received any RDM training, while 6 (18%) researchers indicated that they had, and two researchers could not remember whether they had received RDM training or not.

Respondents were asked to share their RDM training experience in a follow-up question (Question 21). One researcher stated that they received RDM training during the NVIVO training. Another one mentioned that they received data curation training at the HSRC and said that it was a straightforward course. One researcher mentioned that they curated research data

for a research organisation in their previous job. Another researcher stated that they found being trained on how to save on the cloud quite useful, while another disclosed that the training was eye opening, but wished it were more in-depth. In total, 33 participants responded to this question.

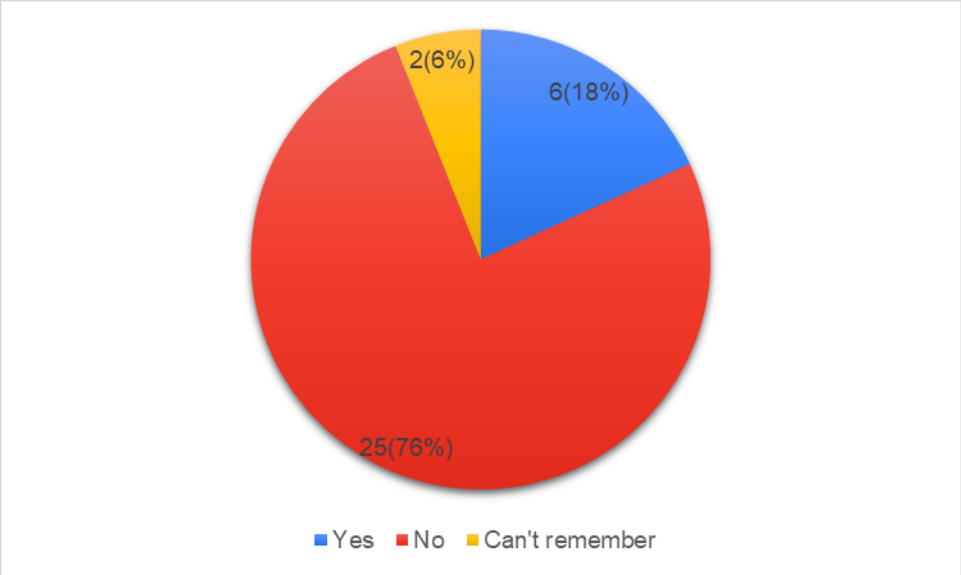

Figure 17. ***Have you ever received any research data management training? N=33***

4.3.3.4. Different RDS that could be offered by UKZN Libraries

Figure 18 shows the number of researchers who indicated interest in training in different areas of RDM. Most (23; 74%) respondents indicated that they would like to receive training in developing a DMP. Fifteen (48%) respondents showed interest in receiving training on data storage, 14(45%) on data repositories, and 13 (42%) respondents on data formatting. Furthermore, 12 (38%) respondents would like to be trained in data documentation and metadata creation, with a similar number (11; 35%) showing interest in training on data sharing and funders requirements, while 10 (32%) of them noted interest in receiving training on open access. Amongst the respondents, two were not interested in training. One researcher selected "Other" but did not specify the kind of training in which they were interested.

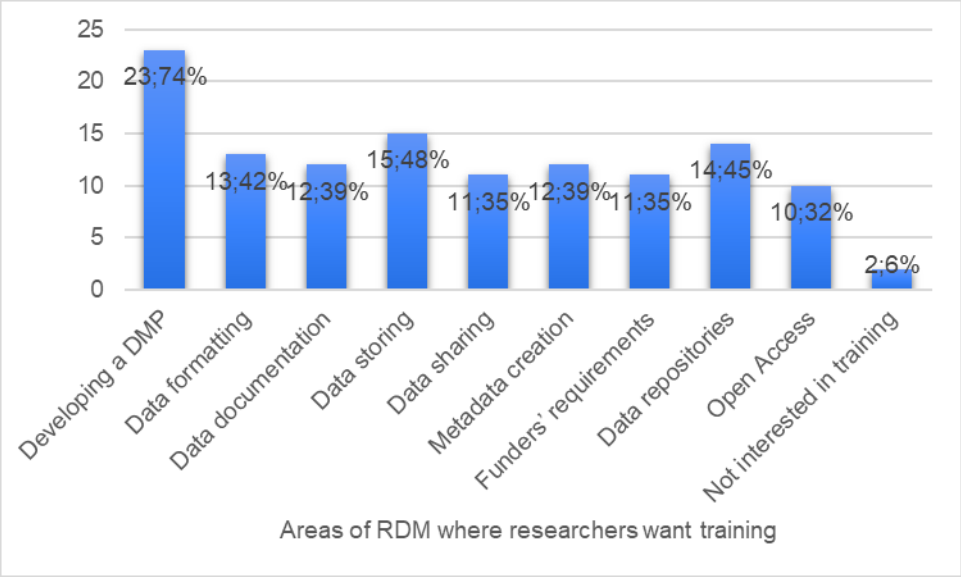

Figure 18. ***RDM training needs N=31***

The penultimate question asked respondents to rate data-related services according to how important they are to research in their discipline. The services were to be rated as "very important", "somewhat important" and "not important". If none of the choices on the Likert scale were applicable, respondents could indicate that they were "unfamiliar with this service". Responses are shown in Table 2.

Table 2. Question 23: Please rate the importance of the following data-related services to assist in improving research in your discipline

Question	Very important	Somewhat important	Not important	Unfamiliar with this service	Total responses
Having increased access to data in my discipline	**28 (84.8%)**	2 (6%)	0	3 (9%)	*33*
Having access to data in other disciplines in my institution and other institutions	**18 (56.3%)**	10(30%)	2 (6%)	2 (6%)	*32*
Guidelines and services supporting researchers in managing their research data	**24 (72.7%)**	8 (24.2%)	1 (3%)	1 (3%)	*33*
Services supporting researchers depositing their research data in digital data repositories	**22(66.6%)**	8 (24%)	1 (3%)	2 (6%)	*33*
Services supporting researchers in creating metadata	**23 (69.6%)**	8 (24%)	1 (3%)	3 (9%)	*33*
Services supporting researchers to use correct data naming conventions	**22 (66.6%)**	8 (24%)	1 (3%)	2 (6%)	*33*
Having teaching and learning materials supporting researchers to work with data	**23 (71.8%)**	7 (21%)	1 (3%)	1 (3%)	*32*
Having the services necessary to assign a digital object identifier (DOI) to my data so other researchers can find and cite my data]	**24 (72.7)**	4 (12%)	3 (9%)	2 (6%)	*33*
Having the infrastructure at UKZN to allow UKZN data to be preserved and made available to others]	**27 (81.8%)**	5 (15%)	0	1 (3%)	*33*

It is worth noting that all services listed on the table were “very important” to most respondents. To sum up responses:

The majority of respondents (28, 84.8%) indicated that having increased access to data in their discipline is very important.

Most respondents (27, 81.8%) indicated that having the infrastructure that allows UKZN data to be preserved and made available to others is also very important.

Most respondents (24; 72.7%) indicated that data related services such as having guidelines and services supporting researchers in managing their research data and services necessary to assign a digital object identifier (DOI) to their data very important.

Furthermore, 23 (69.6%) indicated that having services supporting researchers in creating metadata and having teaching and learning materials supporting researchers to work with data was very important.

Twenty-two (22) respondents, equating to (66.6%), regarded services supporting researchers to deposit their research data in digital data repositories and services supporting researchers to use correct data naming conventions to be very important.

Eighteen (56.3%) respondents rated having access to data in other disciplines in their institution and other institutions very important.

A few respondents ranging from (10 -2) revealed that they found all the data-related services in table 2 somewhat important, while others ranging from (3-1) revealed that they were not familiar with the services. Other respondents ranging also from (3-1) rated the above services unimportant.

4.3.3.5. General views about the implementation of RDS at UKZN

The purpose of the final question served to determine the views of UKZN researchers with respect to the implementation of RDS. Researchers were provided enough space to share their views. Eleven researchers answered this question, but only four stated their views. The other seven researchers had no views.

One researcher mentioned space as one of the challenges RDS implementation would face, while the other one expressed a longing for more lessons on how to handle data. Another researcher posed the question as to how RDS will be promoted by UKZN Libraries.

4.4. Qualitative data analysis and findings of the interviews

4.4.1. Interview responses: Overview

In this section, the data collected from interviewing the respondents are presented and analysed. Respondents are referred to as R1, R2, R3, and R4. After providing some demographic information about respondents, data are presented according to the themes explored in the interview (rather than the order in which questions were asked). Themes have been categorised and presented under the following sub-sections:

1.) RDM awareness (4.5.3);

2.) Implementation of RDS at UKZN (4.5.4);

3.) Possible RDM services (4.5.5);

4.) Researchers and their data (4.5.6);

5.) Institutional capabilities (4.5.7);

6.) Challenges anticipated in implementing RDS at UKZN (4.5.8); and

7.) Any other issues related to the implementation of RDS at UKZN (4.5.9).

4.4.2. Preparation for data checking

Collected data was assessed and transcribed in Microsoft Word. All four library staff members participated and answered all questions.

4.4.2.1. Demographic information

The respondents were three full-time librarians and one IT Technical Support Consultant working in the library. One of the full-time librarians that was interviewed was an Institutional Repository Librarian and two were Subject Librarians. The IT Technical Support Consultant

was chosen because he works in the library and as an Institutional Repository Librarian at times. His knowledge of research and of the library system deemed him a suitable interviewee.

The Subject Librarians support Teaching and Learning in the institution by liaising with academics, making sure that users' needs are met. They are also responsible for providing an appropriate resource collection that supports teaching, learning and research, and for training and educating students, staff and researchers in the principles, knowledge, techniques and efficient use of information. Institutional Repository Librarians are responsible for managing and administering the institutional repository and the data repository; marketing the repositories; and ensuring that the research conducted by UKZN students and academics is uploaded and archived in the two repositories.

The IT Technical Support Consultant is responsible for providing Information Systems support to the Library on a broad variety of technologies, including, but not limited to Library and Information Systems, server technologies, and database management systems. His responsibilities also include keeping up-to-date with current technology trends and coming up with innovative ways to improve Library Information Systems.

Three respondents hold an honours degree in Library and Information Science, while one holds a national diploma in Information Technology.

4.5. RDM awareness

Respondents were asked about their understanding of RDM. When respondents were approached, the nature and purpose of the interviews was explained. As a result, it could be noticed that some respondents prepared their answers in advance, especially the question about RDM (Question 3). Generally, the respondents showed an appreciable level of understanding of what RDM is, as can be seen in the extracts below (cited verbatim). R1 stated that to fully be aware of what RDM is, one has to be introduced to research and research processes:

"It will take a person who has been introduced to research as well as a person who has done research to fully understand what RDM is… but one of the processes is data collection. The data collected through the research project is considered valuable and can be used by other researchers. Research Data Management is a process of preserving, hosting, sharing, and managing the research data digitally or online… and by assigning metadata, they will ensure retrieval." [R1]

As can be seen, R1 was quite knowledgeable about RDM. R1 has been introduced to RDM through short courses and continuing professional development programmes offered by the University of Pretoria.

The other respondents also showed some understanding of RDM but were not as detailed as R1:

"RDM entails the tools and resources involved in organising, storing, preserving, and sharing data collected for a particular research project." [R2]

"RDM makes the research accessible quicker to other researchers." [R3]

"I believe RDM has to do with how researchers deal with their research data, like how they collect, store and share their data…. some institutions have data management plans that ask researchers to write how they will be dealing with their data." [R4]

Above, R4 has also showed an awareness of data management plans (DMPs).

4.6. Implementation of RDS at UKZN

The respondents were asked as to whether there were any plans to implement RDS at UKZN Libraries. All four of them answered with a resounding yes and R1 elaborated with:

"The library recently launched Yabelana which is a Zulu word meaning sharing, given to the UKZN data repository." [R1]

R1 also mentioned the use of DSpace, an institutional repository, to promote the research output of the institution.

R4 mentioned the OpenJournals that UKZN Libraries launched in 2018. The institution hosts an Open Journal System (OJS), an open source software responsible for managing and publishing scholarly journals. R4 considered OJS an RDS because the journals hosted are open access, although OJS do not host data but peer reviewed articles.

4.7. Possible RDM services

Respondents were asked which RDS UKZN Libraries could potentially provide and the extent to which the respondents could support the implementation. They were provided with a list and

were asked to rate the extent of their support as well as give reasons for that rating. The list of RDS provided was:

a) finding, organising, managing and evaluating scholarly data;

b) data repositories and Open Access;

c) data management plan writing;

d) research data management training;

e) metadata creation;

f) data formatting;

g) data storage;

g) data sharing;

h) data archiving.

Data sharing, data storing and data archiving were the main RDS receiving support from librarians, mainly due to the data repository the library has launched. The respondents felt that the existence of the data repository would enable data storage, data archiving and data sharing. With finding, organising, managing and evaluating scholarly data, respondents felt that subject librarians partially provide that service through teaching and learning.

"For now, I support data sharing, data storage, and data archiving, because we can provide all through Yabelana." [R3]

"We already have a data repository and I am in full support so I guess I support data sharing, storing and archiving." [R4]

R1, however, said that, if UKZN does not focus on marketing first, the implementation of any RDS would be futile:

"At this point, the priority is marketing which will cover the importance of managing their research data." [R1]

The respondents had nothing more to say about the rest of the RDS listed.

4.8. Researchers and their data

Respondents were asked some questions that related to researchers. One question asked respondents to describe how the library supports the research process at UKZN (Question 4). A range of answers was received to this question. Respondents mentioned "personalised librarianship" [R4], where librarians provide a personalised service to individual researchers, for example creating researchers' profiles, assisting them with Open Researcher and Contributor ID (ORCID). They stated that librarians also assist in uploading researchers' outputs and data to the institutional repository and data repository for preservation and sharing.

R3 elaborated on how librarians support researchers, mentioning spaces designated for researchers and training provided. The respondent elaborated as follows:

"The library has well experienced librarians to assist the researchers, each campus library has a research commons for researchers and postgraduate students. Subject Librarians also organise trainings each semester where they train the students and researchers on how to do literature review searches and also assist with the proposal structures."[R3]

Respondents were asked if they have supported researchers in the process of managing their data. Only one respondent had previously supported a researcher. The respondent elaborated as follows:

"I assisted a medical lecturer who publishes her articles with a journal - where they are requested to submit data of the articles before they are published in that journal." [R1]

The researcher asked respondents about DMPs, specifically as to whether researchers are required to develop DMPs, which funders require DMPs and as to whether UKZN Libraries provide training for the development of DMPs. Two respondents (R3 and R4) seemed not to know what DMPs are and kept referring to the data repository. The other two agreed on the lack of DMP service as shown in the quotes below.

"I don't know which funders need DMP. Currently the library is not helping with the development of DMPs." [R1]

"We have not implemented DMPs yet….it is on the plans." [R2]

The researcher asked respondents whether they agreed with the growing need for data sharing and whether the library is considering providing support for researchers to share their data. All respondents were in agreement with the need to share data and mentioned that the library is 100% behind researchers sharing their data hence the launch of a data repository.

R2 went on to state their opinion about the value of sharing:

"The purpose of conducting research is to solve problems, have new discoveries and to make an impact. That is valuable information that is not supposed to be hidden and which will assist future researchers to have data that they can refer from or utilise for their research. It can be for comparison, if certain patterns have changed."

Respondents were asked to share their opinions on the need for the library to support researchers to manage and share their data. All respondents felt it to be necessary and important for the library to support researchers to share their data. The general consensus was that the library is an information centre of the institution and therefore it makes sense that they support the management and sharing of researchers' data.

R1 said that the *"library is the information hub and a mutual place on campus visited by students and staff… librarians are experts in describing information, which is crucial for information retrieval."*

R2 shared their views on how data sharing would benefit researchers:

"Personally, I would say the library should support researchers to manage and share their data… because research is meant for sharing and contribute to developing knowledge in a field or study. So, by supporting researchers manage and share their data the library will be making their research journey easier as we know how difficult it is sometimes to be a researcher."

4.9. Institutional capabilities

Questions were posed on staffing, knowledge and skills, and resource allocation (funding) for RDS at UKZN. Among the library staff interviewed, there is a general consensus that the current library staff are not knowledgeable enough to support researchers in RDS. When the researcher was an employee at UKZN, and began embarking on this study, she had informal

discussions about RDM with librarians and most of them were not aware of RDM. Two of the respondents, when asked whether they consider the current library staff knowledgeable and suitable to support researchers in RDS, just stated "no", but one did elaborate on their negative answers as shown below.

"No, they don't possess the skill, because some don't even know what RDM is – but I believe with training and more focus on it, they can do it." [R1]

The other two respondents believed the current library staff are knowledgeable and suitable to support researchers in RDS. However, when asked to elaborate, R3 acknowledged that currently *"it might seem as if they are not suitable because they have never been asked to implement RDS."*

The respondents were asked whether the library has any resource allocation to facilitate RDS. Responses were mixed:

"Yes and no. Yes on the backup server since the cloud system was selected (Figshare), the university had to have its backup server for security, and no in terms of staffing and marketing materials." [R1]

"We have launched the software (Yabelana) but it does not look like we have enough resources to proceed with the project." [R4]

When the researcher asked for views, R1 revealed that UKZN library staff need a change of mindset.

"What I can propose is change of mindset – librarians need to adapt to change and stop clinging onto job descriptions. Sometimes it's not a matter of being knowledgeable about the subject but being willing to learn and easily adapt to change."

4.10. Challenges anticipated in implementing RDS at UKZN

Respondents were asked whether they anticipate any challenges (other than those already noted) with the implementation of RDS at UKZN. Three of the respondents mentioned that they foresee challenges, while one thought there would not be any challenges. The challenges

ranged from library staff not being knowledgeable about RDM, to librarians resisting technology, to advocacy and support for RDS. For instance, R2 indicated that the library might have a "hard time" marketing RDS in the institution. R3 mentioned that researchers might be sceptical about sharing their research data in the data repository. When probed as to why researchers would be sceptical, it was established that scepticism could result from lack of advocacy for and proper knowledge of RDM from librarians. The respondent felt that it would not be easy for librarians to advocate for a subject that they are not fully knowledgeable about. Furthermore, there is a general consensus that researchers do not feel comfortable sharing their data, because they might still want to use them themselves in future studies.

R1 mentioned a challenge that would be posed by librarians' lack of interest regarding technology and researchers' interest in being compensated for their research data:

"Resistance from librarians who do not want to embrace technology and also resistance from academics who want to money out of their research." [R1]

R2 mentioned the shortage of staff as one of the challenges that could be faced when implementing RDS at UKZN.

"We are always short-staffed in the library. I assume Subject Librarians would be the ones taking this role but they have a lot to do already. And advocacy might be the problem; the library might have a hard time marketing the importance of RDS and getting support." [R2]

On the other hand, R3 anticipated uncertainty from researchers, mentioning how they might feel about their data:

"Academics may feel that the data is still important for them, for further research." [R3]

4.11. Any other issues related to the implementation of RDS at UKZN

The researcher asked respondents to discuss issues related to RDS implementation that might not have been addressed by the interview schedule. Three respondents had nothing to add and one mentioned collaboration between librarians and researchers to support Open Access.

"If researchers and librarians can support Open Access to prevent predatory journals to benefit and exploit researchers, also to end the myth that Open Access research is of the less research impact." [R1]

4.12. Summary

Chapter Four presented collected data from researchers and library staff. Quantitative data was analysed and presented using charts and tables while qualitative data was presented according to themes using quotations as illustrations. The study asked researchers and library staff to share their views regarding the implementation of RDS at UKZN.

Chapter Five: Data interpretation, recommendations and conclusion

5.1. Introduction

Taking into account the data analysed in Chapter Four as well as the literature reviewed in Chapter Two, this chapter will respond to the research objectives. After that, it will make recommendations from the findings, discuss study limitations, suggest further studies and conclude.

This study set out to determine the role of UKZN Libraries in the implementation of RDS at the UKZN. The study aimed to achieve the following objectives:

1. to identify the need for research data services among UKZN researchers;
2. to identify the major challenges associated with introducing research data services at UKZN; and
3. to determine the possibility of implementing research data services at UKZN Libraries.

The objectives will be addressed one by one.

5.2. The need for research data services among UKZN researchers

As seen in Figure 14, the majority of researchers indicated that they have never been offered RDM training at their institution. The data indicates the need for training in a number of RDM-related areas as well as for other RDS.

Most researchers have indicated that they do not assign metadata to their data. This is not a good RDM practice, as metadata creation enables easy data accessibility. A bulk of researchers have expressed the need for services for metadata creation, indicating that they might be aware of its importance. A metadata service that offers training on how to create metadata is a necessary RDS need at UKZN.

Both researchers and librarians agreed on the need for data sharing. Librarians indicated that they would be able to respond to this need through their data repository. There is evidence that health science researchers are being encouraged to make their data shareable. For this reason, this RDS would be important to them. Researchers expressed the need for training in data repositories. Some researchers rated services supporting researchers depositing their research data in repositories as very important. Furthermore, only a few researchers used the repository to share their data. Researchers have shown an interest in getting increased access to data in

their discipline, access to others' data, preserving data, and assigning DOIs. This indicates their openness to RDM and the need for RDS and all these needs are associated with a data repository. Librarians saw the data repository helpful for data sharing, data storage, and data archiving.

Most researchers indicated that their formats and software enable sharing and long-term access. This is a good RDM practice. Consequently, though, the study found that their storage is erratic, where 42% of researchers noted having trouble sharing their data before. It is somehow doubtful that their chosen formats and software enable sharing and long-term access, even if they say so. In addition, Delasalle (2013) mentions the researchers' reservations when it comes to data sharing as one of the issues at the University of Warwick. There is therefore a need for training in format and software obsolescence, as well as advocacy and training in data sharing.

One researcher expressed the need for data handling lessons. These are all the needs that UKZN Libraries ought to be aware of in preparation for RDS implementation.

The study has found that researchers are not quite aware of RDM. This can be seen in the irresponsible manner in which they store and back up their data, such as storing their data on computer hard drives, cloud services, and removable storage devices, and mostly using external hard drives and USB's to back-up their data. It is presumed that researchers are not aware of the risks and of the potential other storage places, like repositories, might provide for sharing. Researchers showed interest in training about data storage presumably because they are aware that there might be better ways to store their data. Chiware and Mathe (2015:1) stated that South African libraries are making progress concerning RDM awareness and campaigns. That is not the case with UKZN where researchers are not aware of RDM and librarians showing moderate understanding of RDM. There is therefore a need for RDM awareness campaigns at UKZN Libraries.

It is clear that researchers at UKZN who responded to this research produce different types of data in large volumes. This was seen in the small sample of this study, meaning that data types must be even more diverse across the institution and library staff will have to be aware that one size will not fit all in terms of RDS. Section 4.5.3. of Chapter Four reveals that researchers from different disciplines are managing their data using various storage devices. There is therefore a need for data storage that caters for all data types.

The study found that the majority of researchers do not have a plan regarding how to recover their data if they were to lose them. Therefore, showing that researchers might lose their research data through obsolescence or misplacing their removable storage devices. These findings speak to ineffective practices, and could lead to valuable data not retrievable.

Library staff indicated that UKZN Libraries have not yet developed DMPs. This was confirmed by 74% of researchers, eager for training in developing a DMP. The data shows that DMPs are not compulsory at UKZN, however, some researchers, (although not many) have developed DMPs. DMPs make it easier for researchers to detail how data will be generated, stored and eventually shared. To ensure that researchers achieve more from their research, planning for efficient data creation, management and sharing are of importance. There is therefore a need for training on how to write DMPs, how to access and populate a DMP tool.

Although there are researchers with no recovery plan, it is worth mentioning that quite a few (39%) researchers have a recovery plan for data loss. This is good RDM practice, as they will be able to prevent data loss and ultimately share their data. Although these researchers have a plan for data recovery, it is crucial that all stakeholders interested in taking part in RDM care about research data storage and ascertain that research data are stored in a manner that makes them retrievable in the long term, as indicated by De Waard (2016:51).

5.3. Challenges anticipated in implementing RDS at UKZN

This study found some potential challenges regarding RDS implementation by UKZN Libraries. One of the challenges revealed is RDM awareness among staff. Library staff showed a moderate level of awareness. Although some library staff did not seem as knowledgeable as others did, their responses showed that they are generally aware of the concept.

Regarding a question about various data-related services, a few researchers showed a lack of awareness about them. For example, some researchers indicated that they were not familiar with services that some deemed very important, such as having increased access to data in their discipline, having the infrastructure that allows UKZN data to be preserved and being made available to others, and having the services necessary to assign DOIs and services supporting researchers in creating metadata. Researchers' lack of awareness of RDM might prove a challenge for RDS implementation as they might find these services futile.

One librarian and one researcher anticipate that RDS advocacy by UKZN Libraries will be a challenge. Although the researcher in question only mentioned this and did not elaborate, librarians' concern was focused on the lack of knowledge and skills among staff. A concern was that it would be difficult for librarians to advocate for a concept about which they are not knowledgeable. Cox and Pinfield (2014:308) note that librarians might lack the necessary knowledge and skills to be key players in the field of RDM.

Another challenge was with regards to space, where one researcher mentioned that there is not much space. Although the researcher did not mention the kind of space they were referring to, I presume that they are referring to storage space. The issue of space was also mentioned by Tenopir et al. (2015), who revealed a concern about the library not being able to sustain the demand for server space.

One librarian identified researchers being uncertain about sharing their data in the institutional data repository as one of the challenges. This is confirmed by the quantitative data, where out of 32 researchers, only two share their data using a data repository. This finding is similar to one of the challenges identified by Delasalle (2013) on researchers' uncertainty about the value of data sharing at the University of Warwick.

Other challenges mentioned by library staff were resistance from librarians to embrace technology and staff shortages. Technology resistance by librarians was mentioned as a potential challenge that could hinder RDS implementation. One respondent urged librarians to change their mindset and embrace change. Cox, Verbaan and Sen (2012) mention existing librarian roles that come second to support RDM, addressing the issue of staff shortage to note that it implies that the existing library staff will assume the new role of supporting RDS.

Resource allocation is one of the challenges when it comes to RDS implementation at UKZN. Perrier et al. (2018) mentioned the lack of resources and funding as a challenge that hinders the availability of infrastructure for RDM. When library staff were asked whether there are any resources allocated to facilitate RDS implementation in the future, the majority had no idea, and one librarian mentioned that UKZN Libraries launched the data repository, however, do not seem to have enough resources to proceed with the project. This shows an understanding that some RDS are needed, but there is an inability to sustain or develop them further once set up. This could be because of the lack of planning or lack of expertise.

Gaining support from management, capability of staff to manage RDS and the ability to meet researchers' evolving needs are common challenges around implementing RDS (Raboin, Reznik-Zellen & Salo (2013).

5.4. Possible research data services at UKZN

There is a wide range of RDS that could be implemented and offered by UKZN Libraries. This study found that researchers were interested in a wide variety of services. Policy development would be a necessity to discover how libraries can work close with the funders. Policy development would ensure that RDS are used and also help in establishing ways to reward researchers who engage in data management. Jones, Pryor and Whyte (2013:5) mentioned that the university leadership has a role to analyse policy requirements at different levels such as national, institutional and funder levels. Flores et al. (2015:85) considers researchers' involvement in policy discussions as essential for their awareness of the decisions made concerning their data.

The development of a DMP tool is an important RDS that could be offered by the library as funders require DMPs with grant proposals. According to Jones, Pryor and Whyte (2013:11), libraries can guide researchers in creating DMP templates and customising DMP tools. Libraries may offer a consultancy service to assist researchers with DMP tools.

Researchers and librarians seemed to share the same sentiments when it comes to data storage. Data storage, archiving, and sharing were the main services librarians were interested in providing. Librarians felt they would be able to offer these services because of the existence of Yabelana. The library may offer a consultancy service on how to use Yabelana as a data storage guided them as to how the repository works as a data archive, and the sharing the repository offers to enable use and reuse of datasets. During the consultations, librarians may share the benefits of data sharing which include improved citation rates and data visibility (Jones, Pryor & Whyte, 2013:2). Data storage is one of the activities of the lifecycle that must be carried out in order to enable the curation and preservation process.

As has been noted, offering training on many RDM aspects is needed. UKZN Libraries could offer training for the following researcher needs:

- The library can satisfy the need for long-term and archival storage by offering training on how the data repository can be utilised as a long-term and archival storage. Tenopir et al., (2017:25), when listing some of the RDS libraries could offer, mention the creation of institutional data repositories to store datasets in the long term. These services could be offered at post-project stage, as suggested by Jones, Pryor & Whyte (2013:2). Librarians can guide researchers as to long-term data selection, train them on how to archive, and make data accessible to particular users.
- The library can meet the need for researchers depositing their research data in the digital data repositories of offering them training regarding how to deposit their data into the repository.
- UKZN Libraries could offer training on how to create an ORCID, identification of a suitable metadata schema, and training on how to assign metadata. The study has shown that very few researchers assign metadata to their research data. Creation of metadata is essential in the process of RDM, therefore, it is crucial for UKZN Libraries to train researchers to assign metadata.
- Researchers considered having the services necessary to assign a DOI to data important. The data repository that UKZN Libraries utilises already has the functionality to create DOIs. Librarians may inform researchers of this functionality and train them on how to create DOIs. This could work on their advantage to advocate for data repository.

Libraries could offer consultancy services for the following needs:

- UKZN Libraries can respond to the need for assistance with data formatting and data documentation by offering a related consultancy service. The collected data shows that UKZN researchers work with different types of data, therefore, guidance in appropriate data formats is an essential service and may need to be offered on a one to one basis, rather than in group training.
- Data sharing - this need can be satisfied by providing a consultancy service. The library can take advantage of this service to market the data repository as it enables data sharing. The library can also explain the benefits of data sharing to the researchers.
- Funders' requirements – this need can be met by offering consultancy services. The library can explain funders' requirements regarding DMPs. Therefore, librarians ought to have a wide knowledge of what different funders want with regards to DMPs.

- Guidelines and services supporting researchers in managing their research data. Consultancy services with researchers may be associated with this need. Librarians may use a libguide for to provide guidance and advice on all sorts of aspects of RDM. This is what the University of Pretoria (UP) is doing to direct their patrons to relevant RDM information, according to Van Wyk (2016).

The findings of the study reveal that researchers and librarians agree that effective RDM is crucial to the research process and thus UKZN should implement some RDS.

5.5. Implementation of RDS at UKZN

This section explores the possibility of offering the RDS mentioned above at UKZN. Implementation of RDS might be easier if UKZN makes RDM part of the institutional structure planning as well as having a steering committee for policy implementation. Jones, Pryor & Whyte (2013:3) support the need for a steering committee, and encourage the university leadership to involve a pioneer at the level of the DVC to advocate and be in charge of institution's RDS steering committee. As described by Jones, Pryor and Whyte (2013:5), policy development ought to include university leadership. Policy development is a starting point when it comes to RDS implementation. UKZN Libraries are currently in a process of formulating a data policy. Jones, Pryor and Whyte (2013: 5) state that a policy represents the principles that will direct the decisions and actions to accomplish what the stakeholders have been agreed upon.

Adequate funding and time allocation would need to be made available to build RDM capacity at UKZN Libraries so that RDS could be implemented. Although it would be essential to appoint a permanent staff member to oversee RDM skills development at UKZN Libraries, reskilling of existing staff would also suffice as a starter. According to Lewis (2010:145, cited in Corrall, 2012:3), data management should extend naturally to the existing role of libraries in offering access to the published research. The current library staff at UKZN Libraries do not possess the necessary skills and knowledge to fully support RDS. Although there are current skills that could be utilised, for example, research support by subject librarians, these are not enough for the RDS suggested, therefore, there is a need for capacity building so as to enhance their understanding as well as obtain new and relevant skills to support RDS. One of the objectives of this study is establishing the possibility of implementing RDS at UKZN. Lack of skillset will make it difficult for the library to implement RDS, however, a change of attitude

would play a great role. Librarians will have to embrace change and actually be willing to be re-skilled. One of the librarians mentioned the need for change of mindset and adaptation to change.

UKZN Libraries have not yet developed a DMP tool. When they do, the incorporation of DMPs into internal grant applications might help UKZN emphasise the importance of this service. Seeing that most funders require DMP tools at a proposal writing stage, it is necessary that the library be ahead and incorporate these at the beginning of the application process. UKZN could look into the DCC's DMPonline or current South African templates that are already in use. For example, the UCT DMP template: (https://dmp.lib.uct.ac.za/public_templates).

UKZN will have to market everything that they want to implement, otherwise it will not be used. For example, UKZN Libraries to market Yabelana to both researchers, who will upload their data, and those who will find their data in the repository and encourage them to make use of it. This service will make researchers' data more accessible to other researchers in the institution (and beyond) as they would know where to go to access other researchers' data.

In as much as some RDS that UKZN Libraries will have to implement need resource allocation, they are ahead in certain aspects. For example, they have a data repository that will serve the purpose of data storage, data sharing, and publishing.

It is vital for UKZN Libraries to ensure that researchers have the best RDS to enable data re-use and sharing and subsequently strengthened research at UKZN.

5.6. Study limitations

The researcher acknowledges that the study was met with a number of limitations, which included the following:

- Being a case study, the study was delimited only to UKZN researchers and librarians, therefore it cannot be generalised to other institutions.
- Questionnaire responses were few in number, hence the findings cannot be generalised to the entire institution.
- Colleges and schools were not equally represented; some had very few respondents, where no conclusions could be drawn.

- During the time of data gathering, the institution was closed due to student protesting and national lockdown due to Covid 19. The country was put into various levels of lockdown to contain the spread of Coronavirus. This affected the study in terms of contact with researchers. The online questionnaire was posted in the university notices and there is no way of knowing that all researchers saw it as they were not at work.
- Because this is a minor dissertation, time constraints made it impossible to run a pilot study to test reliability.

5.7. Recommendations for further studies

The recommended future studies are:

- A comparative study of RDM practices of researchers in various disciplines, to provide more understanding on different data types generated by UKZN researchers and how they manage their data.
- A mixed methods study to collect comprehensive data on researchers' perceptions about RDS. A mixed method would not only collect quantitative data but also qualitative data through, for example, interviews, to fully capture researchers' understanding about RDS.
- A qualitative study to collect a detailed information on the perceptions of librarians with regard to RDS.

5.8. Conclusion

This research study about the role of the library in implementing research data services was the first to be conducted at UKZN. The study is believed to have provided the initial step towards filling the knowledge gap by investigating the need for research data services among UKZN researchers, the major challenges associated with research data services at UKZN and the possibility of implementing research data services at UKZN libraries. The study investigated the role of the library in the implementation of RDS, using UKZN Libraries as a case study. The study used a mixed methods approach. Researchers' needs in terms of RDS were investigated quantitatively by means of an online questionnaire and librarians' opinions on RDS implementation by means of interviews.

The study found that, while there are some RDS already at UKZN, researchers would value the introduction of some RDS that catered for the long-term preservation of data, development of DMPs and metadata creation, while librarians believe that services related to data storage, sharing and archiving are the RDS they would be in a good position to support. A policy, proper advocacy and effective communication with all relevant stakeholders will gradually establish the discipline of properly managing research data. The library has a massive role to play in RDS but researchers' needs have to be taken into consideration, different stakeholders have to be invited, resource allocation in terms of library staff and funding have to be realised and ensured and marketing of different RDS has to be prioritised.

References

Akers, K.G. & Doty, J. 2013. Differences among faculty ranks in views on research data management. *IASSIST Quarterly*, 36(2):16-20. DOI:10.29173/iq771.

Allan, B. 2010. *Supporting research students*. London: Facet.

Association of College and Research Libraries. 2017. 2016 top trends in academic libraries: A review of the trends and issues affecting academic libraries in higher education. College & Research Libraries News. 77(6): 274-281. DOI:10.5860/crln.77.6.9505

Aurini, J., Heath, M. & Howells, S. 2016. *The how to of qualitative research: strategies for executing high quality projects*. London: Sage.

Australian National Data Services. 2011. *Annual Report 3, 30 September*. Available: https://www.ands.org.au/__data/assets/pdf_file/0009/386982/ardc_eif_annual_report_2011-12.pdf [2020, December 2].

Babbie, E.R. 2009. *The practice of social research*. 12th ed. Belmont, CA: Cengage Learning.

Bethlehem, J.G. 2002. Weighting nonresponse adjustments based on auxiliary information. In *Survey nonresponse*. R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J.A. Little, Eds. New York: Wiley. 275–287.

Bryant, R., Lavoie, B. & Malpas, C. 2017. *The realities of research data management: A tour of the research data management (RDM) service space, Part 1*. Dublin, OH: OCLC Research. DOI:10.25333/C3PG8J.

Bryman, A. & Bell, E. 2011. *Business research methods*. 3rd ed. Cambridge: Oxford University Press.

Canada. 2011. The 2011 Canadian Research Data Summit. Available: https://www.rdc-drc.ca/wp-content/uploads/Background-to-the-2011-Canadian-Research-Data-Summit1.pdf [2021, July 05]

Carey, N., Justh, N.M. & Williams, J.W. 2003. *Academic Libraries, 2000. ED Tabs (*NCES 2004-317). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Carlson, J. 2014. The use of life cycle models in developing and supporting data services. In *Research data management: Practical strategies for information professionals.* J.M. Ray, Ed. West Lafayette: Purdue University Press. 63-86.

Chiware, E. & Mathe, Z. 2015. Academic libraries' role in research data management services: A South African perspective. *South African Journal of Libraries and Information Science,* 81(2), 1-10.

Cohen, L., Manion, L. & Morrison, K. 2000. *Research methods in education*. 5th ed. London: Routledge.

Constantopoulos, P., Dallas, C., Androutsopoulos, I., Angelis, S., Deligiannakis, A., Gavrilis, D., Kotidis, Y. & Papatheodorou, C. 2009. DCC & U: An extended digital curation lifecycle model. *International Journal of Digital Curation*, 4(1):34-45. DOI:10.2218/ijdc.v4i1.76.

Cox, A., Verbaan, E. & Sen, B. 2012. Upskilling liaison librarians for research data management. *Ariadne*, (70). Available: http://www.ariadne.ac.uk/issue/70/cox-et-al/ [2020, October 27].

Cox, A.M. & Pinfield, S. 2014. Research data management and libraries: current activities and future priorities. *Journal of Librarianship and Information Science*, 46(4):299-316. DOI: 10.1177/0961000613492542.

Cox, A.M., Kennan, M.A., Lyon, L. & Pinfield, S. 2017. Developments in research data management in academic libraries: towards an understanding of research data service maturity. *Journal of the Association for Information Science & Technology*, 68(9):2182-2199. DOI:10.1002/asi.23781.

Cragin, M.H., Palmer, C.L., Carlson, J.R. & Witt, M. 2010. Data sharing, small science and institutional repositories. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 368(1926):4023-4038. DOI:10.1098/rsta.2010.0165.

Creswell, J.D. & Creswell, J.W. 2009. *Research design: Qualitative, quantitative, and mixed methods approaches.* 5th ed. Los Angeles: Sage.

Creswell, J.W. & Plano Clark, V.L. 2011. *Designing and conducting mixed methods research.* 2nd ed. London: Sage.

Creswell, J.W. 2013. *Qualitative inquiry and research design: choosing among five approaches*. 3rd ed. Los Angeles, CA: Sage.

Creswell, J.W. 2014. *Research design: qualitative, quantitative, and mixed methods approaches.* 4th ed. Los Angeles: Sage.

De Vos, A.S., Strydom, H., Fouche, C.B. & Delport, C.S.L. 2011. *Research at grass roots for the social sciences and human service professions*. 4th ed. Pretoria: Van Schaik.

De Waard, A. 2016. Research data management at Elsevier: supporting networks of data and workflows. *Information Services & Use*, 36(1-2):49-55. DOI:*10.3233/ISU-160805***.**

Delasalle, J. 2013. Research data management at the University of Warwick: Recent steps towards a joined-up approach at a UK university. *LIBREAS. Library Ideas*, 23. Available: https://libreas.eu/ausgabe23/10delasalle/ [2019, October 19].

Department of Higher Education and Training. 2018. Report on the evaluation of the 2018 universities' research output. Available: https://www.up.ac.za/media/shared/1/2020/May%202020/report-on-the-evaluation-of-the-2018-universities-research-output_april_2020.doc.zp189504.pdf. [2020, September 20]

Devlin, A.S. 2018. *The research experience: planning, conducting, and reporting research.* Los Angeles: Sage.

Digital Curation Centre. 2010. *What is digital curation?* Available: https://www.dcc.ac.uk/about/digital-curation [2021, January 28].

Digital Curation Centre. 2017. *CARDIO: About - Digital Curation Centre*. Available: https://cardio.dcc.ac.uk/about/ [2019, June 19].

Digital Curation Centre. 2017. *Data Asset Framework*. Available: https://www.dcc.ac.uk/tools/data-asset-framework [2020, April 20].

Digital Curation Centre. 2018. *Data management plans*. Available: http://www.dcc.ac.uk/resources/data-management-plans [2020, June 18].

Donnelly, M. 2012. Data management plans and planning. In G. Pryor, Ed. *Managing research data*, London: Facet publishing. 83-104.

Ekmekcioglu, Ç. & Rice, R. 2009. *Edinburgh data audit implementation project: Final report*. Available:

https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.175.6624&rep=rep1&type=pdf [2020, October 18].

Erway, R. 2013. *Starting the conversation: University-wide research data management policy.* Dublin, Ohio: OCLC.

Flick, U. 2009. *An introduction to qualitative research.* 4th ed. London: Sage.

Flick, U. 2011. *Introductalion to research methodology: a beginner's guide to doing a research project*. London: Sage.

Flores, J.R., Brodeur, J.J., Daniels, M.G., Nicholls, N. & Turnator, E. 2015. Libraries and the research data management landscape. In *The process of discovery: The CLIR postdoctoral fellowship program and the future of the academy.* J. C. Maclachlan, E. A., Waraksa, & C. Williford, Eds. Washington: Council on Library and Information Resources. 82-102.

Friedlander, A. & Adler, P. 2006. *To stand the test of time: long-term stewardship of digital data sets in Science and Engineering. A report to the National Science Foundation from the ARL Workshop on new collaborative relationships.* Allington, VA: Association of Research Libraries.

Gelo, O., Braakmann, D. & Benetka, G. 2008. Quantitative and qualitative research: beyond the debate. *Integrative Psychological and Behavioral Science,* 42(3):266-90. DOI: 10.1007/s12124-008-9078-3.

Gill, P., Stewart, K., Treasure, E. & Chadwick, B. 2008. Methods of data collection in qualitative research: interviews and focus groups. *British Dental Journal*, 204(6):291-295. DOI:10.1038/bdj.2008.192.

Halbert, M. 2013. *Research data management: principles, practices, and prospects*. Washington, DC: Council on Library and Information Resources.

Harrison, R. & Reilly, T. 2011. Mixed methods designs in marketing research. *Qualitative Market Research: An International Journal,* 14(1):7-26. DOI:10.1108/13522751111099300.

Henderson, M.E. & Knott, T.L. 2015. Starting a research data management program based in a university library. *Medical Reference Services Quarterly,* 34(1):7-59. DOI:10.1080/02763869.2015.986783.

Hesse-Biber, S. & Leavy, P. 2011. *The practice of qualitative research*. 2nd ed. Los Angeles: Sage.

Higgins, S. 2008. The DCC Curation Lifecyle Model. *The International Journal of Digital Curation*, 3(1):134-140. DOI:10.2218/ijdc.v3i1.48.

Hopkins, W.G. 2000. *Quantitative Research Design – Sportscience*. Available: http://www.sportsci.org/jour/0001/wghdesign.html [2020, November 13].

Iarossi, G. 2006. *The power of survey design: A user guide for managing surveys, interpreting results, and influencing respondents*. Washington, D.C.: World Bank.

Ingram, C. 2016. *How and why you should manage your research data: a guide for researchers.* Available: https://www.jisc.ac.uk/guides/how-and-why-you-should-manage-your-research-data [2021, January 21].

Institute of Medicine, National Academy of Engineering, National Academy of Sciences, Committee on Science, Engineering, and Public Policy, Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age. 2009. *Ensuring the integrity, accessibility, and stewardship of research data in the digital age.* Washington: National Academic Press.

Jisc. 2015. Meeting the requirements of the EPSRC research data policy. [Online]. Available: https://www.jisc.ac.uk/guides/meeting-the-requirements-of-the-EPSRC-research-datapolicy [2020, September 8]

Jisc. 2018. *How and why you should manage your research data: a guide for researchers*. Available: https://www.jisc.ac.uk/guides/how-and-why-you-should-manage-your-research-data [2021, January 21].

Jones, S. 2011. *Research data policy briefing.* Available: http://www.dcc.ac.uk/sites/default/files/documents/resource/policy/DCC_policy_briefing_2011.pdf [2020, January 29].

Jones, S., Pryor, G. & Whyte, A. 2013. *How to develop research data management services: A guide for HEIs.* Available: http://www.dcc.ac.uk/resources/how-guides [2020, August 7].

Jones, T., Baxter, M. & Khanduja, V. 2013. A quick guide to survey research. *Annals of the Royal College of Surgeons of England*. 95. 5-7. DOI:10.1308/003588413X13511609956372.

Kahn, M., Higgs, R., Davidson, J. & Jones, S. 2014. Research data management in South Africa: how we shape up. *Australian Academic & Research Libraries*, 45(4):296-308. DOI:10.1080/00048623.2014.951910.

Kennan, M.A. & Markauskaite, L. 2015. Research data management practices: a snapshot in time. *International Journal of Digital Curation*, 10(2):69-95. DOI:10.2218/ijdc.v10i2.329.

Kombo, D.K. & Tromp, D.L.A. 2006. *Proposal and thesis writing: An introduction.* Nairobi: Paulines.

Kumar, R. 2011. *Research methodology: A step-by-step guide for beginners*. 3rd ed. London: Sage.

Lee, C.A. & Tibbo, H.R. 2007. Digital curation and trusted repositories: Steps toward success. *Journal of Digital Information*, 8(2). Available: https://journals.tdl.org/jodi/index.php/jodi/article/view/229 [2020, September 11].

Leedy, P.D. & Ormrod, J.E. 2013. *Practical research: Planning and design*. 10th ed. Upper Saddle River, N. J.: Pearson Education.

Lewis, M.J. 2010. Libraries and the management of research data. In *Envisioning future academic library services.* S. McKnight, Ed. London: Facet.145-168.

Macmillan, D. 2014. Data sharing and discovery: what librarians need to know. *The Journal of Academic Librarianship*, 40(5):541-549. DOI:10.1016/j.acalib.2014.06.011.

Macwhinnie, L.A. 2003. The information commons: the academic library of the future. *Portal: Libraries and the Academy*, 3(2):241-257. DOI:10.1353/pla.2003.0040.

Massey, D. & Tourangeau, R. 2013. Where do we go from here? nonresponse and social measurement. *The ANNALS of the American Academy of Political and Social Science*, 645(1):222-236. DOI:10.1177/0002716212464191.

National Research Foundation. 2015. *Statement on Open Access to research publications from the National Research Foundation (NRF)-Funded Research.* Available: https://www.nrf.ac.za/media-room/news/statement-open-access-research-publications-national-research-foundation-nrf-funded [2020, October 18].

Neevers, A .2009. The battle against digital obsolescence: exploring strategies of digital Preservation in New Media and New Media Art. Available:

https://annewilneervens.files.wordpress.com/2009/03/new-media-theories-final-paper.pdf [2019, March 21].

Ngulube, P. 2012. Mapping mixed methods research in Library and Information Science Journals in Sub-Saharan Africa, 2004 - 2008. *African Journal of Library, Archives and Information Science*, 22 (2):117-132

Nueman, L.W. 2012. *Social research methods: Qualitative and quantitative approaches.* 7th ed. Boston: Pearson.

Odiya, J.N. 2009. *Scholarly writing: research proposals and reports in APA or MLA publication style*. Kampala: Makerere University Printery.

Ogier, A., Hall, M., Bailey, A. & Stovall, C. 2014. Data management inside the library: assessing electronic resources data using the data asset framework methodology. *Journal of Electronic Resources Librarianship,* 26(2):101-113. DOI:10.1080/1941126X.2014.910406.

Oxford University. 2018. *About RDM*. Available: http://researchdata.ox.ac.uk/home/introduction-to-rdm/ [2019, November 25].

Pampel, H., Vierkant, P., Scholze, F., Bertelmann, R., Kindling, M., Klump, J., Goebelbecker, H., Gundlach, J., Schirmbacher, P. & Dierolf, U. 2013. Making research data repositories visible: the re3data.org Registry. *PLoS ONE*, 8(11):1-10. DOI:10.1371/journal.pone.0078080.

Patterton, L., Bothma, T. & Van Deventer, M. 2018. From planning to practice: an action plan for the implementation of research data management services in resource-constrained institutions. *South African Journal of Libraries and Information Science*, 84:2:14-26.

Pennock, M. 2007. Digital curation: a life-cycle approach to managing and preserving usable digital information. *Library & Archives*, 1(1):1-3.

Perrier, L., Blondal, E., Ayala, A.P., Dearborn, D., Kenny, T., Lightfoot, D., Reka, R., Thuna, M., Trimble, L. & MacDonald, H. 2017. Research data management in academic institutions: a scoping review. *PloS One*, 12(5):1-14. DOI:10.1371/journal.pone.0178261.

Pickard, A.J. 2013. *Research methods in information*. London: Facet Publishing.

Pinfield, S., Cox, A.M. & Smith, J. 2014. Research data management and libraries: relationships, activities, drivers and influences. *PloS One*, 9(12):1-28. DOI:10.1371/journal.pone.0114734.

Pryor, G. 2013. Options and approaches to RDM service provision. In *Delivering research data management services: Fundamentals of good practice*. G. Pryor, S. Jones & A. Whyte, Eds. London: Facet.

Punch, K.F. & Oancea, A. 2014. *Introduction to research methods in education*. London: Sage.

Raboin, R., Reznik-Zellen, R.C. & Salo, D. 2013. Forging new service paths: Institutional approaches to providing research data management services. *Journal of eScience Librarianship*, 1(3):134-147. DOI:10.7191/jeslib.2012.1021.

Rans, J. & Jones, S. 2013. *RDM strategy: Moving from plans to action*. Available: https://www.dcc.ac.uk/resources/developing-rdm-services/rdm-strategy-moving-plans-action [2019, June 3].

Raosoft. 2004. Sample size calculator. Available: http://www.raosoft.com/samplesize.html [2020, September 8]

Read, E.J. 2007. Data services in academic libraries: assessing needs and promoting services. *Reference & User Services Quarterly*, 46(3): 61-75.

Research Data Strategy Working Group. 2011. *The 2011 Canadian Research Data Summit*. Ottawa Convention Centre. September 14th - 15th. Available: https://www.rdc-drc.ca/wp-content/uploads/Background-to-the-2011-Canadian-Research-Data-Summit1.pdf [2021, January 21].

Romm, N. & Ngulube, P., 2015. Mixed methods research. In *Addressing research challenges: Making headway for emerging researchers.* E.R. Mathipa & M.T. Gumbo, Eds. Noordywk: Mosala-MASEDI. 157-175. Available: https://www.researchgate.net/publication/305728941_Mixed_methods_research [2019, June 17].

Sarantakos, S. 2005. *Social Research*. 3rd ed. New York: Palgrave Macmillan.

Shen, Y. & Varvel, V.E. 2013. Developing data management services at the Johns Hopkins University. *The Journal of Academic Librarianship*, 39(6):552-557. DOI:10.1016/j.acalib.2013.06.002.

Si, L., Xing, W., Zhuang, X., Hua, X. & Zhou, L. 2015. Investigation and analysis of research data services in university libraries. *The Electronic Library*, 33(3):417-449. DOI:10.1108/EL-07-2013-0130.

Smith, J. & Noble, H. 2014. Bias in research. *Evidence Based Nursing*, 17(4):100-101. DOI:0.1136/eb-2014-101946.

Soehner, C., Steeves, C. & Ward, J. 2010. *E-Science and data support services: A study of ARL Member Institutions*. Washington: Association of Research Libraries. Available: https://files.eric.ed.gov/fulltext/ED528643.pdf [2020, May 10].

Strasser, C., Cook, R., Michener, W. & Budden, A. 2012. *Primer on data management: What you always wanted to know*. Available: https://escholarship.org/content/qt7tf5q7n3/qt7tf5q7n3.pdf [2019, October 3].

Struwig, M. & Stead, G.B. 2001. *Planning, reporting & designing research*. Cape Town: Pearson South Africa.

Tansley, S., & Tolle, K. M. 2009. The fourth paradigm: data-intensive scientific discovery (Vol. 1). A. J. Hey, Ed. Redmond, WA: Microsoft research.

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M. & Frame, M. 2011. Data sharing by scientists: practices and perceptions. *PloS One*, 6(6):1-21. DOI:10.1371/journal.pone.0021101.

Tenopir, C., Allard, S., Frame, M., Birch, B., Baird, L., Sandusky, R., Langseth, M., Hughes, D. & Lundeen, A. 2015. Research data services in academic libraries: data intensive roles for the future? *Journal of eScience Librarianship*, 4(2). DOI:10.7191/jeslib.2015.108.

Tenopir, C., Birch, B. & Allard, S. 2012. *Academic libraries and research data services: Current practices and plans for the future*. Chicago, IL: Association of College and Research Libraries.

Tenopir, C., Dalton, E.D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D. & Dorsett, K. 2015. Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PloS One*, 10(8):1-24. DOI:10.1371/journal.pone.0134826.

Tenopir, C., Sandusky, R. J., Allard, S. & Birch, B. 2013. Academic librarians and research data services: preparation and attitudes. *IFLA Journal*, 39(1)70-78. DOI:10.1177/0340035212473089.

Tenopir, C., Sandusky, R. J., Allard, S. & Birch, B. 2014. Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*, 36(32):84-90. DOI:10.1016/j.lisr.2013.11.003.

Tenopir, C., Talja, S., Horstmann, W., Late, E., Hughes, D., Pollock, D., Schmidt, B., Baird, L., Sandusky, R. & Allard, S. 2017. Research data services in European Academic Research Libraries. *Liber Quarterly*, 27(1):23-44. DOI:10.18352/lq.10180/.

Trimble, L. & Macdonald, H. 2017. Research data management in academic institutions: A scoping review. *PLoS One*, 12(5):1-14. DOI:10.1371/journal.pone.0178261.

UCL. 2018. *Sharing data.* Available: http://www.ucl.ac.uk/library/research-support/research-data/best-practices/guides/sharing [2020, May 18].

University of Bristol.2012.*Why manage research data*. Available: http://data.bris.ac.uk/bootcamp/management/ [2020, October 12]

University of Cambridge. 2018. *Funders' policies: research data management*. Available: https://www.data.cam.ac.uk/funders [2019, June 18].

University of KwaZulu-Natal. 2017. *History – University of KwaZulu-Natal*. Available: https://www.ukzn.ac.za/about-ukzn/history [2019, June 20].

University of KwaZulu-Natal. 2017. UKZNDABA Online. Available: https://ndabaonline.ukzn.ac.za/UkzndabaStory/Vol5-Issue38/UKZN%20Ranked%20Third%20in%20SA%20by%20International%20Rankings%20Organisation/ [2021, August 17].

University of KwaZulu-Natal. 2017. *Research report.* Available: http://research.ukzn.ac.za/UKZNResearchReports.aspx [2020, September 20]

University of KwaZulu-Natal. 2018. *About UKZN postdoctoral fellows*. Available: http://research.ukzn.ac.za/postdoc-at-a-glance/postdocataglancepdrf.aspx [2019, March 16].

University of KwaZulu-Natal. 2019. *History*. Available: https://ukzn.ac.za/about-ukzn/history/ [2020, September 19]

University of KwaZulu-Natal. 2021. *Organisational structure*. Available: https://ukzn.ac.za/about-ukzn/organisational-structure/ [2020, September 18]

University of Stellenbosch Library. 2017. *Research data management (RDM).* Available: http://library.sun.ac.za/en-za/Research/Pages/research-data-management.aspx [2020, October 19].

Uzwyshyn, R. 2016. Online research data repositories: The what, when, why and how. *Computers in Libraries*, 36(3):18-21.

Van Den Eynden, V., Corti, L., Woollard, M., Bishop, L. & Horton, L. 2011. *Managing and sharing data: A best practice guide for researchers*. Essex: UK Data Archive.

Van Deventer, M. & Pienaar, H. 2015. Research data management in a developing country: a personal journey. *International Journal of Digital Curation*, 10(2):33-47. DOI:10.2218/ijdc.v10i2.380.

Van Wyk, J. 2016. *RDM and the role of the information professional*. [PowerPoint presentation, October 25]. Available: http://hdl.handle.net/2263/58892 [2019, February 15].

Van Zeeland, H. & Ringersma, J. 2017. The development of a research data policy at Wageningen University & Research: best practices as a framework. *Liber Quarterly*, 27(1):153-170. DOI:10.18352/lq.10215/.

Verbaan, E. & Cox, A.M. 2014. Occupational sub-cultures, jurisdictional struggle and third Space: theorising professional service responses to research data management. *The Journal of Academic Librarianship*, 40 (3-4):211-219. DOI:10.1016/j.acalib.2014.02.008.

Wallace, D. & Van Fleet, C. 2012. *Knowledge into action: research and evaluation in library and information science*. Santa Barbra, California: Libraries Unlimited.

Welman, C., Kruger, F. & Mitchell, B. 2005. *Research methodology*. 3rd ed. Cape Town, South Africa: Oxford University Press.

Whyte, A. & Tedds, J. 2011. *Making the case for research data management. Digital Curation Centre Jisc Briefing Paper*. Available: https://www.dcc.ac.uk/guidance/briefing-papers/making-case-rdm [2020, April 23].

Wilson, J.A., Martinez-Uribe, L., Fraser, M.A. & Jeffreys, P. 2011. An institutional approach to developing research data management infrastructure. *International Journal of Digital Curation*, 6(2):274-287. DOI:10.2218/ijdc.v6i2.203.

Appendices

Appendix A: Ethical clearance

Department of Knowledge and Information Stewardship
Faculty of Humanities
University of Cape Town
Upper Campus

Private Bag X1, RONDEBOSCH, 7701 South Africa
Level 6 Hlanganani, The Chancellor Oppenheimer Library
Tel: +27 (0) 21 650 4546
E-mail: dkis@uct.ac.za
Internet: www.dkis.uct.ac.za

UCTLIS201906-04 15 October 2019

Zizipho Madibi

Department of Knowledge and Information Stewardship

University of Cape Town

Dear Zizipho

I am pleased to inform you that ethical clearance has been granted by the Ethics Review Committee of the Department of Knowledge and Information Stewardship on behalf of the Humanities Faculty of the University of Cape Town for your Master's study entitled: *The role of academic libraries in implementing research data services: a case study of the University of KwaZulu-Natal Libraries.*

I wish you the very best with your study.

Yours sincerely,

Andiswa Mfengu

Appendix B: Consent Form

UNIVERSITY OF CAPE TOWN

FACULTY OF HUMANITIES- DEPARTMENT OF KNOWLEDGE & INFORMATION STEWARDSHIP

Informed Consent Form

Title of Research Study	The role of academic libraries in implementing research data services: a case study of the University of KwaZulu-Natal Libraries
Researcher	Zizipho Madibi MDBZIZ001@myuct.ac.za 0786653586 Master of Philosophy specialising in digital curation, Department of Knowledge and Information Stewardship, UCT
Supervisor	Michelle Kahn, Lecturer, Department of Knowledge and Information Stewardship, UCT Michelle.Kahn@uct.ac.za 021 650 1851
Ethics office: UCT	Andiswa Mfengu 021 650 3092
Ethics office: UKZN	Mariette Snyman 0312608350
Nature of the Research	In order to determine the role UKZN Libraries can play in the implementation of RDS at UKZN, the study seeks: • To identify the need for research data services among UKZN researchers.

- To identify the major challenges associated with research data services at UKZN.
- To determine the possibility of implementing research data services at UKZN libraries.

Library staff will be individually interviewed for about 45 minutes each and researchers will be surveyed by means of questionnaires. These are believed to be the major stakeholders in the development and implementation of research data services in universities. The information provided by the stakeholders will be beneficial to this study. The study motivates further exploration of RDM at UKZN and contributes to the body of knowledge on the implementation of research data services in libraries. Furthermore, the study will provide a valuable case study and contribute to the body of knowledge on research data services for other libraries who might need to implement them.

Please note that there are no risks involved and no remuneration attached to participation in this study. Your personal information will be anonymised and any information you provide will be used for the purposes of this research study only. Participants have a right to withdraw from the study at any point.

DECLARATION

I……. (full name of participant)

hereby confirm that (please tick where appropriate):

I understand the contents of this document and the nature of the research project, and I consent to participating in the research project ☐

I do not give/give (strike through as appropriate) consent to the recording of this interview ☐

I understand that this research might be published in a research journal ☐

I understand that the document (dissertation) of this research will be made available on an Open Access Institutional Repository ☐

I understand that I am at liberty to withdraw from the project at any time, should I so desire ☐

I understand the intention of the research ☐

I hereby agree to participate ☐

SIGNATURE OF PARTICIPANT DATE

…… ……

SIGNATURE OF RESEARCHER DATE

…… ……

Appendix C: Questionnaire

UNIVERSITY OF CAPE TOWN

FACULTY OF HUMANITIES- DEPARTMENT OF KNOWLEDGE & INFORMATION STEWARDSHIP

Title of Research Study The role of academic libraries in implementing research data services: a case study of the University of KwaZulu-Natal Libraries

Questionnaire for Researchers

1. Please indicate your discipline
    - College of Humanities
    - College of Law and Management Studies
    - School of Agriculture, Engineering and Science
    - School of Health Sciences

2. Which school do you belong to?
    - School of Applied Human Sciences
    - School of Arts
    - School of Built Environment and Development Studies
    - School of Education
    - School of Religion, Philosophy and Classics
    - School of Social Sciences
    - School of Engineering
    - School of Agricultural, Earth and Environmental Sciences

- School of Chemistry and Physics
- School of Life Sciences
- School Mathematics, Statistics and Computer Science
- School of Clinical Medicine
- School of Laboratory Medicine & Medical Sciences
- School of Health Sciences
- School of Nursing & Public Health
- Graduate School of Business and Leadership
- School of Accounting, Economics and Finance
- School of Law
- School of Management, IT and Governance

3. What is your highest level of education?
    - Professor
    - Doctorate
    - Master's Degree
    - Honours Degree
    - Bachelor's Degree
    - Other
4. What types of research data do you create or work with?
    - Spreadsheets
    - Documents (PDF, Texts, Microsoft Word)
    - Audios
    - Videos
    - Images
    - Slides, samples, artifacts, specimens
    - Geospatial data
    - Other
5. Do your chosen formats and software enable sharing and long-term access to the data?
    - Yes
    - No
    - Not sure
    - Don't know

6. Have you ever developed a data management plan (DMP) for your research?
    - Yes
    - No
    - Don't know
7. If yes, did you experience any challenges when developing your DPM?
    - Yes (please specify)
    - No
    - Maybe

8. Do you know of any requirements regarding DMPs from your funders?
    - Yes, there are requirements
    - No, there are no requirements
    - Not sure, there might be requirements

9. Where is your research stored?
    - On the cloud
    - Your computer
    - Shared drive
    - USB
    - Server on campus
    - Data repository
    - Email
    - On paper
    - Other *(please specify)*

10. Do you have a plan on how your data can be recovered in the event of an incident?
    - Yes, *(please specify)*
    - No
    - Not sure

11. How often do you back up your data?
    - Hourly
    - Daily
    - Weekly
    - Monthly
    - Fortnightly
    - Annually
    - Ad-hoc
    - Don't know
    - Never

12. If data is backed up, where is it backed up?
    (please select all that apply)
    - External hard drive
    - USB
    - On cloud
    - Server on campus
    - Other – (*please specify)*
    - Not applicable – data not backed up

13. Do you assign metadata for your data (Metadata is data about research data, making it easier to search for)
    - Yes
    - No
    - Sometimes
    - Don't know

14. If metadata is assigned, do you follow any policy/guidelines?
    - Yes (*please specify*)

- No
- Sometimes (*please specify*)
- Don't know
- Not applicable, data not assigned

15. Data sharing includes sharing data with fellow researchers informally and more formally in data repositories, data centers. Whom do you share your data with? Who can access the data you are creating?
    (please select all that apply)
    - Fellow researchers
    - Others in the research unit
    - Colleagues on campus
    - Funders
    - General public
    - Study supervisor
    - No one
    - Other *(please specify)*
16. How often have you been able to provide access to your data?
    - Always
    - Half the time
    - Never
17. Which methods have you used to share your data? *(please select all that apply)*
    - Email
    - A curated digital data repository
    - CD/DVD
    - USB
    - File Transfer Protocol (FTP)
    - A web portal for download access
    - Other *(please specify)*

18. If using a curated digital data repository please provide details on its location, curator, users and usage

    Text box

19. What kind of challenges do you experience in making use of the repository? *(if any)*

    Text box

20. Have you ever received any research data management training?
    - Yes
    - No
    - Can't remember
21. If answer to the previous question is yes, please share your experience

    Text box

22. Which areas would you like to receive training in?
    - Developing a data management plan

- Data formatting
- Data documentation
- Data storing
- Data sharing
- Metadata creation
- Funders' requirements
- Data repositories
- Open access
- Not interested in training
- Other *(please specify)*

23. Please rate the importance of the following data-related services to assist in improving research in your discipline
    - Having increased access to data in my discipline
    - Having access to data in other disciplines in my institution and other institutions
    - Guidelines and services supporting researchers in managing their research data
    - Services supporting researchers depositing their research data in digital data repositories
    - Services supporting researchers in creating metadata
    - Services supporting researchers use correct data naming conventions
    - Having teaching and learning materials supporting researchers to work with data
    - Having the services necessary to assign a digital object identifier (DOI) to my data so other researchers can find and cite my data
    - Having proper infrastructure at UKZN that allows preservation of researchers' data, and made accessible to others

All services listed were rated as either "very important", somewhat important", not important or unfamiliar with this service

24. Are there any of your issues related to the implementation of Research Data Services you would like to discuss? Please comment below

Long answer text

Thank you for this invaluable information. I am grateful that you have made yourself available for this survey. Your contribution will be of great help in ensuring that this research is successful.

Name of researcher: Zizipho Madibi
Email: mdbziz001@myuct.ac.za

Name of supervisor: Michelle Kahn
Email: michelle.kahn@uct.ac.za

Appendix D: Interview schedule

UNIVERSITY OF CAPE TOWN FACULTY OF HUMANITIES- DEPARTMENT OF KNOWLEDGE & INFORMATION STEWARDSHIP

Title of Research Study: The role of academic libraries in implementing research data services: a case study of the University of KwaZuluNatal Libraries

Interview Schedule for Librarians

1. What is your designation in the library?
2. What is your highest level of education?
3. Can you please explain what your role is in the library?
4. Please describe how the library supports the research process at UKZN
5. What do you understand about Research Data Management (RDM)?
   (The researcher will provide an explanation of what RDM entails if the participant is not familiar with the concept, this is necessary for the purposes of guiding the interview)
6. 4. Does the library have any plans of implementing research data services in the future?
7. 5. Does the library have any resource allocation (including funding) to facilitate RDS implementation in the future?
8. In your opinion, is there a need for the library to support researchers manage and share their data?

9. 7. Which research funders require a data management plan (DMP)? - Is there any training/support provided by the library to assist researchers develop their DMPs?
10. To what extent do you agree with the growing need to share data? Is the library considering providing support for researchers to share their data?
11. Would you consider the current library staff knowledgeable and suitable to support researchers in RDS? *(If no, ask a follow up question)*
12. What challenges do you anticipate in implementing RDS at UKZN?
13. Have you ever supported a researcher in the process of managing their research data? *If yes, please elaborate.*
14. Could you share with me which research data services could UKZN Libraries provide and to what extent would you support the implementation?

If some of the research data services are not mentioned, the services below will be mentioned as prompts.

Finding, organising, managing and evaluating scholarly data; Data repositories and Open Access; Data management plan writing; Research data management training; Metadata creation; Data formatting; Data storage; Data sharing; Data archiving.

Any other issues related to Research Data management

15. Are there any of your own issues related to the implementation of Research Data Services you would like to discuss?

Closing Remarks

Thank you for this invaluable information. I am grateful that you have made yourself available for this interview. Your contribution will be of great help in ensuring that this research is successful. I request your permission to contact you again should a need for further questions and clarification arises.

Appendix E: DMP Tool

DMP title

Project Name My plan (DKIS DMP)
Principal Investigator / Researcher Zizipho Madibi
Institution University of Cape Town

Dissertation Title

The role of the library in implementing research data services: a case study of the University of KwaZulu-Natal Libraries.

Other DMP requirements

2.1 Have you already written a DMP?

- No

Data collection/generation

3.1 Data sources

- I am generating primary research data

3.2 Where will you obtain permission to reuse the secondary data and/or collect the primary data for your research?
Permission will be obtained from Department (DKIS) Research Ethics Committee.

3.3 What data will you be collecting, from where, and in which formats? How much data do you expect to generate/collect?
The researcher will be collecting quantitative and qualitative data from UKZN and will be in text format.

Storage

4.1 Where will this data be stored during the research? What security measures will be in place? What are the plans for back-up?
Research data will be stored on external hard drive.
Data will also be stored on Google Drive for backup purposes.

4.2. How will the data be documented and described?
The following naming convention will be used:
<Student number>_<document type, if needed>_<document title>_<year>

Ethical considerations

5.1 What are the main ethical considerations around your data?
Research will involve humans and no harm or risks forseen for the study participants.

Data retention, preservation and sharing

6.1 Once your research is complete, which data (if any) should be retained?
Data will be retained according to discipline-specific or institutional policy.

6.2. For what period of time should data be retained? 5 years.

6.3 Where will this data be stored in the relative long-term (while it is being retained)?
Data will be submitted to an open access data repository.

6.4 Under what conditions could this data be shared with others?
Data will be available entirely on open access. There will be no restrictions, as permission will be received from participants.

Implementation of data management plan

7.1 Are there any resources or additional guidance needed to implement the above plan? No.

Appendix F: UKZN Permission Letter

UNIVERSITY OF KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

7 December 2019

Ms Zizipho Madibi
University of Mpumalanga

Email: zmadibi@gmail.com

Dear Ms Madibi

RE: PERMISSION TO CONDUCT RESEARCH

Gatekeeper's permission is hereby granted for you to conduct research at the University of KwaZulu-Natal (UKZN) towards your postgraduate studies, provided Ethical clearance has been obtained. We note the title of your research project is:

"The Role of Academic Libraries in Implementing Research Data Services: A Case Study of the University of KwaZulu-Natal Libraries".

It is noted that you will be constituting your sample by handing out questionnaires and/or conducting interviews with staff members in the UKZN Libraries.

Please ensure that the following appears on your notice/questionnaire: • Ethical clearance number;
• Research title and details of the research, the researcher and the supervisor; • Consent form is attached to the notice/questionnaire and to be signed by user before he/she fills in questionnaire; • gatekeepers approval by the Registrar.

You are not authorized to contact staff and students using 'Microsoft Outlook' address book. Identity numbers and email addresses of individuals are not a matter of public record and are protected according to Section 14 of the South African Constitution, as well as the Protection of Public Information Act. For the release of such information over to yourself for research purposes, the University of KwaZulu-Natal will need express consent from the relevant data subjects. Data collected must be treated with due confidentiality and anonymity.

Yours sincerely

DR KE CLELAND
REGISTRAR (ACTING)

Office of the Registrar
Postal Address: Private Bag , Durban, South Africa
Telephone: +27 (0) 31 260 8005/2206 Facsimile: +27 (0) 31 260 7824/2204 Email: reaistrar@ukzn.ac.za
Website:
www.ukzn.ac.za a • 2010
100 YEARS OF EXCELLENCE
Founditg Campuses Edgewood Howard College Medical School Pietermaritzburg