University of Zululand

RESEARCH DATA MANAGEMENT SERVICES IN SELECTED PUBLIC UNIVERSITIES IN KWAZULU-NATAL, SOUTH AFRICA

By

Mpilo Siphamandla Mthembu

(201256633)

(BLIS, AMAS)

A dissertation submitted in fulfilment of the requirements for the Doctor of Philosophy in Library and Information Science

in the

Department of Information Studies

Faculty of Humanities and Social Sciences

at the

University of Zululand, KwaDlangezwa

Supervisor: Prof. D.N. Ocholla

2023

DECLARATION

I hereby declare that this is my original research work for the degree of Doctor of Philosophy in Library and Information Science at the University of Zululand's Department of Information Studies. The dissertation has not previously been submitted for any degree or examination at this or any other university. This document's information was obtained and presented in accordance with academic guidelines as well as ethical procedures and practices. In addition, I have fully cited/referenced all materials and results in this work that are not entirely my own.

Student: Mpilo S. Mthembu

Supervisor: Prof. Dennis N. Ocholla

2023

DEDICATION

This work is dedicated to my family, supervisor (Professor D.N. Ocholla), and friends. I am grateful for their encouraging words, which will always ring in my ears. They were there for me every step of the way, and I will be eternally grateful for their contributions throughout the entire research project. I also thank the Librarians, Library technicians/IT specialists, Research Directors, Deans/Deputy Deans Research, Deputy Vice-chancellors Research, Heads of Departments, and academic researchers (NRF rated) who took part in this research and contributed to its success.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Dennis N. Ocholla, for his tireless supervision, guidance, and support throughout the entire research project for ensuring the success of this piece of work. I am thankful to my family. I would also like to thank my fiancée Nomfundo Ndlovu for her unending support and words of encouragement throughout the duration of this research project. I am grateful to my friends, particularly Lungelo S Mbatha and Bethiweli M Mlambo, as well as colleagues from the Department of Information Studies, for their inspiration and support. I am grateful for the unending support and contributions I have received since the day I chose the topic for this study. Your words of encouragement will be remembered forever. I would also like to thank the study participants who made significant contributions to the study's completion. Finally, I would like to thank the National Research Foundation (NRF) for its financial support of this research project. May God abundantly bless you all.

ABSTRACT

Research data is thought to be an essential component that assists researchers in forming the foundation of research output. To increase the amount of reproducibility in research data, the data should be accompanied by extensive documentation, making it more accessible and secure. As a result of the current digital world and the resulting explosion of multidisciplinary research data, Research Data Management (RDM) comes into play. The current study aimed to investigate research data management services in selected public universities in South Africa, particularly those based in KwaZulu-Natal province. This study was underpinned by the issue of varying cultures between disciplines or departments, as well as a lack of skills and infrastructure, as major challenges impeding the provision of RDM services in most higher education institutions. In terms of theoretical foundation, this study used the Community Capability Model Framework (CCMF) in conjunction with the Digital Curation Centre (DCC) Lifecycle Model.

The study used both qualitative and quantitative methodologies and followed the post-positivist paradigm. As research methods, both qualitative and quantitative case studies were used. Semi-structured interviews were conducted online via Teams and Microsoft Teams with 23 study participants (Librarians, Library Technicians, Research Directors, Deans / Deputy Deans Research, Deputy Vice-Chancellors Research, and Heads of Departments) to collect qualitative data during the COVID-19 pandemic. To collect quantitative data, questionnaires were converted into Google forms and emailed to 30 NRF-rated researchers.

The findings of this study revealed that participating institutions (University of Zululand- UZ, and Durban University of Technology- DUT) are still in the early stages of providing RDM services, as (UNIZULU) has a planning department that has begun RDM in the institution, though it is not widely known, and (DUT) RDM service provision is in the pipeline, as its RDM policy is still at the draft stage. However, as regards research activities, both institutions support data creation, storage, preservation, sharing, access, use, and re-use. Both institutions also support grant applications and collaborative research as research-related activities and data services, as well as RDM and data citation. DSpace and DUT Open

Scholar are institutional repositories (IRs) that are used to permanently store, manage, and archive research data such as journal articles, conference proceedings, theses, and dissertations. Unfortunately, most researchers are unaware of how long their institutions keep research data. The findings also reveal that the majority of researchers store and backup their research data on personal computers, emails, and external storage devices like USBs, external hard drives, and CDs. The RDM service is important for new researchers because it promotes research visibility, data sharing, data storage and availability for researchers, data use and re-use. Research supervisors have the potential to assist emerging researchers in locating and publishing in reputable journals. Furthermore, data sharing contributes significantly to increasing the impact and visibility of research, promoting scientific inquiry and debate, and fostering new collaborations between data users and data creators.

Knowledge and skillsin data management, research methodology, data curation, metadata skills, and technical skills were overemphasised as truly necessary RDM competency requirements for both staff and researchers. Fortunately, most researchers understand the importance of ethics such as citation, data sharing, data analysis, searching and retrieval, and data capture. They are, however, lacking in data management and organisation, database development, data curation and re-use, and software skills. Short RDM courses, workshops, and professional body training are all practices that could be used in institutions to develop RDM skills. Furthermore, because the vast majority of study participants had not received RDM training at their institutions, data archiving and curation training are required.

The findings of the study further reveal that UNIZULU currently does not have an RDM policy that facilitates research data management, whereas DUT has a research data management policy that is still in the draft stage. In terms of infrastructure or resources, the findings revealed that UNIZULU lacks resources and well-equipped staff to provide RDM services, whereas the DUT library has adequate resources; however, whether additional resources are to support RDM practices in their institution will be determined by the actual implementation of RDM. However, teaching and research findings at DUT revealed a lack of adequate infrastructure as well as a lack of human resources or capital. There are currently no specific

capacity-building programs or strategies for developing RDM skills in the institutions studied. A lack of resources or infrastructure, skilled personnel, getting researchers on board, and adoption, which is about getting people to adopt the institution's culture of managing research data, were cited as major challenges in providing the RDM service. RDM challenges may be addressed through training, workshops, and institutional policies and procedures. Finally, as viable options for developing an RDM framework for South African public universities, the study suggests RDM policy formulation, a collaboration between universities and relevant stakeholders, and a responsible group or body such as NRF and StatsSA.

This study's originality and novelty stems from its scope, subject matter, and application. The study's findings can be widely used in research, teaching, and learning. Furthermore, the study can possibly inform RDM strategy and policy in South Africa and even elsewhere.

Key words: *Research data, research data management, data preservation, data curation, data archiving, institutional repository, higher education institution, researcher*, *Community Capability Model Framework, DCC Curation Lifecycle Model.*

TABLE OF CONTENTS

Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

ANDS	Australian National Data Service
ANU	Australian National University
ARC	Applied Research Centre
CARDIO	Collaborative Assessment of Research Data Infrastructure and Objectives
CCMF	Community Capability Model Framework
CGIAR	Consultative Group on International Agricultural Research
CHE	Council of Higher Education
CLIR	Council on Library and Information Resources
CPD	Continuous professional development
CPUT	Cape Peninsula University of Technology
CSIR	Council for Scientific and Industrial Research
CSIRO	Commonwealth Scientific and Industrial Research Organization
DAF	Data Audit Framework
DCC	Digital Curation Centre
DD	Deputy Dean
DHET	Department of Higher Education and Training
DIRISA	Data Intensive Research Initiative of South Africa
DMF	Data Management Framework
DMP	Data Management Plan
DRAMBORA	Digital Repository Audit Method Based on Risk Assessment
DST	Delta Sigma Theta
DUT	Durban University of Technology
DVC	Deputy Vice-Chancellor
EIAR	Ethiopian Institute of Agricultural Research
HEI	Higher Education Institution
HOD	Head of the Department
HSRC	Human Sciences Research Council

ICAR	Indian Council of Agricultural Research
IKT	Information Knowledge and Technology
IR	Institutional Repository
IT	Information Technology
JISC	Joint Information Systems Committee
LIASA	Library and Information Association of South Africa
LIS	Library and Information Science
M&E	Monitoring and Evaluation
MANTRA	Man-made Textile Research Association
NCASRD	National Consultation on Access to Scientific Research Data
NRF	National Research Foundation
NSTCC	National Science and Technology Council Committee
OBIS	Ocean Biogeographic Information System
OECD	Organisation for Economic Co-operation and Development
ORDO	Open Research Data Online
ORDS	Online Research Database Service
OU	Open University
PHD	Doctor of Philosophy
RD	Research Data
RD	Research Director
RDM	Research Data Management
RDMP	Research Data Management Plan
RDMS	Research Data Management System
RDRDS	Research Data Registry and Discovery Service
RDS	Research Data Service
SADA	South African Data Archive
SADCO	Southern African Data Centre for Oceanography
SAEON	South African Environmental Observation Network

SAFRING	South African Bird Ringing Unit
SAIAB	South African Institute for Aquatic Biodiversity
SANBI	South African National Biodiversity Institute
SANParks	South African National Parks
SDG	Sustainable Development Goal
SSHRC	Social Sciences and Humanities Research Council
SU	Stellenbosch University
UBC	University of British Columbia
UCT	University of Cape Town
UK	United Kingdom
UNISA	University of South Africa
UNIZULU	University of Zululand
UP	University of Pretoria
USA	United States of America
UWC	University of the Western Cape
WITS	University of the Witwatersrand

CHAPTER ONE: INTRODUCTION AND BACKGROUND

1.1 Introduction

Higher education institutions (HEIs), libraries, and other research centres offer a variety of research data management (RDM) services and activities to their researcher user community. However, providing these services successfully in HEIs requires personnel with the necessary capabilities, as well as appropriate policies, adequate infrastructure, and funding, among other things. In this context, these are critical for managing data and understanding its lifecycle, availability, accessibility, security, use, and re-use in a university. Having such strategies in place at a university also assists the institution to achieve its goals and objectives.

Though the management of data, information and knowledge is central to most institutions and has become an area of keen interest in the current information, knowledge and technology (IKT) driven era, particularly in institutions of higher education and training, it is noted that the provision of RDM services has been widely practised in the developed nations of the world like the United Kingdom, United States, Australia, Canada, to mention but a few (Ng'eno & Mutula; 2018: 28). Unfortunately, it has been reported that this practice is still in its infancy in developing countries (Pasipamire, 2017: 2; Mushi, Pienaar & van Deventer, 2020: 1; Kwanya, 2021: 9), as they are presented with several deficiencies including inadequate infrastructure, lack of resources, financial constraints, insufficient collaboration between stakeholders, a lack of human capabilities and policies, which poses a big threat (Pinfield, Cox and Smith, 2014; Cox & Penfield, 2014; Chiware & Mathe, 2016; Yoon & Schultz, 2017; Cox *et al,* 2017; Ng'eno, 2018; Anduvare, 2019; Avuglah & Underwood, 2019; Cox *et al*, 2019; Chawinga, 2019, Chawinga & Zinn, 2020; Mushi, 2021).

It is in this respect that the present study sought to provide an understanding of the current research data management services, activities, and endorsements of tool adoption as well as shedding light on the areas of support and development in the selected public universities in South Africa. Investigating the current research data

management services in the institutions of higher learning is significant and also relevant for universities that are currently planning on improving their data management services, as well as for those that are already providing such services to their user community of researchers.

1.2 Background to the study

Research data is believed to be an essential component that helps researchers form the basis of research output, and it could consist of survey answers, images, artefacts, observations, and spreadsheets, to mention but a few. Dora and Kumar (2015: 484) point out that research data is raw data produced from labs or surveys, or data that has been processed, cleaned, refined, arranged and combined, such that it becomes useful for research purposes. The authors further explain that research data involves already published data in any scientific communication such as a journal. Chigwada, Chiparausha and Kasiroori (2017: 8), citing Kennan and Markausaite, emphasise that research data is heterogeneous, and hence, presented in many forms, depending on its origin, the problem under investigation, and the discipline of research. For Tripathi, Shukla and Sonker (2017: 417), research data is the "data that is generated when the researchers undertake any research activity". Research data can be qualitative, quantitative, textual, images, verbal context, musical compositions, codes, and so forth (Tripathi, Shukla & Sonker, 2017: 417). Ng'eno and Mutula (2018: 28) declare that "research data is instrumental in supporting innovative solutions to economic and social challenges and holds tremendous potential for productivity, competitiveness and quality of life". From numerous definitions of research data provided, it can be recognised that research data is essential and should be properly managed as a way of promoting research development, growth, visibility and purposes of use and re-use in the universities or other organisations dealing with abundant research data.

To increase the amount of reproducibility in research data, the data should be accompanied by extensive documentation, making it more accessible and secure. In this context, comprehensive documentation is referred to as research data management, which Rice (2022: 2) defines as best practices in the planning, collection, storage, documentation, sharing, and preservation of data generated and

used in research projects. Nwabugwu and Godwin (2020: 2) define RDM as the documentation, curation, and preservation of research data. With this in mind, RDM comes to play in the current digital world and the attendant explosion of multi-disciplinary research data to improve access to research outputs for several fields or organisations such as the educational, commercial and public sectors. Research data management could be considered housekeeping for research data, as it involves the practical steps that are put in place to take care of research data from the beginning of research work. Tripathi, Shukla and Sonker (2017: 418) are of the view that research data management involves activities, as well as the processes involved in ensuring proper documentation, organisation, storage, archival and curation of research data for its availability, ease of access, use and re-use. RDM entails planning, organising, creating explanatory documentation, storing, backing up, as well as providing security for research data and publishing. Bryant, Lavoie and Malpas (2017: 4) note that "Research data management (RDM) has emerged as an area of keen interest in higher education, leading to considerable investment in services, resources and infrastructure, to support researchers' data management needs". Conversely, services, resources, infrastructure as well as policies for supporting RDM activities have been marked as exhibiting some shortfalls in both developed and developing nations, though largely affecting the developing ones. Numerous scholars (Anduvare, 2019; Morgan, Duffield & Walkley Hall, 2017; Avuglah, 2016; Penfield, Cox & Smith, 2014; Molloy & Snow, 2014) among others have cited culture as the greatest challenge hindering the delivery of the RDM services in both libraries and institutions of higher learning. On cultural deficits, Shearer (2015: 5) has observed a lack of infrastructure, services and funding mechanisms to support extensive RDM. Avuglah and Underwood (2019: 13) have also identified a lack of formal RDM infrastructure, with a specific focus on the University of Ghana. On the other hand, Anduvare (2019: 64) noted financial limitations as posing the biggest threat to most African countries in providing RDM services. However, Anduvare also observed financial problems to be common even in developed nations.

Despite the challenges of RDM, some scholars attest to the significant role of managing research data in institutions, as digital data archiving and research data management have become increasingly important for institutions around the world.

Kwanya (2021: 2) observes effective RDM practices to support a solid and reliable foundation upon which to anchor future research and thus advance scholarship; the continued existence of valid data upon which current and future research can be founded; the re-use of research data generated, thereby saving costs; and the saving of researchers' time by avoiding data loss or duplication of efforts by recreating existing data. Koopman and De Jager (2016: 1) have pointed out that the National Research Foundation (NRF), which is one of the major funders of the South African academic research, has recommended the importance of managing research data and archiving digital data they fund in the South African institutions of both higher learning and research. Good management of any data is very important as it aids the verification of research outputs so that researchers can build on existing research. RDM is thus a good research practice for risk management, which is quite essential in the promotion of innovative research in Higher Education Institutions (HEIs) and open science. RDM plays a significant role in fulfilling ethical obligations and avoiding problems such as data loss, theft, confidentiality breaching, and other related problems. Kahn *et al.* (2014: 297) affirm that RDM helps in making research collaboration between institutions, groups and individuals more effective through the reduction of duplication of efforts and data loss. RDM helps in organising data so that it can be easily located on one's computer, and it facilitates the publishing of research data for ease of access. Kahn *et al.* (2014) note that from a researcher's perspective, RDM is vital for the effective management and curation of research data and having access to adequate institutional support. Whyte and Tedds (2011: 2) observe that research data management is significant as it oversees the organisation of data from the initial stage of its entry until the final dissemination and archiving of valuable results. Good management of research data could help fulfil the objectives of HEIs such as research support other than teaching and learning. Additionally, RDM helps in acquiring, processing, analysing, storing and disseminating data, which is essential evidence of research work.

Alongside the wide embracement of the RDM services and activities by the developed nations, some developing nations like South Africa (the focus of the current research) are making significant progress toward offering RDM services in universities, libraries and other research-related centres. Promisingly, among African countries, Van Deventer and Pienaar as cited by Nge’no and Mutula (2018: 39)

identify South Africa as leading in terms of research data management practices. Macanda, Rammutloa and Bezuidenhout (2015) mention higher education institutions (SU, UNISA, UP, WITS, UWC); research councils (CSIR, ARC, HSRC); government entities (DST, NRF, DHET, SANPARKS, SANBI); national research facilities (SAEON, SAIAB, iThemba LABS); and data centres/libraries (projects, funders, themes, disciplines) as RDM units in South Africa. On the other hand, Ng'eno (2018: 7) also observes that South Africa has established some repositories for managing research data, and these include: the South African National Parks; National Health Information Repository and Data Warehouse; and the Data-Intensive Research Initiative of South Africa (DIRISA). Again, Ng'eno and Mutula (2018: 39) point out that "some research councils and institutes, and academic and research libraries in South Africa, have initiated programmes towards the realisation of RDM". It is worth recognising that all these research establishments are making a substantial development of RDM services in the South African nation.

Undoubtedly, globally there are different research services and practices in institutions of higher learning; however, there is still no evidence on how public universities, particularly those based in KwaZulu-Natal (KZN) Province of South Africa, manage their research data. It is for this reason that the current study attempts to establish the RDM practices, services and activities of selected public universities in the country.

1.3 Contextualisation

The structure of the South African university system is known to be determined by the South African Higher Education Act of 1997. Cloete *et al*., (2018: 172) assert that "section 3 of the South African Higher Education Act makes the minister of higher education and training responsible for the governance of the public and private higher education systems at the national level". Previously, the country had 23 public universities, before the introduction of three institutions (Sefako Makgatho Health Sciences University, University of Mpumalanga and Sol Plaatje University) in the year 2014. There are currently 26 recognised public universities in South Africa. The country's universities are divided into three broad categories, namely the universities

of technology, comprehensive universities and traditional universities as presented in Table 1 below:

Table 1. 1: South African public universities by sub-categories

Traditional universities	Comprehensive universities	Universities of technology
University of Cape Town	University of Johannesburg	Cape Peninsula University of Technology
University of Fort Hare	University of Mpumalanga	Central University of Technology
University of the Free State	Sefako Makgatho Health Science University	Durban University of Technology
University of KwaZulu-Natal	Sol Plaatje University	Mangosuthu University of Technology
University of Limpopo	University of South Africa	Tshwane University of Technology
North-West University	University of Venda	Vaal University of Technology
University of Pretoria	Walter Sisulu University	
Rhodes University	University of Zululand	
Stellenbosch University		
University of the Western Cape		
University of the Witwatersrand		

Source: Cloete et al., (2018: 175)

The country has of eleven (11) traditional universities; nine (9) comprehensive universities; and six (6) universities of technology. Cloete *et al.,* (2018: 173) set the orientations for each university category as follows:

- **Traditional universities:** they focus on degree training up to doctoral level and enrol small numbers of diploma and certificate students.
- **Universities of technology:** they concentrate on vocational training through diplomas and certificates and offer a limited number of degree programmes; and
- **Comprehensive universities:** they offer a mix of the programmes of both traditional universities and universities of technology at degree and diploma levels.

Table 1. 2 reflects South Africa's public universities by size and student enrolment.

Table 1. 2: South Africa's public universities according to size and student enrolment

University	Institutional size	Student enrolment
University of South Africa	Large	400, 000
North-West University	Large	74, 355
University of Pretoria	Large	50, 000
Tshwane University of Technology	Large	50, 000
University of Johannesburg	Large	48, 500
University of KwaZulu-Natal	Large	40, 000
University of the Free State	Large	33, 000
Cape Peninsula University of Technology	Large	33, 000
University of the Witwatersrand	Large	32, 703
University of Stellenbosch	Large	30, 150
University of Cape Town	Medium	26, 322
Nelson Mandela Metropolitan University	Medium	26, 000
Walter Sisulu University	Medium	25, 000
Durban University of Technology	Medium	25, 000
University of the Western Cape	Medium	23, 000
University of Limpopo	Medium	20,000
Vaal University of Technology	Small	17, 000
University of Zululand	Small	16, 100
Central University of Technology	Small	13, 534
University of Fort Hare	Small	12, 000
University of Venda	Small	12, 000
Mangosuthu University of Technology	Small	10, 000
Rhodes University	Small	7, 000
Sefako Makgatho Health Sciences University	Small	5, 060
University of Mpumalanga	Small	140
Sol Plaatje University	Small	135

Source: FET college website (2020) and Cloete *et al*, (2018)

Implementing research data management services to support research activities has become a common practice for most universities around the world in the current era, as one of their goals is to support research as well as teaching and learning. Ocholla and Ocholla (2020: 360) recently reported several current trends in academic libraries in relation to providing research support in the 26 South African public universities, including research data services/management (RDS; RDM) among other identified areas. According to Chiware and Becker (2018: 2), the Southern African region has "well-established tertiary education institutions as well as leading

research centres" in comparison to other African nations. Also, Ng'eno and Mutula (2018: 39), citing Van Deventer and Pienaar, vouched for South Africa among other African nations to be the spearhead in terms of RDM practices. South Africa and its distinct leadership on the subject of the RDM is also acknowledged by Avuglah (2016: 22) saying that SA has established some research initiatives, research councils, repositories, research centres and other research-related facilities or units in support of RDM in the country.

Chiware and Mathe (2016: 1) noted that some libraries, primarily university libraries, have made significant progress in providing frameworks for RDM services. Chiware and Mathe (2016) further observe the formulation of policies, establishment of infrastructure, training of the library staff in relevant skills as well as the awareness and advocacy campaigns being mounted for academic staff and researchers to be contributing to the success of the services of RDM. The development of RDM policies in some South African universities is evident in subject literature, such as a report by Macanda, Rammutloa, and Bezuidenhout (2015) revealing that the University of South Africa (UNISA) has completed the investigation of RDM as part of the plan to establish the management of research data in this institution. UNISA currently has a formal RDM policy, which was approved by its council on August 5, 2021. The University of Cape Town (UCT), on the other hand, has an RDM policy called the University of Cape Town Research Data Management Policy, which was published in March 2018 (University of Cape Town Research Data Management Policy, 2018). The University of Pretoria (UP) is another South African university that has an RDM policy in place to govern the management of research data. The Cape Peninsula University of Technology (CPTU) has had an RDM policy in place since July 2014 (CPUT Research Data Management (RDM) Policy, 2016). CPUT collaborates with other universities with the primary goals of establishing international working groups on various aspects of RDM development; developing an open-source platform for managing the RDM cycle; and establishing institutional working groups to contribute to the development of the open-source platform to manage research data. Stellenbosch University also has an RDM policy, which was approved by Senate in November 2020 and is aimed at managing research data at SU to ensure compliance with legislative frameworks, as well as protecting the University, staff, and research participants involved in research through risk

mitigation and management (Stellenbosch University Research Data Management Regulations, 2020: 1). The University of the Western Cape has an RDM policy with a three-year life cycle that was approved in January 2021. Durban University of Technology (DUT) does not have a specific RDM policy; rather, it provides guidelines for research data storage. Furthermore, Stellenbosch University provides some RDM services, although without a formal RDM policy. In Chapter three (section 3.6.1) of the dissertation, which falls under the third objective of the study, a detailed description of RDM policies in universities is provided.

Notably, not all the South African public universities have similar RDM policies; instead, some have RDM procedures and related policies, such as codes of research practice and research ethics that help them define and facilitate the management of their research data. However, it is deemed necessary that RDM activities are in line with research policies in an organisation. This is because a lack of relevant RDM policies in an organisation is more likely to cause some predicaments in the management of research data. Therefore, this study has thrown light on the RDM policies used by the selected public universities in support of RDM to address this gap. Moreover, this study has developed an RDM framework in support of the provision of the RDM services in the public universities in South Africa (see section 7.4.5, Chapter 7).

Despite the establishment of some RDM supporting structures in South Africa, Chiware and Mathe (2016: 1) identify the challenges associated with the provision of RDM services in the country as a lack of resources, insufficient infrastructure, and inadequate data management skills among library staff. Again, most universities appear to have only institutional repositories for managing their research data, rather than the specific RDM services being provided to academics, researchers, staff, and the wider community. In this regard, the current study has shed light on the importance of providing RDM services in the university context, because universities always have a large group of researchers who need proper documentation, storage, sharing, and use of their research data in order to improve their research visibility. Other RDM-related challenges obstructing both universities and libraries are discussed in detail in Chapter 3 of the dissertation (see section 3.7).

1.4 Problem statement

In the current information, knowledge and technological landscape, research data management (RDM) has become a crucial service that is commonly practised in most higher education institutions around the globe. Given that the management of research data is in a burgeoning state, it must have adequate infrastructure, skilled staff, enough resources, and support from institutional management to ensure the effective management of data. However, existing literature reveals that RDM is still in its infancy in most African countries (Nhendodzashe & Pasipamire, 2017; Ng'eno & Mutula, 2018; Chiware & Becker, 2018; Mushi1, Pienaar & Van Deventer, 2020), and could lead to poor accountability and validation of research and insufficient access to previous research work, particularly now that fake research is becoming a challenge in the society.

There are several identified gaps in the literature, including a lack of skills and knowledge; varying cultures among departments or disciplines; a lack of funding; lack of resources/infrastructure; a lack of collaboration between stakeholders; and a lack of necessary policies that are all stumbling blocks to effective RDM practices (CLIR, 2013; Lötter, 2014; Cox & Pinfield, 2014; Shearer, 2015; Avuglah, 2016; Chiware & Mathe, 2016; Nhendodzashe & Pasipamire, 2017; Morgan, Duffield & Walkley Hall, 2017; Gunjal & Gaitanou, 2017; Cox *et al*., 2017; Chiparausha & Kasiroori, 2017; Yoon & Schultz, 2017; Ng'eno, 2018; Cox *et al*. , 2019, Chawinga, 2019; Avuglah & Underwood, 2019; Anduvare, 2019; Chawinga & Zinn, 2020). Despite several identified limitations in providing RDM services, the intended study was underpinned by the issue of varying cultures between disciplines or departments, as well as a lack of skills and infrastructure, as the most cited challenges in the literature. It is hoped that by addressing the study's underlying problems (varying culture, skills gap, and lack of infrastructure), higher education institutions will benefit, as they always have an abundance of research data that must be managed by putting necessary skills and infrastructure in place, as well as adopting a similar culture to facilitate comparable RDM practices. Furthermore, the study could assist research students, particularly postgraduates, in preserving and archiving their research data for future use. Furthermore, the research has provided

an in-depth understanding of the gaps in the literature, as well as areas for development that will improve RDM services in the country.

A failure to address such gaps would result in lapses in the management of research data, in terms of its validation, use, re-use and evidence objectives. Thus far, it is worth noting that South Africa is engaging in most of the RDM-related activities in meeting the demands of the modern era by implementing RDM policies in some HIEs, introducing research councils, government entities, national research facilities, data centres, and libraries as RDM units in the country. Also, South Africa has established some repositories for managing research data and has initiated programmes toward the realisation of RDM. The following is a reflection on the limitations of some previous studies conducted in South Africa that also necessitated the present study, adding to the identified gaps underpinning the present research study.

- The study by Patterton (2016) investigated what organisations like the CSIR could do in ensuring that best practices are applied when managing the CSIR's research data by future researchers. Patterton's study did not cover the issues of policies and procedures, which the current study sought to explore by investigating RDM development strategies including policy, investment, infrastructure and capacity building, in the South African public universities.
- Chiware and Becker (2018) covered the readiness of universities and research libraries in Southern Africa to provide research data management services in their institutions. Chiware and Becker's study, as well as Patterton's, did not propose any framework that institutions could use to improve RDM services. As a result, the current study developed an RDM framework for South African public universities.
- Kahn *et al.* (2014) is another study in the South African context that focused on RDM. However, the study did not propose any RDM framework or model that will aid RDM practices in South African public institutions. Contextually, research data management in KZN universities has not received thorough attention. This study has noted these gaps in previous related studies.

By focusing on the main research question **"How are research data management services delivered in selected public universities in South Africa?",** the current study addressed the identified problems.

1.5 Aim of the study

The study aimed to investigate research data management in selected public universities in South Africa, particularly those based in KwaZulu-Natal.

1.6 Research objectives

The objectives of the study are as follows:

1. To determine the RDM services of the selected public universities.
2. To establish the competency requirements of researchers and staff that are responsible for research data management in the selected public universities.
3. To determine the RDM development strategies of the selected public universities.
4. To establish the challenges facing the selected public universities in managing research data.
5. To develop an RDM framework for public universities in South Africa.

1.7 Research questions

The following research questions were derived from the set objectives:

1. What are the RDM services of the selected public universities?
2. How is the competency requirement of researchers and staff that are responsible for research data management in the selected public universities established?
3. Are there RDM development strategies in place in the selected public universities?
4. What are the challenges facing the selected public universities in managing research data?
5. How can an RDM framework for public universities in South Africa be developed?

1.8 Intended contribution to the body of knowledge

Generally, all research is conducted with the sole purpose of contributing to the existing body of knowledge, bridging existing gaps in the literature, and suggesting further areas of research. There still seem to be very few in-depth studies conducted that solely focus on RDM practices in the South African context, as mentioned by Patterton (2016: 21) that "published information on RDM practices in South African research is scarce, mostly confidential and quite peripheral in nature". One in-depth investigation on RDM practices was conducted by Chawinga (2019); however, it was based on Malawian public universities. Another RDM-related study is by Ng'eno (2018), but its context is Kenyan Agricultural Research Institutes. The study by Patterton (2016) is focused on the South African research council. Chiware and Becker (2018) conducted a study on 'Research Data Management Services in Southern Africa', while Chiware and Mathe (2016) concentrated on 'Academic libraries' role in Research Data Management Services: a South African perspective'. Though these two studies (Chiware and Becker, 2018 and Chiware and Mathe, 2016) are South African-based, they did not cover public universities in this-country per se. This study has contributed to the LIS community in South Africa, and globally theoretically, methodologically and practically.

1.8.1 Theoretical contribution

The current study is significant as it has provided the theoretical foundation for future research in the following domains: public universities, RDM staff, LIS researchers, and LIS curriculum developers. The current study created an RDM framework for South African public universities. The creation of a framework or model would contribute to the advancement of existing theories and models. On the other hand, the current study's theory development could aid in clarifying the relationships between constructs from previously existing theories or models in the South African context.

1.8.2 Methodological contribution

The methodology used by the researchers is thought to have the potential to significantly contribute to the existing body of knowledge. In the current study, a review of empirical studies on RDM as the subject matter aided the researcher in

identifying previously used methods. In-depth related studies, such as Chawinga (2019) and Ng'eno (2018), used the pragmatic school of thought, whereas the current study used the post-positivist philosophy to address nearly the same problem identified in previous research studies. In the current study, the researcher was able to triangulate both qualitative and quantitative data at the data collection and analysis levels by employing post positivist methods. Essentially, the study reviewed and revised existing methodologies, confirming or contradicting their applicability in various situations. As in previous related studies by Chawinga (2019), Ng'eno (2018), Chiware and Becker (2018), and Chiware and Mathews (2018), the current study collected both qualitative and quantitative data, with thematic data analysis used for qualitative data (2018). (2015). On the quantitative side of the collected data, a web questionnaire was completed online, and the data was automatically recorded in an analysable format in a Google spreadsheet that allowed for data tabulation and graphical representation.

1.8.3 Practical contribution

Practically, the present study contributed to the body of knowledge in the LIS research community on RDM practices, in terms of data creation, sharing, archiving, and preservation in the selected public universities in KZN, South Africa. This study also shed light on the competency requirements of researchers and the staff responsible for the provision of research data management services. The study further revealed the RDM development strategies, such as policy, investment, infrastructure and capacity-building programmes. Again, this study also gave information on partnership programmes and policies that support RDM services in the selected public universities. In addition, the study solicited suggestions on the challenges encountered relating to RDM practices. The study outcome could be used widely for research, teaching and learning. Also, the study would inform RDM strategy and policy in South Africa, and perhaps elsewhere.

1.9 Scope and limitations of the study

The study was focused on how South Africa's public universities manage their research data, particularly those based in the KwaZulu-Natal Province of the country. The study was limited to two institutions, viz the Durban University of Technology (DUT) and the University of Zululand (UZ) in terms of the Research Data

Management (RDM) practices, services and activities. Aside from universities in the province of KwaZulu-Natal that do not offer RDM at all, the participating universities' RDM services are encouragingly in the pipeline, with one institution (UNIZULU) having one department that has begun RDM and the other (DUT) having an RDM policy that is still in the draft stage. The rest of the public universities in the country as well as all the private universities or even research units were not involved in the present study due to financial constraints. Then, within the universities of focus, the selection was based on library, teaching and research, whereby librarians, Library Technicians/IT specialists, Research Directors (RD), Deans (D)/ Deputy Deans (DD) Research, Deputy Vice-Chancellors (DVC) Research, Heads of the Departments (HODs) and academic researchers (NRF rated) were selected as the study participants.

1.10 Research methodology overview

This section provides a brief overview of the employed research methods, designs, techniques and procedures comprising the entire research methodology for the undertaking of this study. Previous related studies appear to use pragmatism, mixed methods, both qualitative and quantitative research inquiry, and survey in the form of questionnaires and interviews, followed by data analysis using SPSS and thematic analysis methods. The current study was triggered on the post-positivist research paradigm and employed both qualitative and quantitative methods of investigation. On the quantitative side, a Google forms questionnaire was used to collect data from academic researchers (NRF rated), and there was no sampling drawn for this population because the researcher chose them all. In the qualitative aspect of the current research, semi-structured interviews were conducted online through Zoom and Microsoft Teams in the form of multiple case studies to gather data from the Librarians, Library Technicians, Research Directors, Deans/ Deputy Deans Research, Deputy Vice-Chancellors Research and Heads of the Departments who were selected purposively. On the quantitative side of the collected data, a web questionnaire was filled in online, and the data was automatically recorded in a Google spreadsheet in an analysable format that allowed for data tabulation and graphical representation, whereas the thematic analysis method was applied on the qualitative side (interviews). The reliability and validity of researcher instruments

were observed. Ethical considerations too were duly observed throughout the entire research project to ensure that the study was ethically carried out. A detailed description of the entire research methodology is provided in Chapter four of the dissertation.

1.11 Knowledge dissemination

The effective dissemination of knowledge is about getting the findings from particular research to the right people that can make use of them. In the current study, the researcher engaged with the audience of the study and kept in touch with them throughout the study. The researcher utilised already available opportunities, such as existing conferences (national and international) in the LIS domain to exchange knowledge and raise awareness about the RDM services in the higher education context in South Africa and globally. Some findings of the study were shared earlier as preliminaries, through conference paper presentations. A dissertation was written and submitted to the University of Zululand institutional repository for global access. From the dissertation, some research publications such as journal articles, conference proceedings and book chapters were harvested. To disseminate the findings, the researcher went an extra mile by sharing the findings of the study through lectures, workshops, short courses, demonstrations and simulations. Table 1.3 is a reflection of some of the publications in which preliminary reports have been published from the thesis.

Table 1.3: Publications where preliminary reports have been published from current thesis

Title	Author(s) & date	Publisher/ platform
Development strategies as catalysts for provision of the RDM services in the South African higher education institutions	Mthembu and Ocholla (2022)	*South African Journal of Libraries and Information Science*, vol88, issue 1.
Research data management competencies of researchers at the University of Zululand, South Africa	Mthembu (2022)	6th -8th September 2022, 22nd Hybrid IS annual conference, University of Zululand (presented preliminary findings).
Research data management in higher education institutions and libraries: Literature review	Mthembu (2023)	Submitted a book chapter edited by Dennis Ocholla, Bosire Onyancha and Veli Jiyane 'A handbook for Information Knowledge and

		Technology Research and Teaching in Africa' to be published by UNISA Press
Community Capability Model Framework, Digital Curation Centre Lifecycle Model and Data Management Framework in Research Data Management: A systematic Literature Review	Mthembu, Adesina and Ocholla (2021)	The Second Biennial University of South Africa International Conference on Library and Information Science Research in Africa (UNILISA), virtual conference (presented a paper)

1.12 Definition of key terms

Research data: Research data involves data that has been created by researchers during their work (Tripathi, Shukla and Sonker, 2017: 417). An institution has a curation responsibility for research data, for as long as there are some requirements from record keeping or archiving acts or regulations. For the University of Pretoria RDM policy (2017: 6) and OECD Principles and Guidelines (2007), research data is viewed as factual records, and that could include numerical scores, textual records, images and sounds that are normally used as primary sources for scientific research and acknowledged in the scientific community as necessary to validate research findings. The University of Southampton's RDM policy (2019: 2) research data involves information presented in a digital or paper-based format that could be obtainable in several means, e.g. notes, facts, figures, tables, images, audio, recordings, etc. and can be gathered or generated during a research study through means of trials, surveys, interviews, focus groups and so on.

Research data management: Research data management (RDM) involves the process of organising data, right from the entry-level into the research cycle, till the dissemination and archiving of the valuable results (Whyte & Tedds 2011; Gunjal & Gaitanou 2017; Crane 2018). RDM plays a significant role in making the research process as efficient as it possibly could, while meeting the expectations and requirements of a particular institution, funders of research and regulatory bodies. Charles Darwin University Research Data Management Procedures (2019: 2) assert that "research data management includes all the processes and actions required to manage data using good practice throughout the research life cycle for current and future research purposes and users".

Data preservation: Data preservation is concerned with the conservation, maintenance, safety and reliability of data. It involves formal activities that are exclusively governed by policies and regulations. The University of Pretoria RDM policy (2017: 5) indicates that data preservation involves actions that are taken in ensuring digital collection remains usable, regardless of the future changes in technology as technology is noted for constantly and drastically changing.

Data curation: Data curation involves the organisation, as well as the incorporation of collected data from different information sources. Data curation is much like archiving, thereby allowing the availability of data for re-use and preservation to enhance good research practice within the research community.

Data archiving: Data archiving is the process of moving data that is scarcely used for long-term retention. Scarcely or less used data often remains crucial in an organisation for future reference.

Data re-use: Data re-use is concerned with the utilisation of research data for research purposes, or even for purposes not necessarily related to what the original data was created to do (He and Nahar, 2016). For data re-use, several platforms could be used, such as an institutional repository.

Ingest: Data ingestion refers to the process of collecting and importing data from various sources, which could be in different formats, into a database or other storage systems. This process is essential for businesses and organizations to gather and manage data effectively. In this context, data ingestion is critical to the success of universities, allowing them to manage and analyse data for a variety of purposes, including research.

1.13 Outline of the thesis

Chapter One: Introduction and background

This is the study's introductory chapter, which covers its background. The contextualisation, problem statement, research aim, objectives and research questions, significance, delimitation of the study, an outline of the thesis, and definitions of key terms of the study were all covered in the study's background.

Chapter Two: Theoretical framework

This chapter detailed the community capability model framework (CCMF) and digital curation centre (DCC) lifecycle model used in this study.

Chapter Three: Literature review

This chapter provides insight into the concept of research data management based on a review of global literature (RDM). The review of literature focuses on RDM practices, particularly in universities and libraries; competency requirements for researchers and staff responsible for research data management in universities; RDM development strategies such as policy, investment, infrastructure, and capacity building; and the challenges that public universities face in managing research data.

Chapter Four: Research methodology

This chapter reflects the techniques and procedures used to carry out the entire study, and it includes the research paradigm, approach, method, study's targeted population, sampling procedure, and data collection and analysis methods. This chapter also addresses ethical concerns.

Chapter Five: Data presentation and analysis

Two participating institutions provided both qualitative and quantitative data (Durban University of Technology and University of Zululand). This chapter presented, interpreted, and analysed findings. It provides an interpretation of the data gathered through interviews and questionnaires.

Chapter Six: Discussion of findings

The study's findings were discussed in the chapter. The findings were derived from the study's research objectives.

Chapter Seven: Summary of findings, conclusions and recommendations

The findings, conclusions, and recommendations of the study were summarised in this chapter, which also discusses the study's originality and contribution. Furthermore, this chapter includes suggestions for future research.

1.14 Chapter summary

This was an introductory chapter that laid the groundwork for research data management as the studied phenomenon, which helped to put the entire study into context. The conceptual setting of the study was addressed by providing explanations of research data, its management, importance, and RDM practices in the global context, revealing that developed nations embrace RDM more than developing nations. The study's context was deliberated upon by representing South African public universities, particularly those of interest to the current study, as well as providing motivations for their selection. A brief overview of the RDM services in the South African context was addressed and it was established that South African public universities come across several challenges of RDM, including a lack of funding, skills shortage, inadequate resources/infrastructure, and lack of necessary policies, among others. The present research was therefore underpinned by varying cultures between faculties or disciplines, skills deficits and improper infrastructure as the most identified challenges in the literature. Moreover, the study's objectives, questions and intended contribution to the body of knowledge were deliberated. The chapter has further provided the scope and limitations of the present study and furthered the indication of how knowledge was disseminated. The structure of the thesis was also highlighted. The next chapter details the theoretical frameworks adopted for the study.

CHAPTER TWO: THEORETICAL FRAMEWORK

2.1 Introduction

This chapter discusses the theoretical framework that guided the current study on the investigation of the research data management services in selected public universities in South Africa. Often people, particularly students, tend to use the terms 'Theoretical framework' and 'Conceptual framework' interchangeably or synonymously. However, it should be noted that these two terms differ yet are more similar in their roles and or functions in research. Regardless of their roles and functions being comparable to some extent, the fact remains that they need to be understood as varying concepts. Adom and Hussein (2018: 438) are of the view that theoretical and conceptual frameworks significantly explain the path of research as well as grounding it in theoretical constructs. Both frameworks are believed to make research findings more meaningful and acceptable to the theoretical constructs in the field of research as well as enhance generalisability.

Nonetheless, a theoretical framework assists in examining research data, linking research methodology, questions and literature together. For Osanloo and Grant (2016: 17), a conceptual framework is a logical structure showing interrelated concepts that helps in picturing the phenomenon being studied. Adom and Hussein (2018: 438) assert that a theoretical framework is a blueprint that is usually borrowed by researchers to build their house of research enquiry. Osanloo and Grant (2016: 13) concur that a theoretical framework is a blueprint for the dissertation enquiry. On the other hand, Adom and Hussein (2018: 438) note the significance of theoretical framework in research as helping researchers to undertake their studies which significantly contributes to scholarly and academic research as it acts as a guide or plan that contextualizes formal theories in research studies. Research design and data analysis plans are the result of a theoretical framework (Adom and Hussein, 2018: 438; Osanloo & Grant, 2016: 15). Chawinga (2019: 58) opines that a theoretical framework acts as a guiding principle to answering questions that may sometimes be speculated on by researchers and also helps to provide very conclusive explanations in research.

Previous related studies (Masinde, Chen & Muthee, 2021; Chawinga, 2019; Ng'eno, 2018; Ng'eno & Mutula, 2018; Patterton, 2016; Avuglah, 2016; Shen, 2016; Lyon, Patel & Takeda, 2014) have used the Community Capability Model Framework (CCMF), Data Curation Centre Lifecycle Model, Data Audit Framework (DAF), and Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO) in understanding the services of RDM. The use of CCMF by Chawinga (2019), Nge'no (2018), Ng'eno and Mutula (2018), Shen (2016) and Lyon, Patel and Takeda (2014) acted as the motivation for its application in the present study, to comprehend the nature of the RDM services in the selected public universities in South Africa. The current study did not adopt the CCMF in isolation; instead used it jointly with the Digital Curation Centre (DCC) Lifecycle Model.

2.2 Overview of theories underpinning the study

In the context of the present study, the application of the two frameworks which are the community capability model framework (CCMF) and the digital curation centre (DCC) lifecycle model underpinning the study is linked with the set research objectives of the study wherever possible. It is however not all the objectives of the study that have been covered by these frameworks; instead, some of the elements in the set objectives have been addressed in the literature, in the next chapter (Chapter 3). The eight capability factors of the CCMF were linked with the following set objectives of the study to motivate its application and relevance in the current study:

- To determine the RDM practices of the selected public universities.
- To establish the competency requirements of researchers and staff that are responsible for research data management in the selected public universities.
- To determine the RDM development strategies of the selected public universities.
- To establish the challenges facing the selected public universities in managing research data.

2.2.1 Community Capability Model Framework (CCMF)

CCMF is a tool that was developed by the UKOLN, the University of Bath in collaboration with Microsoft Research to assist institutions, research funding bodies,

as well as researchers, to strengthen and promote the capabilities of their varying communities so that they can perform more data-intensive research (Lyon *et al*. 2012: 5). It is further viewed by Lyon *et al.* that data-intensive research can be performed through:

- Profiling the current readiness or capability of the community,
- Indicating priority areas for change and investment, and
- Developing roadmaps for achieving a target state of readiness.

Higher education institutions as the focus of the current study have been noted as always dealing with intensive research data, as one of their objectives is to support teaching, learning and research. Lyon *et al*. (2012: 10) highlight two characteristics in the CCMF that could be considered the major indicators for data-intensive research in an institution:

- the research involves intense computational analysis of data and.
- research as involving the analysis of a large amount of data, known as big data, that could not be expected to be reviewed by a research team without the aid of software.

According to Lyon *et al.* (2012:10), in the CCMF, a community could be viewed as a "set of people who share a particular location with the structure of an institution in general". In the context of the present study, 'community' includes the staff responsible for RDM and the researchers from their respective institutions as the study's participants. Lyon *et al.* (2012) highlight eight capability factors in the CCMF, including collaboration; skills and training; openness; technical infrastructure; common practices; economic and business models; legal and ethical issues; and academic culture, respectively and thus discussed later. Figure 2.1 shows the diagrammatic illustration of the CCM Framework.

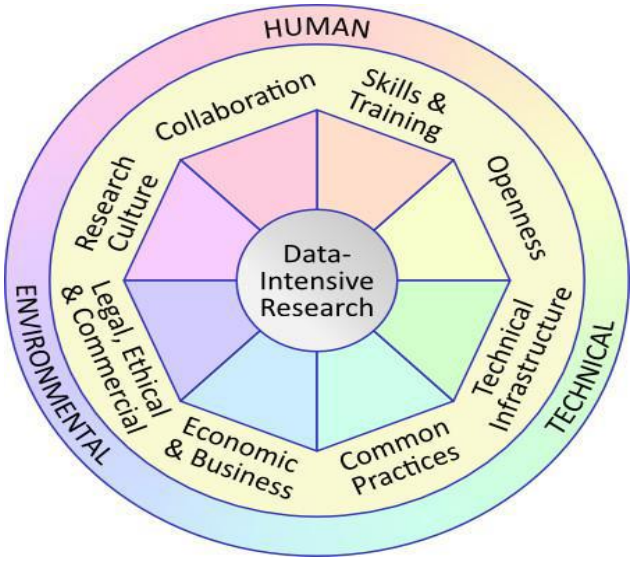

Figure 2. 1: Community Capability Model Framework (CCMF)

Source: Lyon *et al*. (2012: 3)

2.2.1.1 Collaboration

One of the eight capability factors in the CCMF is collaboration, which is often associated with the kind of relationships being formed during research processes. Lyon *et al*. (2012: 21) state that the types of relationships established during research are a significant determinant of the types of research that can be performed in an organisation. Collaboration can be done between a section or department, across departments, within the institutions or with external bodies. In the context of the present study, collaboration is linked with the third objective of the study (to determine the RDM development strategies e.g. policy, investment, infrastructure, and capacity building in the selected public universities) with a specific focus on capacity building, known for its crucial role in the development of activities in an organisation, for example a university. For example, in the next chapter (Chapter 3) some global and joined projects for enhancing good practices of RDM have been discussed (see section 3.6.3). With that in mind, it can be recognised that collaboration between the institutions of higher learning and other recognised research centres could bring about good RDM services, and practices, enhance sound scientific research as well increasing research visibility. Lyon *et al*. (2012: 22) believe that "where the driver for data-intensive research is the sheer scale of the data needed to answer the research question, is a key strategy for drawing together sufficient capacity to perform the collection and processing".

2.2.1.2 Skills and training

Skills and other competencies are a major determinant of a community to performance data intensive research (Lyon, 2012: 25). Therefore, provision of training is considered highly significant and can be provided through early education and professional development. The skills and training section has been covered in the literature (Chapter 3) in section 3.5 by discussing the second objective of the study (to establish the competency requirements of researchers and staff that are responsible for research data management in the selected public universities). Relevant skills and knowledge to be possessed by both staff responsible for RDM and researchers have been covered (see section 3.5.1); whereas the development strategies for skills development have been addressed in section 3.5.2, Chapter three of the dissertation. Lyon et al. (2012: 25) mark tools and technologies (e.g. cloud computing, visualisations, statistical analysis); data description and identification such as (metadata, vocabularies, citation); and policy and planning (data management, business models included)" as significant skills set for researchers. But today, as born-digital data is gradually developing and spreading like cancer cells, the mere fact is that skills or capability requirements can never be overemphasised.

2.2.1.3 Openness

Openness is a capability factor concerned with research communication in open platforms such as institutional repositories. The current study chose librarians and library technicians for their active roles in data selection, acquisition, preservation, institutional repositories, archiving, and sharing in the university context in order to better understand participating institutional research platforms that promote open access to research data. Lyon *et al* (2012: 26) indicate that openness of research results and methods can be achieved during the research study and in the publication stage to reach a wider audience. Opening up research data is viewed as adding value to research as far as validation, reproducibility and re-usability of research are concerned (Lyon *et al,* 2012: 26). In the context of the present study, issues of openness have been discussed in lines of institutional data repositories in support for RDM (see section 3.6.2.4, Chapter three of the dissertation). Section 3.6.2.4 falls under the third objective of the study, which is ‘to determine RDM development strategies’. In addition, this capability factor focuses on making

research data more freely available or with fewer restrictions, and that could contribute to not only furthering research and development but instead avoiding the duplication of efforts in an institution.

2.2.1.4 Technical infrastructure

Lyon *et al*. (2012: 31) assert that technical infrastructure supporting research is marked with the necessary tools and services in line with different stages of the research life cycle. To understand technical infrastructure for supporting RDM services for the participating universities, library technicians were interviewed as they play a significant role in implementing the infrastructure to support research activities, such as acquisition, storage, security, integration, preservation, and archiving, searching among other research-related activities. Lyon *et al*. state that this factor (technical infrastructure) is concerned with meeting the needs of users in their varying activities in research, including data capture and processing; storage of data; curation and preservation; discovery and access, among others. Issues concerning this capability factor have also been discussed in the literature when discussing the study's third objective (see section 3.6.2, Chapter three). In the context of the current study, a good technical infrastructure is one that is capable of supporting RDM service provision in the university setting by being welcoming to its users, who may include both researchers and RDM staff.

2.2.1.5 Common practices

This factor involves a clarification of community practices that have significantly produced some standards in various areas like data formats, methods of data collection, processing workflows, packaging of data, transfer procedures, data description and other areas (Lyon *et al.*, 2012: 36). Lyon *et al.* further point out that the quality of standards is what matters other than their quantity, as the quality could indicate whether standards are capable of promoting and enabling re-use as well as a combination of data in a discipline. To promote RDM practices, common standards are frequently associated with a common research culture of practice in a university. In contrast, as organisations and disciplines have varying cultures in their research practices, the issue of culture has been marked with some barriers and constraints in common RDM practices in the literature. Such impediments have been observed to impede good and common RDM practices in the majority of institutions of higher learning.

2.2.1.6 Economic and business models

This capability factor touches on the aspect of investments in an organisation to achieve the goal of data-intensive research. In the present study, issues about investment were addressed in the literature review (see section 3.6.2) together with infrastructure covered in the third objective of the study. Lyon *et al.* (2012: 39) base the investments on research and infrastructure on funding models. In this respect, research multi-phase thematic investments, bilateral initiatives between national funders as well as major investments in longitudinal data surveys are included. In terms of infrastructure, sustained multi-investments in data centres, collaborative development between international funders and large central investments in network infrastructure could be used as the funding models. Institutions appear to make significant investments in new and existing infrastructure to support the provision of services to their user community, RDM services in the context of the current research. As a result, universities can recognise that they can invest in already existing infrastructure, such as institutional repositories, to support the provision of RDM services.

2.2.1.7 Legal and ethical issues

This capability factor of the CCMF exclusively addresses the ethical issues that are usually embedded in data sharing practices which could then help in the provision of restrictions for data-intensive research in an organisation (Lyon *et al*., 2012: 40). Therefore, Lyon *et al*. postulate the significance of understanding and managing legal issues surrounding data sharing and re-use to promote smooth data-intensive research as well as to ensure that research is undertaken more ethically in an institution. In the present study, Research Directors (RDs) were selected as the study's participants for their active role in ensuring that research processes in the university are ethically carried out. By involving RDs in the current research, insight regarding the legal and ethical issues was gained. Legal and ethical issues were also covered in the third objective of the study, particularly in section 3.6.1 (Chapter three) with a specific focus on RDM policies and procedures in the global context, at universities exclusively.

2.2.1.8 Academic issues/ culture

Regarding the academic issues, Lyon *et al*. (2012: 42) believe that community norms aligned with research practices are a major determinant of the level of support that

could be received by a researcher as moving to more data-intensive research. Lyon further states that a good flourishment of data-intensive research is normally acknowledged in communities that tend to give a high value to data, for instance in organisations in which researchers are rewarded for their data contributions as well as the expectations of the high standards in terms of data entering the research record.

2.2.2 Relevance, application and gaps in the CCM Framework

The application of the Community Capability Model Framework has significantly helped in the evaluation of the selected public universities' current readiness to perform data-intensive research as well as identifying areas where changes need to be made to increase their capability. Though the present study has five set objectives, the application of the CCMF has been addressed greatly in the third objective of the study, which was '*to determine the RDM development strategies (e.g. policy, investment, infrastructure, capacity building) in the selected public universities*.' This is due to the fact that this objective covers the most capability factors of the CCMF including technical infrastructure, openness, and legal and ethical issues, among others.

CCMF has been applied and significantly succeeded in determining RDM services, practices and activities in the previous related studies like Chawinga (2019); Ng'eno (2018); Ng'eno and Mutula (2018); Shen (2016); Lyon, Patel and Takeda (2014). Chawinga (2019), as well as Ng'eno and Mutula (2018), are RDM studies conducted in the African context, whereas, Shen (2016) and Lyon, Patel and Takede (2016) were based in Western nations. Chawinga (2019) used the CCMF in investigating research data management practices in public universities in Malawi with a specific focus on research data generation, organisation, sharing, storage and preservation for re-use and long-term access. Ng'eno (2018) used this framework to investigate the research data management practices at some selected agricultural research institutes in Kenya. Ng'eno and Mutula (2018) used the same framework to understand the nature of Research Data Management (RDM) in the agricultural research institute in Kenya. Shen (2016) used the CCMF to analyse data sharing activities amongst researchers at Virginia Tech. Lyon, Patel, and Takeda (2014) used this model for assessing the requirements of RDM in academic libraries to

develop a simple self-assessment tool primarily for researchers to use and catalyse the collection of disciplinary profiles to inform RDM service development in libraries, research funder investment decisions and policymaking in the data arena.

Though the CCM framework covers most aspects that form RDM, it lacks the element of data curation, which is indeed a crucial part that forms the basis of RDM. Such a gap, therefore, affects the issues of re-use and preservation of data which could negatively impact the value of data at the university. Therefore, the current research bridged this gap by combining this framework with the DCC Curation Lifecycle Model to complement it while moving closer to the understanding of RDM practices and services at selected public universities in the country South Africa.

2.2.3 The DCC Curation Lifecycle Model

The DCC Curation Lifecycle Model was developed in 2007 by the Digital Curation Centre (Constantopoulos *et al*, 2009: 38). Higgins (2008: 134) asserts that the "Digital Curation Centre (DCC) Lifecycle Model has been developed as a generic, curation-specific, tool which can be used, in conjunction with relevant standards, to plan curation and preservation activities to different levels of granularity". The Digital Curation Centre in the United Kingdom uses the model as a training tool for curators, which significantly helps them understand the processes involved in successful curation as well as developing curation and preservation methodologies for their organisations. Higgins goes on to reveal another importance of this model in the DCC SCARP Project3 as they are using it in their immersive case studies, to help subjects in the identification of the possible risks that current practice places on their data as well as the necessary steps in mitigating those risks. Again, the model is also "being used by the DCC Services team to ensure resources cover all areas of the lifecycle" (Higgins, 2008: 136). Figure 2.2 is the DCC Curation Lifecycle Model in diagrammatic representation, indicating an overview of the lifecycle stages required for successful curation in an organisation.

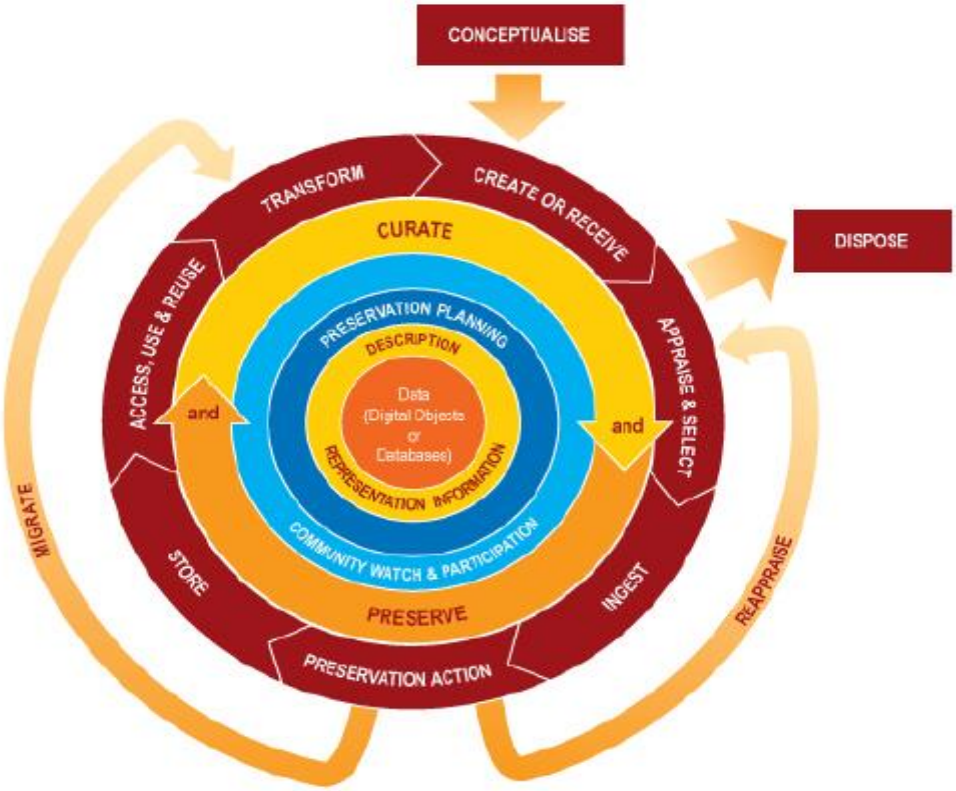

Figure 2. 2: DCC Curation Lifecycle Model

Source: Higgins (2008: 136)

To better understand the processes of curation and preservation in the DCC model, Higgins (2008: 137) highlights its full lifecycle actions, sequential actions and occasional actions as forming the bases of this model and thus will be discussed below.

2.2.3.1 Full Lifecycle Actions

a) Description and Representation of Information

For description and representation of information, Higgins (2008: 137) postulates assigning metadata in lines of administrative, descriptive, technical, structural and preservation through the utilisation of suitable standards for ensuring adequate description and control in the long term. Higgins further adds that there should be a collection as well as assigning of representation information that is needed in understanding and rendering both the digital material and the associated metadata. These actions are done to get the ball rolling. In the context of the current study, it is important to recognise the importance of implementing necessary RDM policies and

procedures, as well as other related policies, to facilitate the entire process of research data description and representation at the university.

b) Preservation Planning

The planning of preservation is an action that involves a proper plan for preservation during the curation lifecycle of digital material. Higgins (2008: 137) views that plan for preservation would entail all curation lifecycle actions in terms of the management and administration. In the context of the current study, preservation planning is linked with the research data management plan (RDMP), which entails necessary steps to be taken before the beginning of any research project. RDMP was broadly discussed in Chapter three of the dissertation (see section 3.6.2.1). In a university context, preservation planning significantly helps a university set preservation guidelines for both the present and future which support the allocation of resources as well as investment.

c) Community Watch and Participation

Higgins (2008: 137) suggests a watch on appropriate community activities that should always be continuous, as well as the participation in the development of shared standards, tools and suitable software. In this context, community activities involve all the research data management-related activities of both researchers and staff responsible for the provision of RDM services in the university setting. In terms of tools and software, staff responsible for RDM, particularly the technicians were interviewed as the ultimate judges of the IT infrastructure in support for RDM services, as IT infrastructure entails storage, network bandwidth, relevant software packages and other elements forming its foundations.

d) Curate and Preserve

In terms of the curation and preservation, which is indeed a crucial aspect that has been noted as lacking in the employed CCM framework, Higgins (2008: 137) asserts that there should be an awareness and the undertaking of management and administrative actions that have been planned to help promote curation and preservation throughout the curation lifecycle. Preservation issues were discussed in several RDM policies (e.g. University of Southampton Research Data Management Policy, 2019; Charles Darwin University RDM Procedures, 2019; Crane, 2018; University of Twente RDM policy, 2018; University of Pretoria RDM policy, 2017;

Cape Peninsula University of Technology RDM policy, 2016) in the literature reviewed for the current study (Chapter three, section 3.6.1). Data curation has been shown to significantly increase the value of data and increase its availability for preservation and re-use within an organisation.

2.2.3.2 Sequential Actions

a) Conceptualise

The creation of data, which includes options and methods of capture and storage, should be well comprehended and properly planned (Higgins, 2008: 138). Conceptualising information or data is of paramount importance to clarify concepts, as a researcher could discuss concepts in the context of their understanding, whereas others from other disciplines may have a completely different understanding of those concepts. Therefore, a research data management plan (RDMP) is a requirement for a sound and successful conceptualisation of any research data. RDMP and its relevance have been deliberated on in Chapter three (section 3.6.2.1) of the dissertation.

b) Create and Receive

The actions of creating data include the administrative, descriptive, structural and technical metadata. In this stage, Higgins (2008: 138) notes the preservation of metadata which may also be added at the time of creation. Higgins further notes the significance of having in place proper documentation for collecting policies, from data creators, other archives, repositories or data centres when receiving data. This marks a substantial need for having in place RDM policies or procedures to facilitate the processes of creation and receiving of data in the university. In relevant literature, the relevance of having in place RDM policies have been documented. In this study, issues related to RDM policies have been addressed in section 3.6.1 of the dissertation, under the third of the objectives of the study.

c) Appraise and Select

An idea of the amount of data should be formed, and the selection for long-term curation and preservation should be considered in the appraisal and selection actions (Higgins, 2008: 138). Higgins emphasises the adherence to documented guidance, policies and/or legal requirements in taking actions related to appraisal and selection of data. In this respect, the relevance of RDM policies in the university

context for selecting research data must never go unnoticed as has been most succinctly noted in actions forming the DCC model. Seemingly, most actions using this model entail the adoption of the policies and or other legal necessities.

e) Ingest

Ingest involves a transfer of data to an archive, repository, data centre or other curators. Again, emphasis is also placed on the adherence to the documented guidance, policies or legal requirements for the actions of appraisal and use. The ingestion process brings into recognition having data repositories in universities and not omitting policies for their significant role in facilitating almost every step in the curation and or preservation of research data. The relevance of institutional data repositories in support of RDM was deliberated on in Chapter three of the dissertation (see section 3.6.2.4). The issues of policies have been emphasised in most sections of the dissertation (see sections 3.6.1, Chapter three; 2.2.1.7, 2.2.3.1 to 2.2.3.5, present chapter).

2.2.3.3 Preservation Actions

Preservation actions involve undertaking actions to ensure long-term preservation and retention of the authoritative nature of data (Higgins, 2008: 138). Preservation actions are believed to facilitate authentic, reliable and usable data at all times, while maintaining its integrity. Among other actions in the preservation of data, Higgins remarks on data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.

a) Store

Data should be kept securely, and its storage should match relevant standards. In literature (Chapter three, section 3.6.2.2), data storage, management and organisation have been discussed. Data storage has been revealed as customarily facilitated by institutions' RDM policies or procedures and most policies stipulate a 5 and 10-year period of data storage, preservation and curation. Storing research data is a crucial aspect of research data management, which prevents data loss or damage in an institution.

b) Access, Use and Re-use

For research visibility, an organisation should always ensure that data is accessible to both designated users and re-users, daily (Higgins, 2008: 138). The publicly available published information can re-enforce the actions of access, use and re-use in an organisation, a university in this context. Though some of the information may be publicly available, Higgins (2008: 137) proposes that strong access controls and verification procedures be applied. Issues relating to data use were also addressed in section 2.2.1.7 (current chapter) in the CCMF model as another employed model in the current research. Again, the use of research data was also covered in Chapter three (section 3.2.2.3) of the dissertation.

c) Transform

Higgins (2008: 138) believes that a transformation of data could be undertaken by creating new data from the original. Higgins states that this could be achieved through migrating data into a different format or by way of creating a subset of a query selection.

2.2.3.4 Occasional Actions

a) Dispose

One of the elements forming the basis of the occasional actions is the disposal of data, which Higgins (2008: 138) views as the data that has not been selected for purposes of long-term curation and preservation in line with the documented policies, and guidance or legal requirements. The relevance of having RDM policies in a university has been discussed in Chapter three of the dissertation (see section 3.6.1) as one of the development strategies which forms the third objective of the current research study. Higgins then points out that data could be transferred to other archives, repositories, data centres or another custodian. However, it has been observed that data is destroyed in some cases. Therefore, Higgins alludes to the fact that the nature of the data may require secure destruction for legal reasons.

b) Reappraise

Reappraisal of data involves the actions of returning data which fails validation procedures for further appraisal and reselection.

c) Migrate

As the term suggests, data may be migrated to a different format. Higgins (2008: 138) claims that the migration of data may "be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence". The migration of data may depend on the institutional infrastructure as well as the preferred format which an institution uses for capture in its system. The process of migrating data involves technicians, as they are familiar with the system used for the entire research data management lifecycle. Technicians were therefore interviewed in the current study to understand the IT infrastructure used and suggest areas of development wherever necessary.

2.2.4 Relevance, application and gaps in the DCC curation lifecycle model

The application of the DCC model has significantly bridged the gap identified from CCMF as another employed model in the current study. DCC accounts for the element of data curation that the CCMF currently ignores; such element contributes to the crucial part that forms the basis of RDM in an organisation. Such a gap, therefore, impends on the issues of re-use and preservation of data which could negatively impact on the value of data at the university. In this regard, this model helped to reflect on how to plan activities within a specific research project in an organisation, a university per se, to ensure that all the necessary stages are undertaken, each of them in their correct sequence. Undertaking stages in the correct sequence contributes to the documentation and integrates curation policies between the stakeholders involved which could include the data creators, curators as well as data re-users in the process.

The DCC Curation Lifecycle Model, like the CCMF, also seem to be a popular and widely used data lifecycle model in most RDM studies. This model was previously applied jointly with other models in the related study by Chawinga (2019), the "investigation on research data management in public universities in Malawi" which sought to understand how research data is generated, organised, shared, stored, preserved, accessed and re-used in the public universities of that country (Malawi). Again, this model was also jointly used with RDM-related models by Ng'eno (2018) in understanding research data management in Kenya's agricultural research

institutes. Ng'eno's study used this theoretical model to propose interventions to improve the management, sharing and re-use of agricultural research output. This model was also used in Barton, Bracke and Clark's (2016) study in investigating the intersections and alignments between scholarly workflow and curatorial workflow, and the implications of these intersections and alignments in collaborative research and curatorial lifecycles. Lee and Stvilia (2017) is another RDM-related study that employed the DCC Curation Lifecycle Model in providing a rich, qualitative description of research data curation and practices in institutional repositories (IRs) of the research universities in the United States of America. Shen and Varvel (2013) applied a similar model to understanding the consultancy roles and requirements within the complexity of data professionals at Johns Hopkins University (JHU). Again, a similar model was adopted by Knight (2012) to gain an insight into the information management practices, policies, and procedures applied by data creators and managers in the business units and the risk factors that may limit access to and use of their digital assets over time.

Though there are some strengths in the DCC as highlighted above, this model is also presented with some shortcomings, like a lack of some aspects that are covered in the CCMF including the technical infrastructure; skills and training; collaborative partnerships and legal and policy issues (Ng'eno, 2018: 34). Palavitsinis, Manouselis and Sanchez-Alonso (2010: 49) reveal that DCC was "yet an emerging field with many different contributions from a great number of scientists that make it even more difficult to define concepts and theories". Palavitsinis, Manouselis and Sanchez-Alonso (2010) and Constantopoulos *et al*., (2009) Barton, Bracke and Clark (2016) observed DCC as lacking a clear insight regarding the action of adding new knowledge to the primary resources kept in digital repositories. On the other hand, Giarlo (2013) highlights the DCC lifecycle model as lacking a guiding principle for recording and maintaining stored statistical information and curated and preserved information for the users. Since both CCMF and DCC have some shortcomings, the current study has used them both to fill in the gaps that one model has. These research lifecycle models are not mutually exclusive and, as such, provide greater benefit when combined. They are also useful tools for researchers and professionals dealing with research data in a university context or any organization dealing with large amounts of data. Combining these models can also significantly increase the

knowledge and skills of researchers. The application of these models to improve quality data preservation over time can also facilitate tool adoption in terms of infrastructure.

2.2.5 Motivation and implications for using theoretical models in the present study

Most related studies like Chawinga (2019); Ng'eno (2018); Ng'eno and Mutula (2018); Patterton (2016); Avuglah (2016); Shen (2016); Lyon, Patel and Takeda (2014) never used either of the models in isolation, instead integrating it with other models to better understand the nature of the RDM in various contexts. As is the case with the present research, the two employed models have been jointly used and complemented each other, which gave a complete insight into the RDM services in the selected public universities in South Africa. Their complimentary benefits lie in the fact that they are having similar elements such as the issues of policies, infrastructure, skills and training needs which form a strong foundation for the RDM lifecycle and its processes. Where CCMF lacks the aspect of data curation, the DCC Curation Lifecycle Model was employed to complement that gap. On the other hand, as the DCC is also presented with some inadequacies, like a lack of technical infrastructure; skills and training; collaborative partnerships and legal and policy issues, CCMF has significantly bridged that gap as it consists of these elements as its attributes. Therefore, employing both models in a single study has contributed by bringing insight into how the South African public universities can successfully implement the RDM services and practices.

CCMF and DCC Lifecycle Model are all capability models which Lyon *et al*. (2012: 10) view as models that are "widely used by industry to help them in identifying key business competencies and activities, helping to determine whether, how easily, and how well a given organisation or community would be able, in theory and practice, to accomplish a given task". In this respect, the adoption of these two models in the present study has provided significant benefit for the staff responsible for RDM to better understand the context of the institutions they are serving, so that they are always equipped with the necessary competencies. Not only the competencies of the staff responsible for RDM support have been revealed, but also those of the researchers. The models have also complemented each other in terms of shedding

light on the current state of RDM service delivery to the participating institutions. Also, the models have enabled the organisations to see where gaps and duplications exist, such as in terms of infrastructure and staff, in the current state of RDM service provisions. Furthermore, a synthesis of the models has helped in the development of a new proposed framework in this study that is hoped will help institutions of higher learning in managing their research data, particularly in the South African context, and perhaps elsewhere.

2.6 Chapter summary

Several models related to RDM were reviewed of which the Community Capability Model Framework (CCMF) and Digital Curation Centre (DCC) Lifecycle Model are most relevant and applicable in the context of understanding research data management services in selected public universities in the South African context. These two models have been found to have some commonalities, differences, gaps and complementary elements. The study revealed that the inherent shortcomings of each model could be complemented by using them jointly. The first model (CCMF) contains the elements of collaboration; skills and training; openness; technical infrastructure; common practices; economic and business models; legal and ethical issues; and academic culture. The second model (DCC) strongly emphasises the curation and preservation of data. This study has therefore found these two capability maturities models' elements potentially useful with regard to investigating the current research data management services, activities, and endorsements of tool adoption, competency requirements. On the other hand, the employed models have significantly shed light on the areas of support and development in the selected public universities in the country, South Africa. Again, these models were also used to aid in meeting the set objectives of this study, particularly those related to the development strategies including policies, infrastructure and capacity building and the one on competency requirements of the staff responsible for the provision of RDM as well as for the researchers. Despite their relevance, application and significant role in the current research, these models were also presented with some shortcomings.

Firstly, CCMF has been identified to be lacking the element of data curation. This gap was complemented by employing the DCC model, which touches on the issues of curating and preservation of data. Closing such a gap was crucial in facilitating easy retrieval of data for future research and/or any other use in the university context. On the other hand, it was significant to understand how the public universities curate and preserve their research data, as it is contributing to the value of digital research data and its lifecycle. Moreover, understanding data curation processes benefits researchers embarking on data discovery and analysis at the university.

Secondly, the DCC model has also been silent on the issues of policies, infrastructure, skills and training, which was bridged through employing the CCMF as covering all of these elements in its eight capability factors. Policies and infrastructure help to properly manage data for a better understanding of its life cycle, availability, accessibility, security, sharing, use and re-use purposes. Moreover, understanding a set of requisite competencies significantly aids in the best performance of staff responsible for RDM, which supports the university's growth and productivity in terms of research.

The next chapter presents the review of literature on the subject of RDM.

CHAPTER THREE: LITERATURE REVIEW

3.1 Introduction

This chapter presents the review of related literature to understand and lay a strong foundation on the status of research data management services of the public universities. The purpose of undertaking research is not to re-invent the wheel, instead it is to identify underlying gaps in literature and solicit suggestions for further areas of research in order to provide solutions to existing problems, as was the case with the present study. So, reviewing relevant literature plays a major role in determining what has been already covered in a particular area of research in relation to an area of interest to avoid redundancy. It is believed that through conducting a literature search, a researcher may become aware of the gaps and inconsistencies that necessitate further research (Welman, Kruger & Mitchell, 2005: 38). For Creswell and Creswell (2018: 26) reviewing literature significantly provides a framework for establishing the importance of the study and benchmark for comparing with other findings of related research.

The present study has reviewed global literature on research data management (RDM) with specific focus to higher education institutions and libraries as the current study's areas of interest. However, the literature on RDM seems to be scarce in the developing nations like South Africa compared to developed countries like the United Kingdom, the United States of America, Canada and Australia to mention a few. Similarly, Kwanya (2021: 5) attests in his "Publishing Trends on Research Data Management in Sub-Saharan Africa: A bibliometrics Analysis" that the production of scientific publications on research data management in Sub-Saharan Africa, like that of the rest of the world, is lower. As a result, the current study has significantly added to the South African literature on RDM. The findings from various cited papers, reports, and other scholarly publications consulted were compared, which was potentially useful in identifying gaps and controversies in order to see and suggest future areas of research to improve good RDM practices and research practices, primarily in the LIS community of researchers.

The review of literature for the present study has covered four broad areas following the set objectives of the study. The first section addressed the RDM services of both universities and libraries. The second covered the competency requirements of

researchers and staff who are responsible for the provision of research data management services in universities. The third section addressed the RDM development strategies through focusing on policy, infrastructure, investment and capacity building programmes. The fourth section addressed the challenges encountered by the public universities in managing their research data. Finally, several RDM frameworks were discussed as a way of paving a path to the development of the RDM framework for public universities.

3.2 RDM services in libraries and HIEs- Global perspective

The global literature provides concrete evidence that RDM is a growing practice for most organisations, including higher education institutions and libraries. This concurs with the observation by Jones, Pryor and Whyte (2013: 2) and Gunjal and Gaitanou (2017: 5) as indicating that many higher education institutions (HEIs) around the globe, in recognition of RDM services, are currently engaged in RDM activities and practices. For example, Ocholla and Ocholla (2020: 360) recently reported on several current trends in academic libraries in South Africa's 26 public universities in terms of providing research support, including research data services/management (RDS; RDM). Though RDM is a global practice, developed nations like the United States of America, United Kingdom, Canada, Australia to mention some, are in a good phase in terms of RDM services as they have adequate resources and proper infrastructure in place to support RDM practices and activities compared to developing nations of the world. Woeber (2017: 9) describes the UK and Australia as the most mature nations in terms of the complete level of service development in RDM. This is evident as the management of research data dates some years back in this country (UK) as Ng'eno and Mutula (2018: 37) acknowledge the establishment of the UK's Data Archives over forty years back, particularly for managing surveys that were paper based. The management of such research data and other data harvests is noted to have been managed on a large scale in this country (UK). On the other hand, the great uptake to support RDM services and activities has been a success in the UK through funding provided by the country's government to e-Science Core Program which is said to be facilitated by the Engineering and Physical Sciences Research Council (EPSRC). Remarkably, UK, through government funding, the country's Research Councils were supported for them to

develop infrastructure, middleware and documentation to support RDM (Ng'eno and Mutula, 2018: 37). Again, in the UK, the Joint Information Systems Committee (JISC) together with e-Science Core Program are noted as significant in the digital curation centre in terms of the development of the RDM with respect to capacity and proficiency. Additionally, Lyon *et al*. (2012: 11) note Digital Curation Centre (DCC) as have been developed in the UK and is a centre in line with the tools like DRAMBORA, Digital Asset Framework (DAF) and CARDIO to support research data activities and could be applied in various levels like project, department, institution, etc. DRAMBORA, DAF and CARDIO are significant frameworks that potentially support the management of research data in an institution. Such numerous afore discussed and highlighted establishments, projects and other facilities in the UK form a fully developed and significant movement for the provision of RDM services in this country.

On the university perspective, most UK universities have been marked with a significant development in terms of RDM. Jones, Pryor and Whyte (2013) note that "Over 40 United Kingdom universities have been involved in developing RDM services within the Jisc MRD programmes and DCC Institutional Engagements". Gunjal and Gaitanou (2017: 5) note that the University of Edinburgh in the UK has implemented RDM services successfully. UK universities are noted to be supported by the Digital Curation Centre (DCC) in the management and exposure of their research data to increase visibility (Davidson *et al*, 2014: 1). Davidson *et al* further highlight the two main services of DCC as DMPonline and the UK pilot research data registry and discovery service (RDRDS). Moreover, the three authors indicate that the DMPonline is significantly used for identifying useful research data while planning for its long-term retention as well as re-use; whereas RDRDS helps in the visibility of research data to promote access and re-use in higher education institutions in this nation (UK) (Davidson *et al*, 2014: 1).

The United States of America (USA) is another developed nation with a significant progress in the provision of RDM. Ng'eno and Mutula (2018: 37) citing Van den Eynden *et al*. observed the USA as have formulated the National Science and Technology Council Committee (NSTCC) as well as the e-Infrastructure Reflection Group in the European Union to respond to the management of data with specific focus on capabilities, capacity and infrastructure as a result of the development

brought by digital research. Given the available programs, projects and agencies in USA, Ng'eno and Mutula (2018: 38) note that the USA government has laid down some strategies for creating "a comprehensive framework of transparent, evolvable and extensible policies, infrastructure, management and organisational structures that provide reliable and effective access, re-use and sharing of research data". Gunjal and Gaitanou (2017: 5) have also observed the Research Data Management System (RDMS) as it had been implemented by the Tufts University in the United States. Moreover, Gunjal and Gaitanou (2017: 5) allude to the fact that the University of Virginia, Cornell University and Princeton University in the US are significantly providing roadmaps, policies, life cycles and a vast array of RDM aspects.

Canada, as another developed country, has been viewed by Miller (2016: 14) citing Shear as lagging in terms of the adoption of RDM strategy. Shearer (2015: 35) observed this nation to have held several consultations and meetings over the past fifteen years about the state of research data management in that country, where some propositions about different solutions were addressed. Some of the meetings and discussions include the National Data Archive Consultation that was held in 2002, which was concerned with the adoption of the national data archive service to collect and preserve research data produced in that country. Another gathering was the National Consultation on Access to Scientific Research Data (NCASRD) which was held in 2004 and sought to address the matters of data access in the fields of physical and life sciences. Again, the Research Data Strategy Working Group was another meeting that was put in place in the year 2008 to help improve the status of RDM in that country. Lately, it is worth noting that the Canadian nation RDM services are supported through three key agencies known as 'Tri-Agencies' which are funded by the government of that country. The government of Canada cited by Ng'eno and Mutula (2018: 38) highlights the three agencies as the Canadian Institutes of Health Research; Natural Sciences and Engineering Research Council of Canada; and Social Sciences and Humanities Research Council (SSHRC). Miller (2016: 14) observes the University of Alberta and University of British Columbia (UBC) as the spearhead of the RDM implementation in institutions of higher learning in Canada.

Among other developed nations of the world, Australia is noted to have an independent initiative known as the Australian National Data Service (ANDS) in response to digital research data (Lyon *et al*., 2012: 10). ANDS is basically focusing

on the management of research data in higher education institutions with a specific focus on sub-groups like senior management, IT support, researcher support as well as the library. In the context of the current study, technicians from the IT support, and librarians from the library section participated in this study among other study participants. It is noted by ANDS Guide (2018: 14) that “ANDS is a partnership led by Monash University in collaboration with the Australian National University (ANU) and the Commonwealth Scientific and Industrial Research Organisation (CSIRO)”.

India as a developing country among nations has been marked as having made good progress towards the services of RDM. KRISHI Research Data Management is a knowledge-based resources information systems hub for innovations in the field of agriculture and an initiative of Indian Council of Agricultural Research (ICAR) which Gunjal and Gaitanou (2017) observe as indeed making a great effort in contributing towards RDM facilities in this country. Such significant progress in RDM by India give evidence that a country does not need to be developed to implement and offer research data management services successfully, whether it be in a university context, library setting or other research centres.

Unfortunately, it is evident in the literature that RDM services are still at an infancy stage in most developing countries as they are still at an implementation stage of the process to support RDM services and activities in different organisations; institutions of higher learning, data centres, and libraries included. Ocholla and Ocholla (2020: 360), have noted RDM as one of the lowest-performing areas in academic libraries on the studied 26 public universities in South Africa. This is supported by Nhendodzashe and Pasipamire (2017: 2) pointing out that “offering research data management services is an emerging trend in most academic institutions in developing countries”, including South Africa. This concurs with Ng’eno and Mutula (2018: 28) who have also indicated that “RDM has been widely embraced in developed countries”, however, Africa is lagging. Similarly, Mushi, Pienaar and van Deventer (2020: 1) also agree that “extensive RDM is not so common in developing countries”, noting that most African higher education institutions and other research institutions have yet to implement data management services. In the study by Cox and Penfield (2014), it has been revealed that libraries were offering limited research

data management services on the participated libraries. On the same note, Patterton (2016) admitted that RDM practices performed by emerging researchers in the Council for Scientific Industrial Research (CSIR) needed improvement in the South African context. Therefore, the current research has contributed to the body of knowledge about RDM in the African literature and has shed light on the importance of embracing the culture of RDM practices, research growth and development in the public universities as they are constantly dealing with abundant research data.

For a successful provision of RDM services in an institution, it is deemed necessary to have in place sufficient resources/infrastructure, skilled labour, policies and other development strategies. Regrettably, Kwanya (2021: 5) observes low publishing trends in Sub-Saharan Africa on research data management due to factors such as insufficient research funding and infrastructure. Anduvare, (2019: 64), too, observed financial constraints as hindering most African countries including South Africa in providing RDM services. Nevertheless, financial limitations have also been noted to be the common challenge, even in developed nations of the world. While most developing countries have been noted to be still battling with the provision of the RDM services and still at the developmental stages, Van Deventer and Pienaar as cited by Ng'eno and Mutula (2018: 39) recognised South Africa among African nations as leading in terms of RDM practices. This concurs with Avuglah (2016: 22) as acknowledging the South African distinct leadership in RDM practices in the higher education sector. South Africa has established some research initiatives, research councils, repositories, research centres and other research-related facilities or units including the South African Data Archive (SADA); Council for Scientific and Industrial Research (CSIR); National Park, National Health Information Repository and Data Warehouse, and Data Intensive Research Initiative of South Africa (DIRISA) among others to support the provision of the RDM services in the country. On the other hand, Koopman and Jager (2016: 2) mark South Africa as having some initiatives for archiving data that date some years back, like the Southern African Data Centre for Oceanography (SADCO) established in the 1960s; AfrOBIS; the African component of the international Ocean Biogeographic Information System (OBIS) established in 1997; the South African Bird Ringing Unit (SAFRING) set in 1948; and the South African Environmental Observation Network (SAEON), established in 2002.

Complications in the management of research data have also been evident in Kenya as another developing country in the African context. Ng'eno and Mutula (2018: 39) note a lack of capacity, resources, collection facilities, analysis, use and re-use in terms of RDM in this country (Kenya). It is, however, promising that Kenya has organisations like the World Agroforestry Centre which plays a crucial role in providing RDM training for agroforestry research scientists in support of human capabilities as well as technical infrastructure for managing data in that country. Again, in the African context, some RDM glitches have been observed in Zimbabwe, particularly at the University of Zimbabwe library which was noted to be halfway ready as it was depicted with a lack of RDM legal obligations and skills to support RDM practices and activities (Nhendodzashe and Pasipamire, 2017: 10). In another developing African country, Avuglah and Underwood (2019: 13) noted that RDM was still emerging at the University of Ghana. Nonetheless, Avuglah and Underwood (2019) have also noted a great potential for significance growth in RDM in this university. Chiware and Becker (2018) found most of the institutions from surveyed countries (Botswana, Lesotho, Malawi, Namibia, Swaziland, Zambia, South Africa and Zimbabwe) not to be fully ready to provide RDM services, noting resources, infrastructure and human capacity to be the main constraints. Their study also revealed countries like Namibia, Malawi, Zambia, Lesotho and Swaziland, as having no RDM-related activities in their universities, research libraries, as well as the national level.

Regardless of the shortcomings experienced in both developed and developing countries of the word, it is evident in the subject literature that there are numerous RDM services and activities in the institutions of higher learning, libraries and other research centres. Flores *et al*., (2015: 91) highlight Access control; Data citation; Data management planning; Intellectual property and copyright; Privacy and confidentiality; RDM workshops; scholarly output as some of the RDM services that could be offered in an academic institution. Avuglah (2016), acknowledges RDM-related activities, such as support for collaborative research; support for data analysis and computational sciences; guidance on RDM, grant applications; and support for storage and high-speed connectivity to facilitate the research at the University of Ghana. Chiware and Becker (2018: 2) highlight the role of academic libraries in enhancing RDM services such as data management planning tool, data

storage and preservation for researchers. Chiware and Becker (2018: 4) citing Cox and Verbaan, postulate that, for successful RDM services, broader institutional approaches are required as researchers generate and use a large quantity of datasets. This marks a pressing need for proper management of research data in response to the ever-changing scientific research landscape, where data is abundant. Chawinga (2019: ii) recognises that participating institutions in the study were actively involved in research activities, leading to intensive research data. Bryant, Lavoie and Malpas (2017: 7) highlight education, expertise and curation as research data management service categories. The three authors go on to indicate that the education service category provides education to both researchers and other stakeholders on the relevance of managing their research data and being responsible for the purposes of long-term curation. This service is also viewed as significantly raising awareness on how to avail research data for future use, which is of importance in documenting and replicating data as well as publishing the outcomes of research. Bryant, Lavoie and Malpas (2017) further note education service as facilitating the policies associated with data management in an institution. With respect to expert service, Bryant, Lavoie and Malpas (2017: 8) assert that it "provides decision support and customized solutions for researchers working through specific research data management problems". On the other hand, the three authors acknowledge that training programs for equipping the internal workforce with necessary skills sets to help in the management of data in institutions to be an example of the expertise service. In respect of curation services, the management of the data through research life cycle is supported by technical infrastructure and other related activities provided through this service (Bryant, Lavoie and Malpas, 2017: 8). Tang and Hu (2019: 9), in their study, indicate that most frequently offered services include RDM planning (81.0%) and data sharing and dissemination (77.8%). Tang and Hu's study also found data preservation; data discovery and access; data visualisation; data organisation and curation and protocol documentation as other services provided by participating institutions. Maurya and Subaveerapandiyan (2022: 263) also highlight data preservation, data discovery and access, data sharing and dissemination, metadata and data visualisation as RDM services expected to be provide in libraries. Remarkably, most organisations do have RDM services that are currently in place, but their successful use by researchers, particularly academics, can never be guaranteed without the required set of

knowledge and skills for carrying out RDM activities. The current study significantly throws light on the competency requirements along with training needs of academic researchers at public universities.

3.3 Competency requirements of researchers and staff responsible for RDM

In today's global and competitive business environment marked by rapidly changing information, knowledge and technology; many organisations are finding it difficult to determine whether their employees are always having current and relevant capabilities for success while meeting the demands of their user community. In this sense, employees should always be equipped with requisite competencies that are current and relevant to create vibrant and relevant organisations. In this regard and context, a university needs to have skilled staff, alongside adequate infrastructure, proper resources, adequate funding and policies, for research data to be properly managed for the purposes of use, re-use, validation and evidence.

3.3.1 Skills and knowledge requirements

Generally, a range of skills coupled with knowledge is really a requirement for a workforce in any workplace. Wilkinson *et al*. (2016: 34) believe that university staff need to be upskilled in all facets of RDM for them to take any RMD-related roles. Wilkinson *et al*. (2016) further state that staff need to engage in training, either delivered in-house or online. For an individual, skills and knowledge are significant as they allow a satisfactory work performance, hence allowing individuals to perform to the best of their abilities, which significantly supports a company's growth, productivity and competition. The current study focused on a set of capabilities required for both researchers and staff responsible for RDM service support, who are constantly confronted with an abundance of research data that must be managed properly.

Several competencies required for better practices of RDM in universities and libraries have been identified by numerous scholars and studies. Skills and training are one of the eight capability factors in the employed CCMF model which Lyon *et al*. (2012: 25) observe as a major determinant to a community to perform data-intensive research. Cox *et al*. (2019: 1434) indicate ongoing staff skill limitations, highlighting

curation skills as among skills required to properly support RDM. Masinde, Chen, and Muthee's (2021) findings also reveal that most respondents had little knowledge of specific data curation activities, and this study recommended immediate data curation training, among other training needs. A need for data curation skills is also evident in the study by Tripathi, Shukla, and Sonkar (2017: 419), who advise that library professionals and researchers should develop digital curation skills. Woeber (2017: 8), too, recognised a need for data curation capabilities and the background of the research environment. Chiware and Becker (2018: 9) agree that the most important skills for the library workforce responsible for providing RDM services in libraries are digital/data curation and IT technology. Similarly, Nhendodzashe and Pasipamire (2017: 7) state that librarians must have digital curation skills in order to support RDM services. Cox *et al.* (2019: 1443) also mention some skills and knowledge that must be developed in staff to support RDM, such as data curation skills, among other types of skills and knowledge.

Apart from curation skills, which are overemphasised by most authors as a requirement for RDM services, Chiware and Becker (2018: 9) identified digital preservation skills and librarians' understanding of identifying and applying appropriate metadata standards as other most important competencies to support RDM services. On the other hand, Anduvare (2019: 45) observed a requirement for project management and communication planning skills to be possessed by librarians in support of RDM. Also noted in Anduvare's study was a need for a strategic initiative such as data librarianship as far as RDM growth is concerned. According to the Unal *et al*. (2019) study, the researchers required training in areas such as data management plans, metadata, dataset version control, consistent file naming, and data citation. Mavodza's (2022) study also confirms data citation abilities. According to Penfield, Cox, and Smith (2014: 28), respondents in their study felt a need for technical, advocacy, liaison, and training skills for staff responsible for providing RDM services. Nhendodzashe and Pasipamire (2017: 7) agree that librarians must have ICT skills as well as advocacy and training skills in order to support RDM. Moreover, Nhendodzashe and Pasipamire (2017: 7) point to a requirement for an understanding or knowledge of the research process, metadata description, RDM workflows and systems as well as data recruitment and different types of research. Tang and Hu (2019: 14) found data/file documentation, metadata

and data management plans to be the most essential knowledge and skills to be possessed by librarians and researchers as far as RDM is concerned. Carlson *et al*, (2011) highlights databases and data formats; discovery and acquisition of data; data management and organisation; data conversion and interoperability; and quality assurance, as core competencies for data information literacy.

On the other hand, Mavodza (2022: 23) observed a need to address data literacy skills among academic librarians and potential data consumers to enhance research data benefits. On this note, Mavodza (2022) points to an ability to recognise the existence and need for data, to applying recognised ways of citing it. Hamad, Al-Fadel and Al-Soub (2021: 77) require researchers to possess the ability to collect, analyse, share, and effectively manage and preserve their research data. The DCC (2015: 3) necessitated a need for a wide variety of capabilities to ensure a proper management of research data and its accessibility to its users from institutions. Molloy and Snow (2012: 251) acknowledge data skills as core academic competency that forms part of the RDM practices and services. Molloy and Snow (2012) note that it is significant to integrate the issue of data handling in the curriculum in HEIs. Kahn *et al*. (2014: 302) agree that there should be an inclusion of the skillsets specific to RDM in existing programmes wherever appropriate. This was raised in the second Research Data Management Forum in November 2008 in the United Kingdom (Kahn *et al*. 2014: 302). Lyon *et al*. (2012) identifies tools and technologies; data description and identification; and policy and planning as skillsets to be possessed for the support of RDM services. On the other hand, Molloy and Snow (2012: 251) highlight sustained awareness of the creation, organisation, validation, sharing and curation of data as necessary skills to be acquired in support for RDM services provision. Jones, Pryor and Whyte (2013: 2) are of the view that the range of skills and knowledge needed to deliver RDM services is determined largely by the individual phases of the research project life cycle, from pre-award, and throughout the project to post-project. For Qin (2013: 215), the research life cycle and processes involve demand management and communication skills from the personnel involved. Shearer (2015: 35) realises a need for improvement in information skills to support the provision of RDM services in universities. Shearer (2015: 35), discussing the German Rectors Conference, an association of 268 universities in Germany in 2013,

indicated that one of the resolutions about research data management in universities is that universities should improve information skills.

Though there are countless skills and knowledge required for both staff and researchers regarding the services of RDM in different organisations, universities included, Cox *et al.* (2019) reveals that skills shortage remains the challenge to most staff working in academic libraries responsible for the provision of RDM services. Similarly, Chiware and Mathe (2016: 1) point out inadequate data management skills among library staff as well as the paucity of resources and infrastructures as the challenges of RDM service provision in libraries. Several studies (Chawinga & Zinn, 2020; Cox *et al*. 2019, Chawinga 2019, Avuglah & Underwood 2019, Ng'eno & Mutula 2018, Cox *et al*. 2017, Pinfield, Chiware & Mathe 2016, Cox & Penfield 2014, Cox & Smith 2014, Tenopir *et al*. 2014) attest to a paucity of staff skills and knowledge in RDM support. Considering the skills and knowledge gap, Avuglah and Underwood (2019) challenge library schools to update their graduate curricula for the benefit of the new and emerging roles of librarians in the data-driven research arena. Lyon (2012) highlights some important skill sets for researchers in CCMF as one of the models used in the current research (refer to section 2.2.1.2, Chapter two).

3.3.2 Training needs and capability development

An organisation like a university could offer a wide range of professional development opportunities for skilling and re-skilling individuals or its workforce. Some development opportunities can be achieved through embedding them in the curriculum, offering them as standalone courses, workshops, in-house training and other conduct to keep up with the demands of the gradually changing world and its latest developments. In Chiware and Becker's (2018) study, workshops and independent study were viewed as the ways of promoting RDM training services for librarians (Chiware and Becker, 2018: 10). Other than workshops and independent study, Chiware and Becker (2018) also acknowledged conference attendance, training provided by professional bodies, short courses in digital technology, hired consultations, training from vendors as well as from internal and external colleagues as platforms for promoting the RDM skills development. Wachtler *et al*. (2021: 12) observed that that the training in RDM is either not provided at all or carried out in an *ad hoc* manner in NFDI-Neuro, Germany. In this regard, Wachtler *et al*. (2021: 12)

propose educating researchers as a key component for improving RDM, observing that developing and implementing a coordinated training concept would be a major task in promoting knowledge and competence in RDM throughout the neuroscience community. Bunkar and Bhatt (2020: 144) also recommend that library professionals receive intensive training in order to establish effective Research Data Management Systems. Kwanya (2021: 9) also suggests appropriate training and publicity programs for Sub-Saharan African researchers interested in adopting the RDM concept. Bunkar and Butt (2020: 144) also advocate intensive library professional training. Relevance of training is also central in the study by Masinde, Chen and Muthee (2021). Furthermore, Wiljes and Cimiano (2019: 7) suggest training in research data management to be an essential part of academic education. Chigwada, Chiparausha and Kasiroori (2017: 3) citing Tenopir *et al*. (2016) and Tenopir *et al*. (2012) also complemented skills development through attending conferences, workshops, some courses on research data, professional development working groups and internal workshops as well as presentations to equip staff responsible for RDM services and activities. However, Unal *et al*. (2019) found a lack of formal training in RDM among researchers at the Manchester Metropolitan University. They further note that the researchers had a need for training in areas like DMP, metadata, version control of datasets, consistent file naming and data citation. It is promising that the University of Edinburgh has a training program on RDM called MANTRA, which is an online course that is free, specifically for those who manage digital data as part of their research project at the University of Edinburgh in the UK (University of Edinburgh website, 2019). Van Wyk (2017) also noted MANTRA to be a course established at the University of Pretoria (UP) in South Africa. Other than MANTRA, Van Wyk (2017) further identifies a subject guide which is known to play a pivotal role in providing links to the RDM policy of this university (UP) as well as some important manuals that facilitate tools, practices and templates on RDM. On the other hand, it is worth noting that Chiware and Mathe (2016: 9) have noted the Cape Peninsula University of Technology (CPUT) libraries have established a skills development plan to support faculty or subject librarians in participating in RDM services which is the plan facilitated by internal staff and external service providers. Equally, Chiware and Mathe (2016: 9) have identified some areas of focus from the CPUT skills development programme as follows:

- subject-specific orientation.

- understanding of institutional research processes and policies.
- knowledge of relevant tools (such as Mendeley, Scopus author management, Data Management Plan Tool, ORCID identifier.
- communication and collaboration (video, web-conferencing technologies).
- current awareness-alerts and RSS services.
- research methodologies (for example, research data lifecycle, data analysis, tools, statistics).
- bibliometrics (for example, impact factor, h-index) and altmetrics.
- Intellectual Property rights.
- publication processes and requirements.
- awareness of DOAG.org and Sherpa Romeo; and
- academic networking

As far as training needs and capability development is concerned, Davidson *et al*. (2014: 251) observe the development of digital curation and RDM training materials in an attempt to address a shortage of skills from researchers and staff responsible for RDM such as support staff. The authors also indicate that the UK Data Archive has tried through providing wide guidance and training materials with specific focus on creation, management, as well as the sharing of research data. Also, Lyon *et al*. (2012: 25) remark the development of community abilities to be encouraged through the provision of training in relevant skills sets. Avuglah (2016: 90), too, notes a relevance of providing RDM skills development for staff, despite them having some idea of what RDM is, because that seems insufficient to providing RDM services in an institution. In addition, considering the skills gap as one of the most cited problems in literature among others, Kahn *et al*. (2014: 304) believe that it could be addressed through training and development, noting the EC-funded FOSTER project and organisations like NeDICC and LIASA to play a role in providing awareness, materials for training, offering practical skills and training itself. Cox *et al.* (2019: 1443) note some skills and knowledge as requiring development among staff to support RDM, and these include data curation skills; knowledge of a variety of research methods; data description and documentation; legal, policy and advisory skills; technical and ICT skills; data storage and infrastructure included understanding of research integrity, reproducibility and transparency principles;

knowledge of the research lifecycle; and subject and or disciplinary knowledge. The necessity of staff development was also acknowledged by Nhendodzashe and Pasipamire (2017: 10) at the University of Zimbabwe, since it was planning to implement a project to support RDM services in that institution. Tang and Hu suggest advanced data management skills and data management tools and infrastructure as areas that required further training RDM (Tang and Hu, 2019: 14). In addition, Tang and Hu's study also revealed a strong need for data services skills training for librarians and data management/data management processing skills for researchers. The CGIAR Internal Audit Unit (2017: 28) postulates that training, awareness as well as continued support could be ensured by effectively implementing an institutional research data management framework that will oversee adequate training to all institutional research staff in terms of policies, procedures and practices aligned to research data management. The CGIAR goes on to indicate that newly appointed research staff could receive necessary skills on RDM through embedding relevant information on RDM to induction training programs. On the other hand, the CGIAR notes refresher training and inductions as some ways to reskill already employed staff. Again, training materials and self-read materials should also be taken into consideration for providing training and support in RDM (CGIAR Internal Audit Unit, 2017: 28). Most authors attest to the relevance of providing training for skilling and re-skilling of staff and researchers to facilitate better practices of RDM in higher education institutions and libraries.

3.4 RDM development strategies

If there is no strategy in place for managing research data in an institution, it is like a ship without a rudder. In this respect, it is considered necessary to have policies along with strategies, infrastructure, adequate resources and programmes in place to properly manage data for a better understanding of its lifecycle, availability, accessibility, security, use and re-use purposes in the university context as the focus of the current research.

3.4.1 RDM policies and procedures in the global context

It is critical to have RDM policies in place or to develop and implement RDM policies in organisations such as libraries (Chiware and Becker, 2018: 6). This is because RDM policies aid in the delivery of services by ensuring conformity among

researchers and the provision of uniform services by responsible personnel such as library professionals and those in the IT section (Chiware and Becker, 2018: 6). In the context of the current research, library professionals were interviewed (librarians and technicians) as study participants among other groups of respondents who participated in this study. Shearer (2015: 8) indicates that RDM policy can also indicate the necessity for the deposit of metadata into a local or national catalogue. Avuglah (2016: 32) identifies two crucial RDM activities to be the strategy and policy development which he observes as significant in helping to achieve coherent data service development in an institution. Jones, Pryor and Whyte (2013: 7) put forth some key actions to be taken into consideration when developing an RDM strategy and policy, as follows:

- One needs to understand his/her current position and where he/she wants to be to define the strategy.
- Map out a programme of activity to deliver infrastructure and services.
- Draft a policy based on external drivers and local context to establish core RDM principles.
- Ratify the policy then undertake advocacy work and pilot studies to aid implementation, and.
- Consult broadly and iteratively throughout to gain consensus and secure support.

Whyte and Tedds (2011) highlight public good, planning for preservation, discovery, confidentiality, first use, recognition and public funding as commonalities in principles of data policy of the UK's research councils.

Table 3.1 below is Shearer's (2015) common elements of an RDM policy between institutions:

Table 3. 1: Common elements of RDM policies

Policy requirements	
Data quality and standards	In order to enable access and reuse, investigators should stick to international standards. For data to be understood by others, it should be complemented by data documentation and metadata.
Data access and sharing	Investigators should be accountable for data availability to enhance sharing more especially during results publication.
Data retention and preservation	Data needs to be retained for a particular minimum time period. Deposition of data in a long-term archive from

	investigators is encouraged to enhance long term preservation of data.
Data management plans	A Data Management Plan (DMP) should be a requirement in research proposals.
Common provisions to policies	
Privacy	The protection of the rights and privacy of individuals who participate in research should always be ensured. Thus, data made available for broader use should be free of identifiers that would permit links to individual research participants and variables that could lead to deductive disclosure of the identity of individual subjects.
Traditional knowledge	Where local and traditional knowledge is concerned, rights of the knowledge holders must not be compromised.
Data of a sensitive nature	Where data release may cause harm, specific aspects of the data may need to be protected.
Intellectual property/Data ownership	It may be necessary on occasion to delay publication for a short period to allow time for applications to be drafted.
Other aspects	
Principles	Data policies adhere to a set of overarching principles that articulate their value.
Scope/Coverage of Policy	Describe the scope of data covered by the policy.
Roles and responsibilities	The policy identifies the various parties responsible for managing data across the different stages of the lifecycle
Monitoring and enforcement	The means by which policies will be monitored or enforced are outlined in the policy.

Source: Shearer (2015: 8)

Chigwada, Chiparausha and Kasiroori (2017: 3) assert that RDM policy is potentially useful as it helps in answering claims of misconduct in research studies and assists in the guarding of intellectual property. Shearer (2015: 36) believes that RDM policies cannot be adopted in isolation; therefore, sound research data management practices depend on multiple factors including incentives, skills/expertise, services, infrastructure, funding and procedures. Such factors are trusted to create a background that supports RDM in an institution, a university in this context. Though RDM policies are perceived as significant for the provision of RDM services in libraries, universities and other research-related centres have some deficiencies associated with them. For example, the study by Ng'eno (2018: 28) brought into attention that "RDM legal framework did not exist in the institutes surveyed; the RDM policies and regulations were outdated; there was limited RDM awareness and advocacy; the institutes lacked RDM security systems; the institutes suffered from

lack of or inadequate RDM guidelines on standardisation and technical infrastructure". In this sense, Kahn *et al*. (2014: 304) emphasise a need for RDM policy to be developed by more than one stakeholder and this could include government, publishers, information professionals and researchers, among others.

3.4.2 Universities' RDM policies

RDM policies are still in the early stages in many countries, and most universities have either full or partial solutions in place (Liu, Zotoo, & Su, 2020: 21). Promisingly, some universities, such as the University of Southampton, Open University, University of Twente, Oxford Brookes, University of Cape Town, University of Pretoria, University of the Western Cape, Stellenbosch University, University of South Africa, Ethiopian Institute of Agricultural Research, and others, have RDM policies in place that govern the management of their research data. However, some universities, such as Charles Darwin University and Durban University of Technology, use RDM procedures instead. According to Unal *et al* (2019), the United Kingdom is the most advanced nation in terms of RDM research and development of technologies, tools, and policies.

The University of Southampton in the UK has a Research Data Management Policy which was approved by Senate in February 2012 (University of Southampton Research Data Management Policy, 2019: 7). This university's policy went for review in July 2014; there were no changes made and it was once again reviewed in July 2015; again, with no changes. The review was done again in June 2016 with no changes. Recently, the review of the policy was done in May 2019 and the Senate approved it in June 2019. The University of Southampton Research Data Management Policy (2019: 2) is aimed at creating a model research data management practice for all researchers at the University of Southampton as well as ensuring the proper and secure storage, retention, accessibility and disposal of research data in line with legal, statutory, ethical, contractual and funding requirement in place. The University of Southampton Research Data Management Policy indicates a minimum of 10 years for research data from its collection, creation or publication of results in accordance with section 6 of the policy. This policy further stipulates a retention period of research data to be 10 years or longer in the case of a public interest or heritage-based value.

The Open University is another university in the UK having an RDM policy. Crane (2018: 1) highlights that the Open University (OU) located in the UK has an RDM policy that helps “define the standards that govern the management of research data at the Open University. It defines research data as the evidence base on which academic researchers build their analytic or other work and it applies to all engaged in research at the OU, most typically research staff and research students.” In terms of availability of research data, Crane (2018: 6) stipulates that there should be available data that supports published research data and should be available on the first date of online publication. Crane further explains that published data should have clear conditions regarding the access of supporting data. It is also stated by Crane that supporting research data could be available as much as it is still useful to the researcher and research community. Furthermore, the author remarks the period for such data to be even longer at up to ten (10) years; however, this could also depend on the funding requirements. Crane additionally indicates that researchers should comply with OU curation policies in terms of retention, reviewers and destruction of the data when storing research data within ORDO.

The University of Twente (UT) in the Netherlands has a Research Data Management policy which was accepted in 2015. However, in 2018, this policy had to be reviewed due to the realisation of handling research data, an increase of electronic data as well as open science and integrity. This then brings us to the current University of Twente Research Data Management policy, which places emphasis on the benefits of RDM, basically for researchers and facilitates guidelines on good RDM at this university. This policy is aligned with the UT privacy policy, privacy statement as well as privacy guidelines for research regarding the privacy regulations in research at the UT. “Selected research data and related materials must be archived at least for ten (10) years, unless legal or contractual regulations demand another term” (UT RDM policy, 2018: 7). The UT RDM policy (2018: 5) stresses the need for a good data management plan (DMP) for every research project. Regarding data storage, the UT RDM policy stipulates that research data collected should be kept in the ISO 27001- and NEN 7510 certified facilities.

Unlike previously discussed universities, Charles Darwin University has Research Data Management Procedures other than the policy itself. This university’s RDM procedural document first draft was approved in 2010. It was revised in 2013, 2014,

2015 and again in 2019, which is the current version of the procedures until the next amendment, scheduled to be made in December 2020. This procedural document helps both the researchers and research administrators in RDM practices and links them to the legal, statutory, ethical and funding body requirements. Also, Charles Darwin University RDM Procedures (2019: 6) indicate a five (5) year period as the minimum retention period after a research project's completion or publication. Charles Darwin's procedural document further highlights the following retention periods in different research types:

- For short-term research projects that are only for assessment purposes, such as research projects completed by students, retaining research data for twelve (12) months after the project's completion is usually sufficient.
- Clinical trials involving humans require that research data be kept for a minimum of fifteen (15) years after publication.
- Research data, such as patient records, must be kept indefinitely in areas such as gene therapy, and.
- If the work has high community or heritage value, the research data should be kept forever.

The Ethiopian Institute of Agricultural Research (EIAR) has an established RDM policy in place. Kirub (2017) asserts that the Ethiopian Institute of Agricultural Research (EIAR) has an RDM policy as a result of the fact that it is conducting an abundant, ever-growing data-oriented agricultural research data that needs to be securely and efficiently collected, managed and shared for researchers and partners in the EIAR. In terms of storage, Kirub (2017: 11) indicates that EIAR RDM policy applies the following to research data across research directorates of EIAR:

- Data stored with appropriate metadata to facilitate re-use.
- Data secure, safe, and backed up regularly (where local storage is chosen, this should be on EIAR networked systems); and
- Data stored in a manner that is compliant with scientific and legal obligations, and the requirements of the program/ commodity.

Additionally, Kirub (2017: 11) notes that research data should be retained for about ten (10) years after the completion of the research project in a correct format as well as the facility; however, that could depend on the nature of the data. Also, Kirub

(2017) explains that destroying research data should be in line with ethical, legal and requirements of research together with EIAR confidentiality and security measures.

As is the case with other institutions in developed countries, not all South Africa's public universities have RDM policies and procedures in place, but they do define their research and ethical policies. Similarly, in the South African context, the University of Cape Town (UCT) has an RDM policy, called the University of Cape Town Research Data Management Policy, which was issued in March 2018 (University of Cape Town Research Data Management Policy, 2018). The DVC (Research & Internationalisation) is the responsible executive for this policy. The policy document states the purpose of the policy as "to transform the way research is conducted at UCT by accelerating discovery, increasing the value of research decision-making, and catalysing changes throughout the economy and society that are of value to all citizens". The UCT RDM policy highlights the following criteria as general guidelines for the selection of research data:

- The data that substantiate published research findings.
- Significant data generated by the project.
- Unrepeatable observations.
- Longitudinal studies of human or natural events.
- Experimental results that would be impossible or expensive to reproduce.

The University of Pretoria (UP) is another South African university with a functional RDM policy. Van Wyk (2017) states that the first draft of the UP RDM policy was proposed in August 2014, then the proposed policy went through several rehearsals before its approval in September 2017 by the UP Senate. The UP RDM policy is aimed at governing the management of research data at the University of Pretoria and ensuring that all research data generated at this University is properly managed and curated effectively and efficiently in order to support the University's positioning as an international research-intensive university (UP RDM policy, 2017: 6). The UP RDM policy also specifies a minimum of ten (10) years for keeping research data sets after the original research project is completed, but there is an involvement of intellectual property. However, in the case of contractual requirements, the policy states a necessity for a longer period. It is further indicated by the UP RDM policy

that funding bodies may require that all raw data be kept indefinitely in the case of research that involves human subjects. UP policy (2017: 6) is aligned with the University of Pretoria Intellectual Property Policy; Contract Research and Consulting Policy; Code of Ethics for Research; Policy and Procedures for Responsible Research; Information Governance Policy Framework; and Information Governance Matrix. UP policy is also aligned to the Promotion of Access to Information Act No. 2 of 2000 and the Protection of Personal Information Act No. 4 of 2013 as government legislation. Regarding ownership, the University of Pretoria RDM policy (2017: 6) states that “All primary research materials and data created, collected and/or generated by students, employees and affiliates of the University belong to the University and must be stored throughout the entire life cycle of a research project, in line with a project Data Management Plan”.

Stellenbosch University also has an RDM policy, which was approved by its Senate in November 2020 and is aimed at managing research data at Stellenbosch University (SU) to ensure compliance with legislative frameworks, as well as protecting the University, staff, and research participants involved in research through risk mitigation and management (Stellenbosch University Research Data Management Regulations, 2020: 1). According to SU RDM policy, research data generated as part of the research process is preserved with appropriate high-quality metadata for its lifecycle. However, unless otherwise specified, the policy states that the default period for research data retention is ten years from the date of the last requested access, publication, or public release.

The University of the Western Cape has an RDM policy known as the University of the Western Cape Research Data Management Policy, which was approved in January 2021 and has a three-year life cycle. The goal of this policy is to make research data management easier at the University of the Western Cape, and to ensure that all research data generated at the University is managed and curated in order to support the institution's goals of becoming a research-intensive university (University of the Western Cape Research Data Management Policy 2021: 3). This policy seeks to promote sound RDM practice throughout the institution by outlining the processes, roles and responsibilities, support services, and facilities available. According to the policy, non-digital research data must be kept securely within the Department, either in the researcher's office or in the laboratory where they were

generated. When data is confidential, it must be appropriately secured in accordance with ethical and legal requirements. In terms of research data storage, the policy states that research datasets must be kept for at least five years after the completion of the original project, and a longer period may be required where intellectual property is involved, or if there are specific statutory or contractual requirements.

Macanda, Rammutloa and Bezuidenhout (2015) report that the University of South Africa (UNISA), as the largest university offering distance learning in Africa, has completed investigation on RDM as part of the plan to establish data management for their research. Currently, UNISA has an RDM policy, which was approved by its Council in August 2021. The goal of this policy is to ensure that research data is stored, preserved, retained, made accessible for use and re-use, and/or disposed of in accordance with the requirements of legal, statutory, ethical, and funding bodies. (Policy on Research Data Management at the University of South Africa, 2021: 2) This policy seeks to ensure consistent research practice in relation to data management principles that support effective data sharing, open access, and the discovery, accessibility, re-usability, and interoperability of data in accordance with quality standards. This policy requires that research data be kept for a period of ten years before being evaluated for further retention. This policy is related to other institutional policies such as the Library Collection Development Policy, the Master's and Doctoral Degree Policy, the Data Privacy Policy, the Information Security Policy, and the Intellectual Property Policy, to name a few.

The Cape Peninsula University of Technology (CPTU) also has an RDM policy that regulates the management of research data in this university. The CPTU RDM policy is like a three-legged stool, as a result of the fact that it consists of culture; infrastructure and services and policies (Chiware, 2015). The CPTU RDM policy is aligned with other institutional policies such as the policy on intellectual property; research policy; contract research policy; institutional operation policy; research management plan; research and innovation and open access policy. The CPUT RDM policy was approved in July 2014. For CPUT to manage research data, it collaborates with other universities with the main goals of setting up international working groups on various RDM development aspects; developing an open-source platform for managing an RDM cycle and to set up institutional working groups to feed into the development of the open-source platform (CPUT Research Data

Management Policy, 2016). The CPUT RDM policy stipulates that data should be retained for at least ten (10) years after a completion of the project or publication; however, the period could be longer depending on the funding period (CPTU RDM policy, 2016: 6).

The Durban University of Technology (DUT) currently does not have an exact RDM policy; instead, it provides guidelines for research data storage to facilitate sound research data storage and maintenance practice. According to these guidelines, numerical or statistical data should be stored in raw data format for five years after the project is completed. Data should be destroyed after five years, unless it is going to be used in a longitudinal study. The data from interview notes/questionnaire responses/transcribed interviews should be stored in their original form for the same amount of time as the numerical and statistical data (five years) after the project is completed, unless the data is still to be used in later longitudinal studies, in which case it may be destroyed after this time. The guidelines highlight that work that informs national policymaking should be archived after ten years. Images/audio and video recordings are typically kept in their original form, which is deemed important where they are later improved. When possible, both original and enhanced images/audio and video recordings should be kept for the same period as the data types after the project is completed. However, if the data is to be used in longitudinal studies, it may be destroyed after this time frame. According to University guidelines, blood and plasma samples should be anonymised, stored for 3-6 months while analysis is performed, and then disposed of in the most appropriate manner. Data for longitudinal studies are kept for a longer period than other types of data. Data collected as part of a known longitudinal study should be kept for the duration of the study and for ten years after completion. Participants should be kept up to date on how long the study is expected to last. If the study is extended, all participants should be contacted and informed that their data is still being saved and may be used.

Most reviewed universities' RDM policies show that they keep, retain, and archive their research data for at least ten (10) years. The most common underlying policies of universities, on the other hand, can be found in the areas of access, retention, sharing, storage, and ownership (Liu, Zotoo and Su, 2020: 32).

3.4.3 Infrastructure and investment for supporting RDM services

A good infrastructure to support RDM services in an organisation like a university in this context could mean an infrastructure that is conducive to its users such as researchers and workforce or even a building on its own. Technical infrastructure is one of the CCMF capability factors and has been covered in section 2.2.1.4 of the dissertation (Chapter two) and Lyon *et al.* (2012: 31) declare that technical infrastructure supporting research should have necessary tools as well as the services in line with different stages of the research life cycle. Again, technical or IT infrastructure forms the basis of the Data Management Framework CMM, which is another model employed in the current research. Technical infrastructure has also been discussed in all the five levels of the Data Management Framework CMM (Chapter two, see section 2.2.3.1- 2.2.3.5) of the dissertation. Normally, companies or organisations make a lot of investments and use their capital to invest in new and existing infrastructure to support the provision of services to their particular user community. In this regard, organisations like universities should first take into consideration the needs of researchers, proper identification and integration of the relevant technologies and capabilities required before investing in infrastructure to support the development of RDM services. As Chiware and Mathe (2016: 5) suggested in their study, the Cape Peninsula University of Technology libraries had run a pilot project with some research groups before attempting to pursue the full-service infrastructure of the university to support RDM. On the other hand, Avuglah (2016: 32) notes several linked nationwide services for supporting RDM services in HIEs and other data centres like UK Research Data Services (UKRDS), UK Data Service, Australia's Research Data Service (RDS) and Data Intensive Research Initiative of South Africa (DIRISA). Therefore, these projects can significantly help universities choose a solid infrastructure to support the provision of RDM services. Infrastructure may vary from institution to institution. Conversely, Shearer (2015: 37) notes divided infrastructure for research data management, where some fields are having very good coverage whereas others enjoy very little in Canada. Shearer goes on to indicate that a suitable coverage is usually provided to those fields in contact with some national repositories like astronomy, ocean science, Statistics Canada Data Centres, polar/arctic research data and genomics. Again, Shearer (2015) has observed portage to be a recent initiative for RDM in Canada, which is managed by the Canadian Association of Research Libraries and is significantly improving on the

issues of infrastructure and services gap. Additionally, Shearer (2015: 37) realised that the national network of expertise and a national preservation and discovery system form the components of the portage. Qin (2013: 216) views technologies; data and metadata standards as well as policies that oversee data management, sharing and use as the three (3) dimensions of the data infrastructure. With regard to technology, Qin (2013) indicates several technologies covered, including those for data collection, storage, processing, organising, transmitting as well as preservation purposes and not omitting the communication and collaboration platforms. For data and metadata standards, she observes grouping of scientific data based on disciplines and type such as physical and chemical; earth and astronomical; and life sciences; whereas with respect to policies, she includes the issues of sharing, intellectual property, ethics and open access for scientific data to cover a wide range of topics with respect to national and global viewpoints (Qin, 2013: 216-218). Using the European context, Whyte as cited by Avuglah (2016: 29) highlights institutional, national and continental as the three (3) levels of infrastructural development. The ANDS guide (2018: 11) highlights that there is a checklist for assessing IT Infrastructure capability for data management which is believed to be basically a starting point for institutions in helping them assess their IT infrastructure capability's maturity. In terms of research data management, the ANDS guide (2018) further notes the checklist as helping to address questions in relation to the storage of data and metadata; identification of management, authentication and access; internal and external network connectivity; access to discipline specific tools to support analysis; software development; visualisation; collaborative environments; and High-performance computing (HPC). Through the evidence from literature, it could be argued that a proper infrastructure should support all the processes of the research lifecycle right from the beginning of the research data management plan to the publication of results.

3.4.3.1 Research data management plan (RDMP)

Generally, a plan is a proposal for achieving something. In the context of the present research, emphasis is placed on the research data management plan (RDMP) which is l a plan that paves a path of necessary steps to be taken before the beginning of any research project. It is actually a key element that sheds light on the entire research project. According to Stellenbosch University Research Data Management

Procedures (2020: 18), a data management plan can assist researchers in documenting every stage of the research data lifecycle. Again, a data management plan is also an important tool for ensuring that researchers are aware of and have a plan in place to comply with policy and regulatory requirements before they begin collecting data. RDMP is closely associated with the preservation planning covered in the employed DCC Curation Lifecycle Model as one of this model's full lifecycle actions (refer to section 2.2.2.1, Chapter three). A data management plan (DMP) significantly portrays representative requirements for a successful research project during its process and after its completion for a researcher (Mushi, Pienaar and Van Deventer, 2020: 3; University of Twente RDM policy, 2018:14; University of Pretoria RDM policy, 2018: 4). Charles Darwin University's RDM Procedures (2019: 4) emphasise the relevance of developing a research data management plan by all researchers and research administrators in the beginning of a research project in order to address the issues of research data in terms of ownership, custodianship retention, storage and disposal. Borghi *et al.* (2018) in their 'Support Your Data: A Research Data Management Guide for Researcher' put forth RDM rubric which shows various RDM-related activities as occurring over the course of a research project with six activities including planning, organising, saving, preparing, analysing and sharing. DCC Checklist (2017) significantly gives an overview of the considerations one needs to make as well as what to include when starting to write a data management plan. Michener (2015: 8) portrays a data management plan (DMP) as a road map that should be easy to follow in guiding and explaining a research project throughout all the processes involved as well as after the completion of the project. Michener (2018) also believes that DMP should determine the research sponsor requirements; identify data to be collected; its organisation; documentation; quality assurance; storage and preservation strategy; policies involved; dissemination of the results; assigning roles and responsibilities; and budget. Corti (2014) suggests some key planning issues including knowing legal, ethical and other obligations towards research participants, colleagues, research funders and institutions; understanding the institution's policies and services: storage and backup strategy, research integrity framework, IPR policy, institutional data repository; assigning roles and responsibilities to relevant parties; incorporating data management into research cycle; and implementing and reviewing management of data during project meetings and review. Mushi, Pienaar and Van Deventer (2020: 3)

also observe some basic information on DMP such as data collection, sharing and preservation. On the other hand, Shearer (2015: 21) suggests that a data management plan should have an assessment of existing data; information on new data; quality assurance of data; backup and security of data; expected difficulties in data sharing; copyright/intellectual property rights; responsibilities and preparation of data for sharing and archiving detailed guidance about preparing data management. Whyte and Tedds (2011) postulate some data management planning approaches like DCC's tool DMP Online to help project teams in creating data management plans prior to and during research projects; DMP-ESRC Project; ERIM Project and HALOGEN Project at the University of Leicester. Jones (2013) proposes a checklist for a data management plan with the following condensed questions to help researchers in developing their research data management plans:

- What data will you collect or create?
- How will the data be collected or created?
- What documentation and metadata will accompany the data?
- How will you manage any ethical issues?
- How will you manage copyright and Intellectual Property Rights (IPR) issues?
- How will the data be stored and backed up during the research?
- How will you manage access and security?
- Which data should be retained, shared, and/or preserved?
- What is the long-term preservation plan for the dataset?
- How will you share the data?
- Are any restrictions on data sharing required?
- Who will be responsible for data management?
- What resources will you require to deliver your plan?

3.4.3.2 Research data storage, organisation and management

Any published data, research data in this context, requires to be kept securely and backed-up to facilitate its ease of access and long-term stability. This therefore brings the recognition that data storage, backup, organisation and management highly depend on the infrastructure in place to support all these activities. Infrastructure has been also addressed in section 2.2.1.4 (Chapter two of the dissertation) as one of the capability factors in the CCMF model employed in the

current research. A proper infrastructure facilitates a sturdy structure and organisation of data so that it may be easily controlled. Research data like any other data type, can be stored in different locations such as a laptop, external storage devices (e.g. USB, external hard drive, CDs), emails, institutional server, clouds, personal website, etc. Though there are numerous storage facilities for research data, it is evident from some studies (Koopman and De Jager, 2016; Chawinga, 2019; Patterton, 2016) that hard drives are a highly favoured location for data storage and backup purposes. This concurs with the findings by Koopman and Jager (2016): they also found a hard drive to be the most used location by 83% of respondents for data back-up, followed by a PC/ laptop (55%) and less in other locations like cloud storage, flash drives, servers and others (Koopman and De Jager, 2016: 5). Similarly, Chawinga (2019) also found most (90.3% and 92.9%) respondents to have previously used the external hard drive as their data backup platform (Chawinga, 2019: 148). Again, Patterton's study (2016) also found external hard drive to be the most common backup location for researchers, with 71% among other backup locations like cloud, USBs, server units, etc. Conversely, Ng'eno (2018: 132) noticed personal computers or laptops to be most used for data storage and backup purposes. According to the findings of Nge'no's (2018) study, most researchers did not use various storage tools such as local servers, central campus servers, web-based servers, data repositories, and automatic server generation. It is observed in the Charles Darwin University RDM Procedures (2019: 7) that there should be a back-up strategy for both researchers and research administrators to recover data in case of loss or to recover data from a particular time period. Again, "backups of more than one copy should be performed regularly and should be housed remotely from the main data storage" (Charles Darwin University RDM Procedures, 2019: 7). Charles Darwin University RDM Procedures (2019) further states that data restoration processes could be ensured through properly labelling and organising backups. For research data storage and management, the University of Southampton Research Data Management Policy (2019: 5) indicates the significance of regular backing up of data in line with stipulated research practice in the designated field of research practice by all researchers in all research data both in digital and computer-readable form.

3.4.3.3 Sharing and use of research data

Data sharing could be regarded as sound behaviour, hence significantly facilitating finding, use and re-use of research data. Issues relating to data sharing and use were also discussed in section 2.2.1.7 (Chapter two) as one of the capability factors in the employed CCMF model which Lyon *et al.* (2012: 40) believes to be concerned with the ethical issues being embedded in data sharing practices that helps in the provision of restrictions for data intensive research in an organisation, a university in this context. Miller (2016: 7) recognises that a library can play a significant role in facilitating data sharing and use processes regardless of the RDM structure present in an institution. Such roles (sharing and use) can be facilitated by institutional repositories which are usually operated by an institution's library. The use and sharing of data in an institution normally operates under some standards, policies and procedures like those related to RDM. In Canada, Shearer (2015: 5) notes an increase in interest and support in open government and open science initiatives by the Canadian federal government to enhance research data management and sharing practices. She then acknowledges the National Institutes of Health (NIH) as one of the first funding agencies to have a policy about research data sharing. Shearer (2015: 8) goes on to acknowledge that, for data access and sharing, an RDM policy could significantly indicate the requirements of the investigators in terms of availing data to be shared; for example, upon publication of results or shortly thereafter. Macdonald and Martinez-Urib cited by Woeber (2017) explain that an Online Research Database Service (ORDS) allows the centralised creation, storage and sharing of humanities datasets. In Tang and Hu (2019) study, data sharing as well as dissemination was found to be the second most (77.78%) provided services by participating institutions (Tang and Hu, 2019: 9). It is also observed by Tang and Hu (2019: 9) that data sharing is crucial in promoting research visibility and output that researchers should consider. Though Tang and Hu (2019: 9) find data sharing to be an area that requires training for researchers. This indicates a significant gap in terms of data sharing for researchers. The study by Woeber (2017: 26) found data sharing to be more common on AHRC-funded projects and among certain disciplines. However, Woeber (2017) found copyright of digital materials to be a major barrier to data sharing in humanities research. Whyte and Tedds (2011) recognise the FISHnet Project as having developed a platform for curating and sharing research data in freshwater biology. Sharing research data is also

acknowledged by Whyte and Tedds (2011) as one of the organisational responsibilities for research data. Unal *et al.* (2019) notes the prominence of sharing and re-use of research data as improving the economic growth and stability in an institution, increasing research visibility, increasing resources, efficiency and other immense benefits relating to open access to research data. Unal *et al.* further marks a substantial move in the UK in terms of management and research data sharing which has been brought by the introduction of some policies like Research Councils UK (RCUK). Chiware and Becker (2018: 8) highlight data sharing or preservation (93%) as one of the main reasons provided by participating libraries as to why they have initiated RDM services to their libraries. Chawinga and Zinn (2019: 118) indicate that data sharing benefits the research community by driving scientific progress, reducing research fraud, and lowering research costs.

3.4.3.4 Institutional data repositories in support of RDM

A sound infrastructure in an institution, should be capable of supporting its data repositories as far as RDM services are concerned. Institutional data repositories are open access platforms which could be viewed as suitable platforms for supporting the management of the scientific knowledge which enhances knowledge generation, preservation, use and sharing as well as increasing the scale of research performance in a research community. Openness has been covered in Chapter two of the dissertation (see section 2.2.1.3) as the third capability factor of the CCMF model employed. An institutional repository (IR) could also be viewed as an archive for collecting, preserving and disseminating the digital copies that represent the products of the intellect in an institution, specifically a university. Ocholla (2020) notes the existence of 219 institutional repositories in Africa. He acknowledges 4.1%, of which 5309 in the global context, with 44 (20%) based in South Africa. Ocholla further remarks the "University of Pretoria's IR as leading in Africa in terms of number and variety of content". An organisation could use either an internal or external repository or both to manage their research data. Institutional repositories may house monographs, theses and dissertations, scholarly journals, conference proceedings, and other related scholarly publications. An institutional repository should be registered with the registry of open access repositories. Green, Macdonald and Rice (2009: 3) suggest the following activities that can be pursued by institutions and libraries in setting up data repositories:

- Raising awareness of data issues within institutions and the benefits of actively managing research data.
- Assisting in developing policies about data management and preservation.
- Providing advice to researchers about data management early in the research life cycle; influencing the way researchers will be creating their data, the formats they will use and building a commitment to use the repository to publish/preserve their data.
- Working with IT service colleagues to develop appropriate local data management capacity.
- Training and introducing data management and curation concepts to research students.
- Exploring methods of moving data from work-in-progress storage spaces to repositories in more seamless ways.

Ng'eno and Mutula (2018: 33) are of the view that collaborative work of librarians, information technologists, archivists, policymakers and research institute administration can help in achieving an effective data repository in an organisation. The Department of Science and Technology (2016: 20) points out some data repositories in the South African context such as those managed by the Human Sciences Research Council (HSRC), the South African Data Archive at the National Research Foundation and Statistics South Africa among others. Chiware and Becker (2018: 12) also confirm the National Research Foundation (NRF) in South Africa as an example of an accredited institutional repository. Qin (2013: 214) marks the growth of some data repositories like the Global Biodiversity Information Facility (GBIF), Dryad and GenBank in the USA in recognition of research data management as far as big data is concerned. Shearer (2015: 30) note Edinburgh DataShare to be an example of an institutional data repository at the Edinburgh University and Oxford University, where both academics and PhD students are encouraged to deposit their data. On the other hand, Shearer (2015: 37) asserts that in Canada the government sustains "repositories that house data in many areas deemed of national importance". Shear further notes both Figshare and Dryad as multidisciplinary repositories available to Canadian researchers. Rather, Ganale and Akhoon (2019: 121) view many institutions across the US as actively creating institutional

repositories and several field-specific online collections, particularly in the biological sciences. Rather, Ganale and Akhoon (2019: 124) therefore suggest that Library Information Science authorities need to play an active role with regard to the development and management of open access journals as well as repositories in their respective institutions to properly implement Open Access (OA) services in developing countries, South Africa included. Martinez-Uribe and Macdonald (2009) believe that researchers' needs should be considered first in order to develop repository services that can accommodate the management as well as the curation of research data. It is acknowledged by DCC (2015: 12) that repositories and databases can accomplish similar outcomes to a license through a provision of terms and conditions (T&Cs) of a contract that a user should abide by when wanting to access or deposit data. It is further indicated by DDC that "an institutional RDM service should be able to help researchers judge whether it is appropriate to agree to the T&Cs a repository attaches to data access and use, as this may affect later data sharing". With this in mind, it should be recognised that data repositories are paramount in managing research data and promoting suitable research practice as Chigwada, Chiparausha and Kasiroori (2017) also recommend establishing research data repositories or the use of already existing ones in ensuring the adherence to the standards of managing research data when doing research (Chigwada, Chiparausha and Kasiroori, 2017: 7). Higgins *et al*. (2018: 5) describes institutional repository as one of the operations and professional practices under their strategies for engaging with data and research data, respectively. Chiware and Becker (2018: 2) note a provision of research data access and re-use since the introduction of institutional repositories from some academic and research libraries in the South African nation. Nevertheless, Chiware and Becker (2018) indicate the presence of institutional repositories on their surveyed libraries, but the fact remains that some of the libraries are yet to explore DSpace and other commonly used platforms to manage data sets and metadata of any deposited data in either platform (internal or external) or both (Chiware and Becker, 2018: 14).

In this respect, it could be realised that data or institutional repositories exist to provide ready access to institutional research output through open access fashion and that gives global visibility to the research works of an institution. Operating an institutional repository could commit to a long preservation of data, information and

knowledge for future generations. As the world moves to an open knowledge condition, repositories are playing a valuable role in disseminating scholarly communication and research without having to pay any fees as opposed to subscription-based platforms. Though repositories serve significant functions in institutions, Gordon *et al.* (2015: 3) note creating and operating them pose some challenges like structural features on how a repository is integrated within the host institution and how it interfaces with content producers; new practices of managing integrated library systems and new skills from the workforce.

3.4.3.5 Capacity building

Capacity building is sometimes used interchangeably with capacity development. Capacity building could be referred to as a process by which individuals and organisations obtain, improve and retain skills and knowledge required to do their jobs competently. Often, people tend to define capacity building as training; however, it is more than that, due to its complexity. Capacity building or development could involve things like partnering with others and strategic planning. The Leeds Building Capacity Project (LBC) is one of the global projects in which libraries participate in the RDM practices and which is aimed at applying the "outputs and outcomes from existing JISC projects, services and intelligence to enhance research, learning and teaching at Leeds" (Gunjal and Gaitanou, 2017: 2). Kahn *et al*. (2014: 298) affirm that there is already available activity with regard to awareness and capacity building in South Africa. Kuhn *et al.* (2014) further highlight the Network of Data and Information Curation Communities (NeDICC), which is said to "arrange seminars, workshops and conferences to promote awareness around digital (including data) curation aimed at practitioners and managers involved with digital object management and encourage the growth of knowledge in this area". Another example of a capacity building platform is the UCT's Library and Information Studies Centre (LISC), which offers a Master's degree module in Digital Curation. It is indicated by Kuhn *et al.* (2014) that, since there has been growing interest in RDM in South Africa, with the attendant need for capacity building in this area, the LIASA's Higher Education Libraries Interest Group hosted a RDM workshop that was facilitated by the Digital Curation Centre (DCC), United Kingdom (UK), in Cape Town in March 2014. Capacity building programmes can enhance collaboration across departments,

within institutions or with external bodies in terms of the relationships they form during research processes. Collaboration is one of the eight capability factors in the CCMF and CCMF is one of the employed theoretical frameworks in the current research. Collaboration has been discussed in section 2.2.1.1 in Chapter two of the dissertation. Lyon *et al*. (2012: 21) state that types of relationships established during research are a major determinant of the types of research that can be performed in an organisation. Chiware and Mathe (2016: 7) point out that the Cape Peninsula University of Technology (CPUT) has set out several working groups at collaborative and institutional level to help in the implementation of RDM. The working groups include communication tools and researcher profiles; a reference management system and visualisation of data; research plan and electronic project journal; and search functionalities, semantic methods, and search engine optimisation. Van Wyk (2017) names five (5) pilot projects that were implemented through utilising Alfresco as a testing platform in 2013-2015 at the University of Pretoria in order to get an insight into researchers' needs regarding RDM services and activities. Those projects include the Institute for Cellular and Molecular Medicine (ICMM); Neuro-Physiotherapy; Potato Pathology Programme; Powdery scab; and Psychiatry Dissociation. Wiljes and Cimiano (2019: 1) indicate that the Bielefeld University in Germany has been teaching an interdisciplinary full semester course on RDM for a six-year period. Further to that, Wiljes and Cimiano (2019) point out that this university also offers a workshop on RDM, particularly designated for researchers through the recognition of wide-ranging capabilities of managing research data in various disciplines.

Gunjal and Gaitanou (2017: 2) highlight some of the global research projects in which several libraries participate. Those research projects include the Research Data Pilot Project at the Helsinki University Library; the DataRes Project; the re3data.org Registry Making Research Data Repositories; the ADMIRAL Project; the FISHnet Project; the Sudamih project at Oxford University and the joint German/Brazilian project on Mangrove Dynamics and Management (MADAM). Gunjal and Gaitanou (2017) also indicate that the Research Data Pilot Project based at the Helsinki University Library is yet to be developed to explore possible ways which could be used by libraries to contribute towards data management processes through gathering, storing and using research data and information on the current

practices and prevailing issues of researchers. The DataRes Project is a project that is aimed at information specialists or library professionals' perceptions which is funded by the Institute for Museum and Library Services (IMLS). In terms of library professionals, the focus is on those in a position to support RDM in terms of the central funding agency mandates. In the current study, library professionals included the librarians responsible for supporting RDM services as study participants. Gunjal and Gaitanou point out that the re3data.org Registry Making Research Data Repositories play a significant role in laying bases of the diverse landscape of the data repository and is a well-known project. The ADMIRAL Project is based at Oxford University and is known to have developed the infrastructure for pilot data management to support researchers in life sciences. The FISHnet Project has been of major importance in paving a platform in freshwater biology discipline for their research data curation as well as sharing. The Project of Infrastructure for Integration in Structural Sciences (I2S2) is basically focusing on the field of Chemistry and helps in exploring the requirements of the infrastructure in terms of data-driven data and is funded under the Research Data Management Infrastructure strand of the JISC's Managing Research Data Programme. The Sudamih project which is also based at Oxford University is aimed at supporting researchers in the field of humanities. Additionally, the joint German/Brazilian project on Mangrove Dynamics and Management (MADAM) is beneficial in research as well as the continuing long-term survey which contributes greatly to a sustainable resource management.

3.5 RDM related challenges in HIEs and libraries

Organisations such as universities come across all sorts of research data management challenges as digital data is ever growing and the current era demands more digital management of data than traditional or manual data, which in return demands proper infrastructure, funding, policies and skilled personnel in place.

Chawinga and Zinn (2020: 8) found key challenges in providing RDM to include a lack of policy frameworks, incentives, curation skills and training, a lack of storage network infrastructure, finding data re-usable data, and a lack of support from the university. Hamad, Al-Fadel and Al-Soub (2021: 81) observed librarian training, marketing and lack of communication between the research community and the library as the challenges that libraries might face when developing RDM services.

Anduvare (2019: 64) observes financial problems facing South Africa in terms of supporting and delivering RDM services, which she also notes as a problem facing developed countries. Anduvare's study also reveals that most universities seem to come across a cultural challenge when adopting RDM in the sense that not every stakeholder will find it easy to migrate from one form of managing and handling data to another quite easily. Avuglah (2016: 30) also notes that culture is the most serious aspect presenting an RDM challenge, ranked from the institutional level to the disciplinary level. In this regard, Avuglah (2016: 22) citing Van Deventer and Pienaar suggests a significant need for considering the cultural environment when institutions are implementing RDM initiatives. On the other hand, Molloy and Snow (2012: 249) imply "changes to the way research data are managed imply cultural change in the way research is practiced". Again, Morgan, Duffield and Walkley Hall (2017) also note a cultural change that could be quite problematic, particularly to academics. Penfield, Cox and Smith (2014: 28) too, observe varying cultures as challenging in the implementation of RDM initiatives. It was further observed by Penfield, Cox and Smith (2014) that there were several cultures being embedded across various professions and academic disciplines from participating institutions in their study. On the other half, Cox *et al*., (2017) found data skills and capacity-building to be the greatest hindrances in the proper delivery of the RDM services in academic libraries. Zang and Hu (2019:1) report a lack of bandwidth or capacity to enhance RDM activities, and therefore recommend institutional commitment to resources and training opportunities. Pinfield, Cox and Smith (2014: 28) note several factors affecting the development of RDM in libraries like acceptance, cultures, demand, incentives, roles, governance, politics, resources, projects, skills, communications and context. Gunjal and Gaitanou (2017: 10) note managing research data as the greatest challenge facing institutions of higher learning. Lötter (2014) highlights collection development, depositor support, digital object management, promotion of secondary data discovery, and use, as the challenges of RDM.

Though skills and knowledge have been documented and emphasised to be a requirement for the proper handling and efficient managing of research data in the institutions of higher learning and other research centres, it is unfortunate that Kahn *et al.* (2014: 302) acknowledge database development and software skills as the

most serious knowledge and skills gap in RDM. Conversely, Cox and Penfield (2014) reveal a skills gap and resourcing as major challenges in libraries when providing RDM support and services. On the other hand, Cox *et al*., (2019: 1434) agree that there are on-going limits to staff skills as highlighting curation skills, knowledge of research methods and data description to properly support RDM. Chawinga and Zinn (2020: 3) agree that a lack of skills isone of the most prominent challenges obstructing RDM progress. Chawinga (2019) observes a poor preservation of research data in most universities due to a lack of necessary skills among staff responsible for RDM. Ng'eno (2018) too, notes a lack of skills to enhance RDM activities. Again, Cox *et al*. (2017) discovered data skills to be the major challenges hindering the delivery of RDM services in academic libraries. Avuglah and Underwood (2019: 1) also found a skills deficit in terms of managing research data at the University of Ghana as revealing skills to be limited, not coordinated, as well as not formally established. Again, skills and knowledge were noted to be the most influential factors in terms of an RDM programme (Pinfield, Cox and Smith, 2014: 28). Chiware and Mathe (2016: 8) citing Cox, Verbaan and Sen assert that limited skillsets and a lack of understanding of the heterogeneous nature of research data in various disciplines remain the major challenges facing academic and research libraries in terms of offering RDM services. Additionally, Cox *et al*, (2017) found data skills to be the greatest hindrance that impacts on the delivery of RDM services in academic libraries.

RDM development strategies seem to be presented with numerous challenges in institutions of higher training and learning. The Council on Library and Information Resources CLIR (2013: 4-7) notes a lack of funding, organisational structures, professional preparation, the priority among researchers and institutional mandates as barriers to RDM in higher education institutions in the United States (US). Lötter (2014) notes policies, procedures, technology, capacity, as well as financial resources as constraints of the RDM. Anduvare (2019) observes storage failures and technological obsolescence as hindering the good practices of RDM as research data is ever growing and abundant in nature. Chawinga's study outlines a lack of training workshops for librarians, lack of data infrastructure, skills, incentives and recognition, as well as collaboration (Chawinga, 2019). Ng'eno (2018) reveals a lack

of collaboration and partnerships between institutions in order to enhance RDM activities. Similarly, a study by Masinde, Chen and Muthee (2021: 12) reveal most study participants to be dissatisfied with internal, external and internationally collaborative partnerships. Avuglah and Underwood (2019: 13) indicate that the University of Ghana has no formal RDM infrastructure or programmes currently in place, though this university's management acknowledge it as an imperative element that forms the basis of good research practices in the institution. Cox *et al*.(2017) observed capacity-building together with data skills to be the major challenges in delivering RDM services in academic libraries. Qin (2013: 214) notes a lack of linking mechanisms from already available resources for supporting RDM. Shearer (2015: 5) describes Canada as lacking in infrastructure, services and funding mechanisms to support extensive RDM. Molloy and Snow (2012: 25) remarks on a gap in data management training in the United Kingdom. Halbert in Gunjal and Gaitanou (2017: 10) observed several factors that could challenge the processes of RDM, such as lack of funding, lack of organisational structures, lack of professional preparation, lack of priority among researchers and lack of institutional mandates. Among challenges related to RDM development strategies, infrastructural deficiencies seem to be most popular in most studies. The following are challenges that Gunjal and Gaitanou (2017) felt significant to be addressed in libraries for successfully adopting RDM.

- User training of library staff
- Training for stakeholders
- Adoption of policies, such as Archival Policies, embargo, access rights, and other related policies.
- Support from stakeholders, such as technical expertise and an academic fraternity
- Voluntary submission of publications to the repositories
- Data curation
- Database creation
- Compliance of policies with funding agencies
- Upgrading of features and software
- Culture change
- Change management.

3.6 RDM frameworks for public universities

Several theoretical models have been used to understand the nature of RDM services both in Higher Education Institutions and libraries. Models and frameworks play a significant role for researchers in integrating RDM into their practice. Among others, the models used include The Data Audit Framework (DAF); Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO); Data Curation Centre Lifecycle Model.

Data Audit Framework (DAF)

Jones, Ross and Ruusalepp (2009: 6) and Jones, Ball and Ekmekcioglu (2008: 113) assert that the Data Audit Framework (DAF) was developed by the JISC-funded DAFD to help Higher and Further Educational institutions to identify data assets they hold as well as exploring the way they manage those data assets. Jones, Ross and Ruusalep further point out that DAF is “structured around audit at departmental or unit level with results being amassed to obtain an institutional or national perspective” (Jones, Ross and Ruusalepp, 2009: 6). Whyte and Allard (2014: 12) are of the view that DAF offers a quick approach to discovering data management practice through employing online and face-to-face surveys and interviews. Other than helping institutions understand the type of data they hold, the framework also indicates the location for data assets, the body responsible body for it, as well as offering tools for managing this information and sharing it with other organisations in a controlled environment (Jones, Ross and Ruusalepp, 2009: 5). The scope of this framework is only on research data assets as per the decision that was made between DAFD project management and the JISC with the purpose of restricting the scope of this framework. Jones, Ross and Ruusalep (2009: 10) acknowledge the information collected by this framework as sufficient in providing a clear overall picture of organisational data collections, current data curation and preservation policies and staff awareness of data issues. Though this framework is related to the current research, it focuses chiefly on how data can be audited in institutions, which is not the case in the present study.

Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO)

The Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO) is a tool and associated workflow for performing data management benchmarking across a data context that may range in size from a small project to a complete institution (DCC, 2011: 3). Whyte and Allard (2014: 13) affirm that CARDIO is a "benchmarking approach that can be used to assess the gaps between current and required support capabilities". Whyte and Allard further deduce that both CARDIO and DAF tools seek information based on current support regarding research data management. The two authors then indicate that DAF collects semi-structured information, whereas CARDIO uses rating scales in assessing offered support, with the intention of comparing over a period or across different groups. DCC (2011: 3) indicates CARDIO as having three principal sections, namely organisation, technology and resources. This concurs with Whyte and Allard (2014: 13) as indicating CARDIO as a tool using a rating scale to assess organisational, resourcing and technology elements in terms of providing services of RDM. With regard to organisation, issues of administration, policy and legal accountability are included, whereas technology includes the issues surrounding technological infrastructure as well as information security. In respect of resources, adequacy of staffing and financial sustainability is crucial.

DAF and CARDIO are two frameworks that are commonly used in RDM studies, but they were not used in this study. Instead, the current study used both the CCMF and DCC curation lifecycle models, as discussed in the dissertation's previous chapter (Chapter two).

3.7 Appraisal of the chapter

This chapter provided concrete evidence that research data management (RDM) is a global practice in most public universities as well as their libraries. It is also evident that the studied phenomenon (RDM) has received most attention in the Western nations as the reviewed literature was dominated by European publications. This correlates with the observation made by Kwanya (2021: 5) that "RDM has been a research issue in Sub-Saharan Africa for about 35 years". There is limited research about research data management in the African literature, therefore the current study

has contributed to the body of knowledge in the African context. The review of literature has established several gaps in terms of providing the RDM services at universities and other related centres.

The first gap is the paucity of financial support to deliver the RDM services. Financial constraints are experienced in both developing and developed nations of the world; however, this is most prevalent in the developing ones. It can be recognised that institutional resources or infrastructure highly depend on finance. The current research has significantly highlighted the importance of institutions of higher learning to secure sufficient funding for the support of the RDM services. On the other hand, this study has indicated the pivotal role that could be played by third parties in terms of sponsorship and collaboration.

The second limitation is a lack of resources or proper infrastructure to support the services of RDM. A proper infrastructure is really a requirement in support for RDM; however, most organisations feature improper infrastructure due to a lack of funding and support from parent organisations and other external bodies. The current study has contributed towards bridging this gap through investigating infrastructure as one of the development strategies in the public universities.

The third shortcoming is associated with skills deficits. Most authors affirm that staff responsible for the provision of RDM as well as researchers lack some relevant skills and knowledge. The present research has shed light on the relevant training needs and capability development programmes, where training some options were highly emphasised as most effective for skilling and re-skilling staff and researchers.

Fourthly, not all the public universities have RDM policies to support RDM services, and some have procedural guidelines that help to support a provision of RDM activities. The present research study has significantly thrown light on the RDM policies used by the selected public universities in support of RDM to address this gap. A lack of relevant RDM policies in a university is more likely to cause some difficulties in the management of research data as research data is ever growing and abundant in nature.

The fifth deficit is associated with varying cultures between disciplines or departments in universities. The current study has significantly deliberated on the relevance of having a uniform culture among disciplines to facilitate better practices

of managing research data in a university. This study has also developed an RDM framework in support of the provision of the RDM services in public universities in South Africa.

Table 3. 2: SWOT analysis of the reviewed literature

Strengths	**Weaknesses**
• Several universities such as the University of Southampton, University of Twente, Charles Darwin University, University of Pretoria, Cape Peninsula University of Technology, , University of Cape Town to mention but a few have developed RDM policies to help in the support of service delivery to ensure conformity among researchers and the provision of uniform services by researchers. • Numerous skills sets as required by both staff offering the RDM and the researchers has been documented in literature and curation, digital preservation skills and the understanding of research methods have been highly emphasised as really a requirement for better practices of RDM. • There are numerous RDM services and activities offered inthe institutions of higher learning, libraries and other research centres including collaborative research; support for data analysis and computational sciences; guidance on RDM, grant applications; data citation; data management planning; intellectual property and copyright; privacy and confidentiality; RDM workshops; scholarly output, etc. • South Africa is making significant progress in terms of offering RDM services among other African nations. • Developed nations like USA, UK, CA, AU among others, are in a good perspective in terms of RDM services as they have adequate resources and proper infrastructure in place to support RDM practices and activities.	• Varying cultures among departments or disciplines affect the way research data is handled and managed. • Lack of skills is one of the most noticeable challenges clogging RDM progress to most staff responsible for the provision of RDM services as well as for the researchers. • There is a paucity of sufficient resources for the offering of RDM in most institutions of higher learning and libraries, predominantly in African nations. • There are major infrastructural deficiencies in most organisations hindering the provision of RDM. • Most African countries are still in the infancy or implementation stage of RDM. • Namibia, Malawi, Zambia, Lesotho and Swaziland have no RDM- related activities in universities and research libraries. • Not all the universities have the RDM policies, strategies and procedures for facilitating the processes of RDM.

Opportunities	Threats
• There are some available global projects like the Leeds Building Capacity Project (LBC) which enhances the libraries' participation in RDM practices. • There are a number of joined nationwide services for supporting RDM services in HIEs and other data centres like the UK Research Data Services (UKRDS), UK Data Service, Australia's Research Data Service (RDS) and Data Intensive Research Initiative of South Africa (DIRISA). • There are available joined nationwide services for supporting RDM services in HIEs and other data centres like UK Research Data Services (UKRDS), UK Data Service, Australia's Research Data Service (RDS) and Data Intensive Research Initiative of South Africa (DIRISA). • Some institutions of higher learning like UCT, UP and University of Edinburgh offer some courses like MANTRA and training programs on RDM as f professional development opportunities for skilling and re-skilling individuals. • Many institutions of higher learning are investing in numerous infrastructures, both new and already existing ones to support the provision of the RDM services. • Several universities like UP, UCT, CPUT, DUT, Charles Darwin University, UT, OU and other have in place RDM policies and procedures that help in the proper management of data for a better understanding of its lifecycle, availability, accessibility, security, use and re-use purposes in the university context. • There are numerous RDM frameworks or models that significantly help in understanding the research lifecycle, capabilities, development and improvement of RDM in universities, libraries and data centres. • The Digital Curation Centre (DCC) is a	• Financial constraints hinder the offering of RDM in both developing and developed nations of the word. • There is a lack of collaboration /partnership between the stakeholders involved in the offering of the RDM services.

well-known centre in the UK that helps in the management and exposure of research data to increase visibility. • In South Africa, there are some established research initiatives, research councils, repositories, research centres and other research-related facilities or units such as SADA; CSIR; DIRISA among others that support the provision of the RDM services in this country. • The UK also has the JISC and e-Science Core Program which is crucial in the development of the RDM in terms of capacity and proficiency in an institution. • Australia has implemented the Australian National Data Service (ANDS) in response to digital research data. • Kenya has organisations like the World Agroforestry Centre which plays a crucial role in providing RDM training for agroforestry research scientists.	

3.8 Chapter summary

This chapter has reviewed global literature on the subject of RDM, putting more emphasis on higher education institutions and academic libraries.

In terms of the RDM practices and services, this chapter revealed many organisations as engaged and continuously engaging in providing RDM services to their user communities, higher education institutions included. It is however noted that the practice of the RDM services has been widely embraced in developed nations compared to developing ones. Most African institutions are halfway ready in providing RDM services as presented with numerous challenges such as a lack of RDM legal obligations, resources, infrastructure and inadequate skills. Among African countries, there is a significance degree of importance that South Africa is leading in terms of providing RDM services.

In terms of the competency requirements, there are a broad range of competencies required for staff responsible for RDM as well as for the researchers to ensure proper

management of research data and its accessibility. Among other capabilities, information skills, data curation and IT technology skills or ICT skills have been noted as the most urgently required skills. Other necessary skills include those in line with data lifecycle like the creation, organisation and validation of data. In terms of knowledge requirements, both knowledge of research methods and curation are really a requirement both for staff responsible for RDM and researchers. It is observed that required skills and knowledge can be embedded in the school curriculum or initiated through training, workshops and other skills development programmes, projects and strategies.

With regard to the development strategies, policy, investment, infrastructure, capacity building has been noted as significant in supporting the delivery of better RDM services in institutions like libraries and universities to enhance scientific research and its productivity and production. Most HIEs policies and procedures indicate that research data or any data should be kept, retained and archived for at least ten years. It is also worth noting that there are some global joint projects like The Leeds Building Capacity Project (LBC), UK Research Data Services (UKRDS), UK Data Service, Australia's Research Data Service (RDS) and Data Intensive Research Initiative of South Africa (DIRISA) currently in place to support nationwide services for RDM in HIEs and other recognisable data centres around the globe.

There are notably plenty of challenges facing higher education institutions and other data centres in managing their research data; however, varying cultures among disciplines, a lack of skills and resources or proper infrastructure have been cited to be common major and daunting constraints hindering the support of RDM services in most HIEs and libraries. These were the problems that triggered the undertaking of the current study, which were also discussed in Chapter one (section 1.4) of the dissertation. Other prominent challenges include a lack of collaboration between organisations and stakeholders involved, and a lack of necessary policies and procedures. Yet not all the identified gaps were covered by the present study. The present study sought to cover the problem in relation to varying cultures between departments, a lack of skills among the RDM personnel as well as the gap in infrastructure as the most cited challenges in literature, which signifies a great need for addressing them in order to support the provision of RDM services in the selected public universities in South Africa. On the other hand, some studies (like

Nhendodzashe and Pasipamire, 2017; Ng'eno and Mutula, 2018; Chiware and Becker, 2018; Mushi1, Pienaar and Van Deventer, 2020) noted RDM services as still being at their infancy stage in the developing countries, which could possibly mean most of the South African universities are still facing identified constraints regarding the offering of RDM services. It is for this reason we need to investigate the RDM services in the selected public universities in the country (South Africa) to bridge the existing gaps while adding to the existing body of knowledge to enhance good research practice in the LIS community of researchers in the country or even globally.

CHAPTER FOUR: RESEARCH METHODOLOGY

4.1 Introduction

The previous chapter (Chapter three) reviewed global literature on RDM, predominantly on higher education institutions and libraries. The present chapter is a roadmap describing the undertaking of the current research on the current status of research data management services in selected public universities in South Africa, particularly those based in the province of KwaZulu-Natal.

Research methodology is generally understood as methods, techniques, procedures and tools aimed at providing the work plan on how the entire research will be carried out. In other words, methodology throws more light on the methods that can be used to undertake a research project. The methodology is closely related to epistemology, as they are both concerned about how we come to know. However, the methodology is more practical while epistemology is simply about the philosophy of knowledge. Ngulube (2015: 127) views research methodology as "central to the research process because it is the lens through which a researcher looks when making decisions on acquiring knowledge about social phenomenon and getting answers to the research questions". He further stipulates that methodology helps to specify the types of research designs, as well as research methods that may be applied in gaining knowledge about a phenomenon of interest. Struwig and Stead (2001) define research methodology as a well-planned systematic process used by a researcher to collect and analyse data in order to solve a problem. Philosophical worldviews/research paradigms, research approaches/designs, research methods, study population, data collection tools/instruments, and methods of analysis are typically covered in research methodology. Previous related studies (Cox and Penfield, 2014; Patterton, 2016; Chiware and Mathe, 2016; Tripathi, Shukla and Sonkar, 2017; Ng'eno 2018; Ng'eno and Mutula, 2018; Chiware and Becker, 2018; Chawinga, 2019; Tang and Hu, 2019) have significantly acted as the motivation for employing some similar methods and techniques in the present research. Igwenagu (2016: 5) believes that research methodology has the following merits:

- Provision of tools for carrying out the research.

- Development of a critical and scientific attitude, and disciplined thinking to observations.
- Enrichment of the research process and provision of the opportunity for in-depth studies and understanding of the subject.
- Ability to evaluate and use research results with reasonable confidence and for decision-making.
- Ability to learn to read and think critically.

This chapter discusses post-positivism as a research paradigm in use. Following the adoption of post-positivism, both qualitative and quantitative research approaches were described. This chapter also discusses the research methods used for this study, including the consideration of multiple case study designs. The chapter also discusses the study's population and sampling. Data collection and analysis are also covered. The chapter concludes with a discussion of research trustworthiness, validity, and reliability, ethical considerations as well as how they were met in the context of the current study.

4.2 Research paradigm

The word research paradigm is normally used interchangeably with the term philosophical worldview or perspective, and largely influences research practice (Creswell and Creswell, 2018: 5). Research paradigms reflect how we view our social world in relation to a social problem. Neuman (2014: 96) asserts that a paradigm is an organising framework for theory and research, which encompasses basic assumptions, key issues, models of quality research, as well as methods for seeking answers. There are three types of research paradigms, which are positivism, constructivism and pragmatism. Most authors agree on their categorisation with minor re-organisations. However, the three categories can be further extended into four or five research paradigms. For example, Creswell and Creswell (2018: 5) identify four philosophical worldviews in research as post-positivist, constructivist, transformative and pragmatic. Ngulube (2015: 127) enumerates positivism, pragmatism and interpretivism paradigms in research. Neuman (2014: 96) highlights positivist social science, interpretive social science, and critical social science. Bryman (2012) enumerates positivism, post-positivism, realism, interpretivism and

pragmatism as research paradigms. Blanche, Durrheim and Painter (2006:6) identify positivist, interpretivism and constructivist paradigms. Positivism is associated with quantitative research, interpretivism is more qualitative, while pragmatism combines the elements of both qualitative and quantitative methodologies and tends to employ a mixed-method approach. Before choosing a research paradigm to inform a study, a researcher needs to take into consideration the assumptions of that worldview, related research design and methods, so that it becomes possible to translate an approach into practice (Creswell and Creswell, 2018: 5). Related studies (Ng'eno, 2018 and Chawinga, 2019) employed the pragmatism school of thought, but the present study was informed by a post-positivist worldview. The next section (4.2.1) discusses post-positivism as the employed paradigm and motivates choosing it in the current study.

4.2.1 Post-positivist

The post-positivist paradigm is also called the scientific method, science research positivist/post-positivist research, empirical science and post-positivism (Creswell and Creswell, 2018: 6). Post-positivism is believed to have emerged as a result of positivism. Panhwar, Ansari and Shah (2017: 253), citing Fischer (1998), observe that the whole idea of post-positivism is not to refute the elements of positivism, but to emphasize the perspectives of any research from the position of multi-dimensions as well as the methods. Post-positivists believe that not everything is completely knowable. Fox (2008) asserts that post-positivism is a modest approach to assessing reality and it describes an approach to knowledge inquiry. Post-positivists are of the view that the way we think in our everyday lives is not distinctly different from those of scientists. This surmises that scientific reasoning and common-sense reasoning are similar. Creswell and Creswell (2018) point out that post-positivists believe that some laws govern the world, which needs to be verified so that humans can better understand the world. The post-positivist stance acknowledges triangulation across sources, to get a clearer understanding of what is happening. This is supported by Clark and Fischer, as cited by Panhwar, Ansari and Shah (2017: 254), in the acknowledgement that post-positivism promotes the triangulation of both qualitative and quantitative approaches, which they believe allows the exploration of diversified facts that are researchable when different methods and tools are employed. In the current study, both qualitative and quantitative data were

gathered and triangulated at the data collection and analysis levels of the study, to have a better insight into the RDM services of the South African universities. Trochim (2020) observes that most post-positivists are constructivists who opine that the way we view the world is constructed based on our perceptions of it. Fox (2008) agrees that the post-positivism perspective offers the basis for both interpretivism and constructivist stance. Panhwar, Ansari and Shah (2017: 253) argue that post-positivism balances the positivism and interpretivism perspectives. Trochim acknowledges the relevance of the context of the wider community of truth seekers, such as scientists and scholars, like that of scrutinising the works of others, to improve our objectivity on how we do things in research and reality. Post-positivism holds an ontological belief that social reality exists as natural reality does, though imperfections exist, which could also be the research case. Phillips and Burbules, as cited by Creswell and Creswell (2018: 7), highlight some of the assumptions underlying the post-positivist perspective as follows:

- Data, evidence and rational considerations shape knowledge. In practice, the researcher collects information on instruments, based on measures completed by the participants, or by observations recorded by the researcher.
- Research seeks to develop relevant and true statements; ones that can serve to explain the situation of concern, or that describe the casual relationships of interest.
- Researchers must examine methods and conclusions for bias. Hence, being objective becomes a crucial element of competent inquiry in research.

Motivation for choosing a post-positivism worldview

The motivation behind the adoption of the post-positivist worldview was in line with the nature of data collected from the participants of the study. This study collected both qualitative and quantitative data from respondents, after which generalisation was made. Prathapan (2014: 45) advises that when using post-positivism in research, data can be collected using both qualitative and quantitative methods that complement each other in order to gain a thorough understanding of the research problem while getting closer to the truth. True statements in research are one of the assumptions by Phillips and Burbules (2000) that underlie the post-positivist paradigm. Qualitative data was collected from the staff responsible for RDM through

interviews, while the quantitative data was collected, through questionnaires, from academic researchers (NRF rated). Both qualitative and quantitative data were triangulated at data collection and analysis levels of the study.

4.3 Research approaches

In literature, the research approach is sometimes called research design. The research approach reflects the strategy of inquiry in research and could be a qualitative, quantitative or mixed-method study. A research approach is a plan and procedure for research that spans from broad assumptions to detailed methods of collecting, analysing and interpreting data (Creswell and Creswell, 2018: 3). Kumar (2011) is of the view that a research design is a procedural plan that is adopted by the researcher to answer questions validly, objectively, accurately, and economically. Kumar (2011) also advises that research design sets out specific details of inquiry in research. There are three types of research approaches/designs in research, namely qualitative, quantitative and mixed methods. Normally, the qualitative research approach is often associated with the use of words, instead of numbers that are used in quantitative research approach, whereas mixed methods research tends to combine the elements of both methods. Most related in-depth research studies (Avuglah, 2016; Patterton, 2016; Ng'eno, 2018; Chawinga, 2019) tend to employ both qualitative and quantitative approaches, as was the case with the present study. Employing both approaches is normally done to complement each other so that a complete picture of the study's problem can be portrayed as each approach has its strengths and weaknesses. The adoption of both methods/approaches in the present study was informed by the adoption of post-positivism as a philosophical worldview.

4.3.1 Qualitative research approach

The qualitative approach works very well with questions, sentences and words. The qualitative research approach is understood as an approach for exploring and understanding the meaning that groups or individuals ascribe to a social problem (Creswell and Creswell, 2018: 4). Data is collected in the participants' setting, as data were collected from the selected public universities in the country and the researcher made meanings through the interpretation of findings (Creswell and Creswell, 2018: 4). As data was collected during the COVID-19 pandemic, qualitative

data was collected from staff responsible for RDM, including librarians, library technicians, research directors, deans or deputy deans research, deputy vice-chancellors research, and department heads, from participating institutions (University of Zululand and Durban University of Technology) through interviews using Microsoft Teams and Zoom. Qualitative research approach is associated with inductive reasoning, whereby data analysis helps in building themes. This approach includes narrative research, phenomenology, grounded theory, ethnographies, and case studies (Creswell and Creswell, 2018: 12). The present study used multiple case studies, which are explained in section 4.4.1. Though there were five main research questions in the current research, that cover both the qualitative and quantitative aspects of the study, the application of the qualitative research approach has helped the researcher to predominantly answer the following main research questions, whereas the rest were addressed through the quantitative research approach:

1) How is the competency requirement of researchers and staff that are responsible for research data management in the selected public universities established?
2) How can an RDM framework for public universities in South Africa be developed?

4.3.2 Quantitative research approach

The quantitative form of inquiry is often known to relate to numbers or statistics and closed-ended questions. The quantitative approach tests theories deductively (Creswell and Creswell, 2018: 4) and builds in protection against bias. Findings can be generalised and replicated in quantitative research. In the present study, quantitative data was collected by sending out questionnaires on Google forms through emails to academic researchers (NRF rated) from the two selected institutions of higher learning (DUT and UZ). The findings obtained were then generalised to other higher education institutions in the province of KwaZulu-Natal. Creswell and Creswell (2018: 12) highlight experimental designs, non-experimental designs (e.g., surveys), and longitudinal designs, as forms of inquiry under quantitative research. However, the current research used quantitative case study through descriptive method (Shareia, 2016: 3840). The use of a quantitative model

of inquiry eliminated research bias in the current research. As discussed in section 4.3.1, some research questions were covered by applying the qualitative research approach. Qualitative and quantitative research approaches are believed to be complementary, rather than competitive. As the qualitative research inquiry covered the two highlighted research questions (see section 4.3.1), the quantitative approach helped to answer the following research questions in the current research:

1) What are the RDM practices of the selected public universities?
2) Are there RDM development strategies in place in the selected public universities?
3) What are the challenges facing the selected public universities in managing research data?

4.4 Research method/design

For Creswell and Creswell (2018: 16), a research method is a form of data collection, data analysis, and the interpretation that a researcher proposes for his/her study. Kara (2012) defines a research method as a systematic investigation for fact-finding and information collection that will improve knowledge. Often, people confuse and get lost in the maze of research methods and research instruments; however, these are two varying concepts. A research method can be understood as a method/mode of collecting data, whereas a research instrument is a tool used for data collection. Phenomenology, grounded theory, ethnography, case studies, and narratives are frequently used in qualitative approaches, whereas surveys and experiments are frequently used in quantitative approaches. For data collection, the current study used multiple case studies, both qualitative and quantitative.

.

4.4.1 Multiple case study

A case study investigates a phenomenon in its context. The case study is primarily a qualitative method in which a researcher investigates a specific activity, event, or situation in depth (Creswell and Creswell, 2018: 247; Neuman, 2014: 42; Kumar, 2011: 123). The current study used multiple case studies to investigate the research cultures of two public universities in the South Africa in the province of KwaZulu Natal, namely the University of Zululand (UNIZULU) and Durban University of Technology (DUT), in terms of RDM services, practices, and activities. UNIZULU is a

comprehensive university, whereas DUT is a university of technology, and their differences have significantly strengthened the study's findings. Gustafsson (2017) indicates that a multiple case study allows a researcher to analyse data within and across situations, which can significantly lead to a good understanding of the similarities and differences between cases for strong and reliable findings. Multiple case studies, on the other hand, are seen as allowing for a broader discovery of research questions and theoretical evolution. Neuman (2014: 42) confirms that case studies are more likely to generate better research theories. This has enabled the researcher to have a complete picture of the status of the RDM services from each university type, as Neuman (2014: 42) approves that a case-study research method can significantly incorporate an entire situation. The adoption of multiple case studies in the current study is informed by its adoption in a related study by Anduvare (2019) who used it to investigate the role of private university libraries in supporting e-Research in Kenya.

4.5 Targeted population and sampling of the study

4.5.1 Population of the study

Population involves all the elements that meet a specified selection criterion in a research study. Maree (2016) asserts that a targeted population is the entire group of people or objects from which a researcher is interested in drawing a sample that can be generalised to conclude the study. There are currently 26 public universities in South Africa and nine (9) provinces in the country, and each province has at least one university. Of the 9 provinces in the country, Gauteng province has the largest number (7) of universities, followed by Western Cape, KwaZulu-Natal and Eastern Cape, with 4 universities in each province; then Free State and Limpopo with 2 in each province, and the least with a single university each: North-West, Northern Cape and Mpumalanga. Among the 26 public universities in the country, the present study was undertaken at two that are specifically located in KZN, namely Durban University of Technology and the University of Zululand. The selection of the two universities was based on the provision of RDM services, as one institution (UNIZULU) is in the implementation stage of the process because one department has begun RDM, and the other (DUT) is in the pipeline because its RDM policy is still

in the draft stage, as well as easy access to the researcher. Furthermore, they responded positively to the researcher's request for participation in the study because the study also targeted the University of KwaZulu-Natal, but they only provided the researcher with the gatekeepers' letter. Due to financial constraints, the remaining public universities in the country were not included in the current study. The sampled universities are presented in Table 4.1.

Table 4. 1: Targeted public universities

University	Type	Campuses	Faculties	Departments
Durban University of Technology (DUT)	University of Technology	7	Applied Sciences,	10
			Accounting & Informatics	6
			Engineering and the Built Environment,	12
			Arts and design	9
			Health Sciences	11
			Management sciences	11
University of Zululand (UZ)	Comprehensive University	2	Faculty of Humanies & Social Sciences	18
			Commerce, Admin & Law	6
			Education	12
			Science & Agriculture	14

The Durban University of Technology (DUT) comprises six faculties. It is a medium-sized university with a total student enrolment of about 25, 000. DUT has seven campuses. The University of Zululand (UZ) is a small university type, with four faculties, and enrolls about 16, 100 students. All the selected universities have libraries that support the teaching, learning and research activities of students, lecturers, and non-academic members of staff, as well as members of the surrounding communities. These libraries also play a crucial role in housing the Institutional Repositories (IRs), which are a collection of all university thesis and dissertations that are available electronically through the internet websites or the library catalogue. Other than theses and dissertations, IRs also efficiently store,

manage, preserve, and maintain the research output and scholarly publications, such as research papers, conference proceedings, etc., of the universities' researchers. Therefore, IRs help both local and global research communities to have electronic access to universities' research outputs.

4.5.2 Sampling

Sampling involves taking a group of people, objects, or items from a larger population of interest for measurement. The current study used a multistage probability sampling method to obtain a desirable sample. According to Sedgwick (2015), multistage sampling consists of two or more random sampling stages based on the hierarchical structure of natural clusters within the population. Sedgwick defines clusters as "natural groupings of people." For example, in current research, the NRF rated researchers, librarians, library technicians, deans/deputy deans, and others. In the first stage, the researcher compiled a list of provinces within South Africa. Then, from among the country's nine provinces, one (KwaZulu-Natal) was chosen as the primary sampling unit due to its ease of access to the researcher. The researcher then listed all the public universities in the province of choice in the second stage. The secondary sampling units were two public universities (University of Zululand and Durban University of Technology) chosen from four in the province. Apart from universities in KwaZulu-Natal that do not provide RDM, the RDM services of participating universities are showing promising progress. One university (UNIZULU) has already started RDM in one of its departments, while the other (DUT) has a draft RDM policy. The researcher obtained a list of NRF-rated researchers from both participating institutions in the third stage. To obtain quantitative data, all rated researchers were chosen to participate in the study. This stage also included choosing librarians, library technicians, deans research, DVC research, and department heads to collect qualitative data. The study's ultimate sampling units were NRF-rated researchers and the rest of the population groups.

The selection of a sample to form a study's population in quantitative and qualitative research tends to vary. In quantitative research, Kumar (2011: 176) advises that a researcher attempts to select a sample in such a way that it is unbiased and representative of the population from where it is selected. Then, for qualitative studies, Kumar (2011: 176) observes numerous considerations that may impact the

selection of a sample, such as the ease of accessing the potential respondents; the researcher's judgment that the person has a wide range of knowledge about the phenomenon of interest, among other concerns. Sampling is usually grouped into two major categories, namely the probability and non-probability sampling methods. Probability sampling is associated with quantitative research while non-probability sampling is typically used with the qualitative approach. In the current study, the purposive sampling technique, which is based on non-probability sampling, was used to collect qualitative data from staff responsible for RDM support in participating institutions, whereas no sampling was used to collect data from academic researchers; instead, the researcher took them all. The study of Chiwanga (2019) and Ng'eno (2018) influenced the use of purposive sampling in the current research.

4.5.2.1 Purposive sampling

Purposive sampling is when a researcher is using his/her judgement to select cases that will help to answer the research questions. Leedy and Ormrod (2013: 215) assert that in purposive sampling, people or units are chosen for a particular purpose, as the name implies. In this context, purposive sampling was used to select three universities located in KwaZulu-Natal. The two participating universities were purposively selected in accordance with their orientation and type as well as the fact that the provision of RDM activities and services is in progress. The University of Zululand is a comprehensive university that offers a mix of the traditional university and university of technology programmes at both degree and diploma levels, and the Durban University of Technology is the technical university that concentrates on vocational training, and awards diplomas and certificates, but a limited number of degree programmes. Then, within the universities of focus, the selection was based on library, teaching and research. From the library, the staff responsible for RDM was selected, and those were the librarians and Library technicians/IT specialists. With respect to teaching and research, research directors (RD), deans (D)/ deputy deans (DD) research, deputy vice-chancellors (DVC) research, heads of departments (HODs) and academic researchers were invited to participate in the study. Detailed motivation for choosing each participating group of respondents is provided in the next sections (see sections A- G).

a) *Librarians*

The first group of respondents were librarians in charge of RDM support at the two public universities chosen. Participating librarians oversaw their respective libraries' electronic resources and metadata management. Librarians are typically information workers who provide access to information, information literacy programs, and library service promotion. However, in a university context, their duties and responsibilities tend to vary. For example, some Librarians are specifically dealing with information resources, institutional repositories, technical issues, and subjects in this environment. Therefore, in the present study, the librarians designated to support RDM were selected for their active role in the selection, acquisition, preservation, institutional repositories, archiving, and sharing of data in the university context.

b) *Technicians/ IT specialists*

One technician/ IT specialist from each participating institution formed part of the second group of respondents. In the university setting, technicians or IT specialists work closely with the information librarians responsible for RDM. This population of respondents was therefore chosen for their active and significant roles in implementing the infrastructure to support research activities, such as acquisition, storage, security, integration, preservation, archiving, searching and retrieval, use, and reuse, which are processes forming an integral part of RDM.

c) *Research Directors (RDs)*

RDs were the study's third participating group. Every university has researchers who ethically undertake research. In that sense, research directors have active roles to play to ensure that research processes in the universities are ethically carried out. RDs actively promote and facilitate cutting-edge research, which includes collaborative interdisciplinary research that reflects the goals of the university. They build, capacitate, and provide sufficient research support, thereby increasing internal and external research opportunities for faculties, post-doctoral fellows, and postgraduate students. In addition, RDs promote networking with other external bodies to promote research in universities.

d) *Deans/ Deputy Deans Research*

Deans/Deputy deans research were selected as operational heads of the faculties that ascertain that research matters are always on the right track. Deputy Dean's Research was the fourth population group of respondents selected as operations managers of research that oversee the processes and operations that are aligned with the institutional research office. In this sense, they help to provide direction, resources, and oversight that ensure that research is conducted following the university policies and procedures or that of sponsoring agencies. D/DDs research also play immense roles in encouraging the ethical conduct of research by overseeing issues that are related to the treatment of human and animal subjects as study participants; conflicts of interest; acquisition of data, management, sharing and ownership; publication activities along with their practices; responsible authorship; and shared research and reporting.

e) *Deputy Vice-Chancellors (DVC) Research*

DVCs research were the fifth population of respondents, and they were selected as the heads of research from their respective institutions that guide the development of research. DVCs research is well known for its pivotal role in delivering leadership and guidance to achieve universities' strategic priorities for research and innovation and focusing on quality assurance and enhancement. Again, DVCs research ensures active and ongoing involvement with research and postgraduate committees, as well as working with stakeholders to enhance research reproducibility and visibility in an institution.

f) *Heads of departments (HODs)*

HODs were selected for their active roles in promoting good practices and ensuring the safety and wellbeing of both staff and students that engage in research, knowledge use, and exchange or sharing of research. HODs also ensure that decisions are made to strengthen departmental research culture and philosophy, which could include using funds to support research performance and/or withdrawing them from areas where research is not strong enough. They also promote departmental research participation at regional, national and international levels through conferences, webinars and workshops. They also promote the engagement

between their departments and national bodies, industry, and political/community leaders for the development of research.

4.5.2.2. Total enumerative/ census sampling

Academic researchers

There were 60 NRF rated researchers from both the University of Zululand and Durban University of Technology. There was no sample drawn about this population group, as several scholars (e.g., Gay, 1996; Nueman, 2000; Ngulube, 2005) advise that sampling populations less than 100 is pointless. Kalusopa and Ngulube (2011: 132), advise that taking an entire population is called census. Initially, it was planned that this population group was going to be reached physically, however, as far as the COVID-19 pandemic and its restrictions and regulations were still in place during the study, research questionnaires were then transformed into Google forms and sent through emails. Though academics are carrying out their teaching roles, they are also expected to conduct research at the university. This explains their being selected for the current study, as research tasks are part of their everyday duties. They undertake research with the goal of publishing in scholarly journals. For academics to stay relevant, they are usually linked with research, publication and promotion. The next section highlights the sampling size and sampling frame of the current research.

Table 4. 2: Sampling size and frame

University of Zululand (UZ)		
Population	**Total number**	**Targeted number**
Librarians	15	2 (only librarians dealing with metadata)
Library Technicians	1	1
Research Director	1	1
Deans/Deputy Dean Research	8	4
Deputy Vice-Chancellors (DVC) Research	1	1
Heads of Departments	41	8 (randomly selected from each faculty)
Academic researchers (NRF rated)	21	All
Total	**88**	**37**
Durban University of Technology (DUT)		
Population	**Total number**	**Targeted number**
Librarians	3	3
Library Technicians	1	1
Research Directors	1	1

Deans/Deputy Dean Research	12	6
Deputy Vice-Chancellors (DVC) Research	1	1
Heads of Departments	59	12
Academic researchers (NRF rated)	39	All
Total	**116**	**24**

4.6 Data collection

The collection of data in research is determined by the nature (qualitative or quantitative) of the data to be collected and the problem underpinning the study. Normally, a qualitative form of inquiry collects data in the form of words or pictures, while a quantitative category collects data in the form of numbers. A study's research design determines what data will be collected and how it will be collected (Maree, 2016). Research instruments are measurement tools that are specifically designed to collect data on a specific topic of study. (Neuman, 2014). "Anything that becomes a means of collecting information for a study is called a tool or research instrument, for example, observation forms, interview schedules, questionnaires and interview guides" (Kumar, 2011: 41). The present study used interviews to collect qualitative data and questionnaires for its quantitative data gathering. The employed instruments are discussed in sections 4.6.1 and 4.6.2.

4.6.1 Interviews

An interview is a primary data collection tool involving a purposeful conversation between two or more people, e.g., a researcher and respondent, to get answers from a respondent on the problem under investigation. There are three types of interviews: structured, semi-structured, and unstructured interviews. The present study used semi-structured interviews, which were conducted with six population groups of respondents including librarians, technicians/IT specialists, research directors (RD), deans (D)/deputy deans (DD) research, deputy vice-chancellors (DVC) research and heads of departments (HODs) to elicit their views and opinions on RDM. The interviews covered the structured and unstructured sections with standardised and open-type questions. This aligns with Leedy and Ormrod's (2013: 190) suggestion that interviews should be relatively open-ended. As of the Covid-19 pandemic and its regulations, as well as the study participants' safety, all interviews

were conducted online via Microsoft Teams and Zoom. Because most universities used Microsoft Teams as a communication platform during the pandemic, most respondents preferred it. The researcher introduced himself and explained the study at the start of each interview. Before each interview, permission to record was requested following the introduction and briefing. The use of semi-structured interviews in the present study was informed by previous related studies (Avuglah, 2016; Patterton, 2016; Chiware and Mathe, 2016; Ng'eno, 2018; Chiware and Becker, 2018; Chawinga, 2019). Employing a semi-structured interview type enabled the researcher to add probing questions, where relevant, to elicit more in-depth information about the matter that was under investigation; that is RDM services. Prathapan (2014: 116) submits that interviews allow the researcher to get in-depth information about a studied phenomenon. Interviews also allow a researcher to establish rapport with participants, which motivates their cooperation (Leedy and Ormrod, 2013: 190).

Though interviews are presented with quite several merits, they are not without their shortcomings in research. Interviews, particularly face-to-face, can provide indirect information filtered through the views of interviewees (Creswell and Creswell, 2018: 188). Furthermore, Creswell and Creswell observe that the researcher's presence may sometimes prompt biased responses from study participants. This observation concurs with Kumar (2011: 142) in the submission that when using interviews for data collection, there may be researcher bias in the framing of questions and the interpretation of responses. Kumar also observes that "if the interviews are conducted by a person or persons, paid or voluntary, other than the researcher, it is also possible that they may exhibit bias in the way they interpret responses, select response categories or choose words to summarise respondents' expressed opinions".

4.6.2 Questionnaires

A questionnaire is more like a form containing a list of questions for data or information gathering in surveys (Kumar, 2011: 138). Initially, questionnaires were to be physically administered to academic researchers to obtain primary quantitative data. However, due to the Covid-19 pandemic and its restrictions and regulations

and for the safety of the study participants, the questionnaires with embedded links for Google forms were sent through emails. The questionnaire for academic researchers consisted of 38 series of both closed and open-ended questions that covered sections like biographical information; RDM practices; competency requirements of researchers and staff that are responsible for RDM; RDM development strategies, including the issues of policy, investment, infrastructure, and capacity building; and the challenges facing the selected public universities in managing research data. The adoption of questionnaires in the intended study was informed by related studies (Patterton, 2016; Cox et al., 2017; Ng'eno, 2018; Chiware and Becker, 2018; Chawinga, 2019). Prathapan (2014) asserts that the adoption of the questionnaire in a research study gives sufficient time to respondents to answer questions in a manner that gives a sense. In this regard, the current study intended to gain the same benefit as articulated by Prathapan. Moreover, using questionnaires in the current study means there was no person-to-person interaction, which enriched anonymity. Participants responded to questions with some assurance that their responses would not come back to haunt them in any manner (Leedy and Ormrod, 2013: 191). Kumar (2011: 141) agrees that the use of questionnaires for data collection offers greater anonymity since there is no face-to-face interaction between respondents and the interviewer.

Nevertheless, the use of questionnaires is also with some inadequacies. Usually, not all sampled population who receives questionnaires return them (Leedy and Ormrod, 2013: 191; Prathapan (2014: 117). Also, responses from questionnaires normally reflect respondents' literacy (reading and writing) skills; hence, questions can be misinterpreted. This concurs with the assertions of Kumar (2014: 181) and Kumar (2011: 141) that questionnaire administration requires a population that can read and write. Again, Kumar (2011: 141) also notes that the inability to clarify issues is a shortfall in using a questionnaire in research.

4.7 Data analysis

The presentation, analysis and interpretation of data form an integral part of a research process, being the stage when collected data is critically interpreted to make complete sense of the phenomena of interest. Williman (2011: 86) states that

data analysis can be used to identify trends, track progression, and replicate results in order to build a strong case for a specific problem. The presentation and analysis of the collected data in a research study are determined by the research instruments used for data collection. In the present study, the nature of the data collection tools called for both qualitative and quantitative methods of analysis. Primary data collected from librarians, technicians, research directors, deans/deputy deans research, deputy vice-chancellors research and the heads of departments, through interviews, were analysed qualitatively through the thematic data analysis method. The second primary data acquired through questionnaires from academics was analysed quantitatively with SPSS. Sections 4.7.1 and 4.7.2 provide a detailed description of both qualitative and quantitative data analyses for the present research.

4.7.1 Qualitative data analysis- Thematic

Qualitative data analysis involves the identification, examination, comparison and interpretation of patterns and themes (Ngulube, 2015: 20). There are numerous means of qualitative data analysis, and Ngulube (2015: 20) highlights discursive, thematic, structured and instrumental methods of data analysis. Thematic data analysis, among others, was employed for the current research. Qualitative data obtained from the staff responsible for RDM, through interviews, were analysed qualitatively according to themes. Thematic data analysis is considered a fundamental analysis for qualitative research studies. The use of qualitative thematic data analysis is informed by related studies (Avuglah, 2016; Ng'eno, 2018; Chawinga, 2019). Collected data from interviews was organised, sorted, arranged and transcribed. All the data were carefully read to get a general sense of the gathered information on RDM, which helped get a reflection of the overall meaning. After studying the data, it was then coded. Coding is the process of categorising the text data collected and labelling them with terms, which are usually the actual language of the participants (Creswell, 2014: 198; Kumar, 2011: 331). The coding of responses for this study was achieved by categorising the collected data into themes and subheadings. This was done to create a text-based version of the original audio that was recorded during the interviews. This was achieved using direct quotes and by paraphrasing some responses of the study's participants. Creswell (2014: 200)

believes that generated themes commonly appear as major themes that are often used in the Findings sections of dissertations; thus, the present research also used generated themes from the set objectives of the study to discuss the chapter on findings (Chapter six). Nevertheless, the collected data was unstructured and required to be sorted through, and organised, before the researcher could make sense of it. This is always the case with qualitative data analysis in research.

4.7.2 Quantitative data analysis- Google forms

Academic researchers (NRF-rated) provided quantitative data. It was originally intended that physical questionnaires would be used to collect data, which would then be entered into the statistical package for the social sciences (SPSS). Due to COVID-19 pandemic restrictions and regulations, however, questionnaires were converted into Google forms and emailed to study participants. A web-based questionnaire was then completed online, and the data was automatically recorded in an analysable format in a Google spreadsheet that allowed for data tabulation and graphical representation.

4.8 Trustworthiness of the study

The quality, authenticity, and truthfulness of qualitative research findings are referred to as trustworthiness. Lorelli *et al.* (2017: 3) citing Lincoln and Guba (1985), believe that the criteria for trustworthiness in qualitative research are credibility, transferability, dependability, and confirmability, which are like reliability and validity in quantitative research.

4.8.1 Credibility

For Stahl and King (2020: 26), one way of enhancing credibility in qualitative research is through triangulation, which could be methodological, data, investigator, theoretical and environmental. Methodological and theoretical triangulation were used in the current study. The researcher used methodological triangulation to collect and analyse both qualitative and quantitative data from study participants. Interviews were used to collect qualitative data, while online questionnaires were used to collect quantitative data. The researcher used two theories for theoretical triangulation: the community capability model framework (CCMF) and the digital curation (DCC) lifecycle model. On the other hand, in order to ensure that the study's

interpretations and findings were properly articulated, the researcher held debriefing sessions with the research supervisor for support and guidance.

4.8.2 Transferability

Transferability is established by showing readers that the research study's findings are applicable to other contexts, situations, times, and populations (Lincoln & Guba, 1985; Mamba, 2019; Stahl & King, 2020). As a result, a researcher cannot guarantee that the research study's findings will be applicable; instead, a detailed description of the studied settings is required. The researcher provided a description of the study context (see section 1.3) and study participants (see section 4.5) of the dissertation in the current study. On the other hand, a thorough explanation of the findings was provided, with sufficient evidence presented in the form of direct quotes from participant interviews (refer to section 5.2).

4.8.3 Dependability

According to Lincoln and Guba (1985), one of the aspects for dependability that improves clear documentation is peer review of the researcher's work. Peer review was sought and received from a departmental colleague in this study, who assisted in ensuring that collected data was clearly documented. Stahl and King (2020: 27) are of the view that dependability is important for trustworthiness because it establishes the research study's findings as consistent and repeatable. Stahl and King go on to say that researchers want to ensure their findings match the raw data they have collected. The current study ensured this by recording the interviews with study participants and taking careful notes during the interviews.

4.8.4 Confirmability

Confirmability is the fourth element of trustworthiness in qualitative research and is established when credibility, transferability, and dependability are all achieved. Lorelli (2017: 3), citing Tobin and Begley (2004), indicates that confirmability is concerned with establishing that the researcher's interpretations and findings are clearly derived from the data, which necessitates the researcher demonstrating how conclusions and interpretations were reached. Confirmability was achieved in the current study through the correlation of a study's findings, conclusions, and recommendations with data obtained from study participants.

4.9 Reliability and validity in research

Any research needs to be reliable and valid. According to Morse *et al*. (2002:14), "without rigour, research is worthless, it becomes fiction, and loses its utility, therefore, all research must have true value, applicability, consistency, and neutrality, in order to be considered worthwhile." As a result, reliability and validity are given special consideration in all research methods. In general, reliability refers to the measurement's consistency, whereas validity refers to whether the measurement measures what it is supposed to measure. Neuman (2014: 212) believes that reliability and validity are ideas that significantly help in establishing the truthfulness, credibility, or believability of findings in research studies. Maree (2016: 238) suggests that the reliability of a research instrument means that if the same instrument is given to different respondents from the same population, the results should be the same. Maree (2016) further points out that "reliability is the extent to which a measuring instrument is repeatable and consistent". Numerous scholars (Blanche, Durrheim and Painter, 2006; Creswell, 2014; Maree, 2016) hold the belief that the validity of a research instrument is about its measuring what it is supposed to measure. However, Maree (2016: 239) notes that validity is sometimes problematic in human sciences, as human emotions may need to be measured, and this could be highly unreliable.

Reliability constitutes the basis of good research. To ensure reliability in the present study, the researcher ensured that there was no participant error, basically on the semi-structured interviews conducted with the study participants: librarians, technicians, research directors, deans/deputy dean's research, deputy vice-chancellors research and heads of department. The interview was scheduled for a time convenient for the participants and took place in an environment conducive to them. Participant bias, which could have led to false responses, was reduced by not conducting group interviews. Research error was also reduced by not conducting all the interviews in a single day. Researchers, as human beings, are subjective, and may sometimes interpret results from respondents in a subjective way. However, in the current study, researcher bias (subjectivity) was minimised by interpreting responses objectively and not leaving out some responses. In the context of this study the reliability and validity of the findings were ensured by collecting data from various population groups (librarians, technicians, research directors, deans/deputy

deans research, deputy vice-chancellors research, heads of departments, and academic researchers) using different research instruments (interviews and questionnaires). Collecting data from these diverse population groups improved the reliability, allowing the problem under investigation to be approached from multiple perspectives. Triangulation was used in the study at both the data collection and analysis levels. Because avoiding ambiguity is an important characteristic of good research questions, research questions were structured to be concise. Respondents were also assured of confidentiality, allowing them to respond freely without fear of being identified. This elicited trustworthy and valid responses.

4.10 Ethical considerations

As research is always about collecting data from people or by people, it is always necessary that a researcher anticipates ethical issues that may arise when conducting research. In other words, for good research practice, ethics should always be taken into consideration. Ethics in research should be duly observed throughout the entire research project, and that includes before conducting the study, starting the study, data collection, data analysis, as well as reporting, sharing, and storing the data (Creswell & Creswell, 2018: 89-90). Neuman (2014: 530) suggests that ethical researchers should protect research participants, conduct research honestly in accordance with codes of ethics, avoid interference from sponsors, and publicise results openly and clearly.

The current study followed research ethics by obtaining permission from the University of Zululand Research Ethics Committee to conduct the entire study via an ethical clearance certificate (see appendix D). To gain access to the research participants, a motivational letter (see appendix A) was prepared and sent to the University of Zululand, Durban University of Technology, and the University of KwaZulu-Natal, as the selected public universities in the province of KwaZulu-Natal. A motivational letter was sent to the participating institutions' research offices and institutional libraries. The University of Zululand responded to the researcher via email, saying that he could continue with the research because he was a student there. The researcher was granted gatekeeper letters by the Durban University of Technology research office and library (see, appendix B and C). The University of

KwaZulu-Natal research office also granted the researcher a gatekeepers' letter, but there were no responses from study participants. The research instruments (see appendix E, F, and G) were examined to ensure that they met the ethics committee's standards. Because the interviews were conducted online, study participants' permission to record the interviews was obtained before each interview began. The confidentiality and anonymity of the participants were ensured through the coding of their responses. Creswell (2014: 99) emphasises that there should be respect for the privacy of participants. The researcher also ensured that not only the positive results were disclosed; instead, all the collected responses were recorded and interpreted unfavourably. As Creswell (2014: 99) remarks, withholding important results or casting results that are not favourable to the researcher or participants is academically dishonest. In the current study, the researcher ascertained that there was no falsification, fabrication or misconduct in the interpretation of the collected data, to enhance trustworthiness and verification. All these factors have significantly contributed to the validity and reliability of this research.

4.11 Methodological challenges/limitations

As is the case with most successful studies on research, the present research study also encountered some hindrances. Fortunately, these shortcomings were not such that they could significantly affect the findings or hinder the success of the entire study. The identified shortcomings could be potentially useful in the LIS field, in terms of selecting methods and techniques, when undertaking similar studies and contribute to the debate on the research methodologies in this domain.

One limitation encountered in the current study was caused by the COVID-19 pandemic. With the restrictions and regulations put in place during this study, the pandemic has been a significant challenge. Originally intended to physically reach study participants, the pandemic forced everything to be moved online. The study planned to conduct physical interviews with librarians, technicians, research directors, deans/deputy deans of research, deputy vice-chancellors of research, and department heads from both participating institutions (UNIZULU and DUT), but due to the pandemic's rules and regulations, interviews were eventually conducted online via Zoom and Microsoft Teams. Questionnaires, on the other hand, were converted

into Google forms and emailed to NRF-rated researchers. However, the response rate in questionnaires (quantitative data) was low; the researcher targeted 60 NRF-rated researchers, but only half (30- 50%) responded. Furthermore, respondents took an unusually long time to complete questionnaires, delaying quantitative data analysis. The study, on the other hand, initially targeted three institutions in the province of KwaZulu-Natal, including the University of KwaZulu-Natal (UKZN). Unfortunately, there was no response other than receiving a letter from the gatekeepers demanding that the researcher exclude this institution (UKZN).

Another drawback relates to the fact that it was not possible to survey all 26 public universities in South Africa due to insufficient project funding and time constraints. As a result, this study chose only two public universities, primarily those in the province of KwaZulu-Natal. Using more than one method in a single study, as the current study does, is thought to have complementary benefits. As a result, both qualitative and quantitative research approaches were used through questionnaires and interviews to gain a comprehensive understanding of how South African public university libraries manage their research data. The study has noted related studies in the domain in the country and their shortcomings.

4.12 Chapter summary

This chapter has shed light on the study's research methods, designs, techniques, and procedures. Previous related studies appear to use both qualitative and quantitative research inquiry, surveys, and data collection tools such as questionnaires and interviews, with data analysis methods such as SPSS and thematic analysis. The post-positivist paradigm served as the foundation for this study. The researcher was able to collect both qualitative and quantitative data from a variety of participants, including librarians, technicians, research directors, deans/deputy deans of research, deputy vice-chancellors of research, department heads, and NRF researchers. The data collected from these various population groups was triangulated at both the data collection and analysis levels. A quantitative case study was conducted with academic researchers (NRF-rated), while semi-structured interviews were conducted with librarians, technicians, research directors, deans/deputy deans research, deputy vice-chancellors research, and heads of

departments using the multiple case study method. Google forms were used to analyse the data collected through questionnaires. Data gathered through interviews was analysed thematically. The research instruments' dependability and validity were assessed. To ensure that the study was carried out ethically, ethics were maintained throughout the entire research project. a

Table 4.3 is the methodological alignment / dashboard of the study and Table 4.4 presents the outline of the methodologies in the previous related studies. Most studies on RDM practices in the global context appear to use both qualitative and quantitative research approaches, as well as surveys in the form of a questionnaire. These related studies have significantly influenced the application of some methods, techniques, and procedures in the current research, as they have previously been successfully applied in RDM in a variety of contexts such as universities, research centres, and libraries, to name a few. Before attempting to apply the current research methods to understand the current state of RDM practices in South African public universities, some methods were reviewed and revised.

Table 4. 3: Methodological alignment of the study

Research Questions	**Research paradigm & Research approach**	**Research design/ method**	**Target**	**Data collection instrument**	**Data analysis**
What is the competency requirement of researchers and staff that are responsible for research data management in the selected public universities established?	Post-positivism (qualitative)	Multiple Case study	Librarians, Technicians, Research Directors, Deans/ Deputy Deans Research, Deputy Vice-Chancellors Research, Heads of the Departments	Semi-structured interviews	Thematic
What are the RDM practices of the selected public universities?	Post-positivism (quantitative)	Quantitative case study	Academic researchers	Structured questionnaires	Google forms
Are there RDM	Post-positivism	Quantitative	Academic researchers	Structured	Google

development strategies in place in the selected public universities?	paradigm (quantitative)	case study		questionnaires	forms
What are the challenges facing the selected public universities in managing research data?	Post-positivism	Quantitative case study	Academic researchers	Structured questionnaires	Google forms
How can an RDM framework for public universities in South Africa be developed?	Post-positivism (qualitative)	Multiple Case study	Librarians, Technicians, Research Directors, Deans/ Deputy Deans Research, Deputy Vice-Chancellors Research, Heads of the Department	Semi-structured interviews	Thematic

Table 4. 4: Outline of the methodologies in the previous related studies

Study aims	**Paradigm, method/design/C1**	**Data collection/ C2**	**Data analysis/C4**	**Study population/ C3**	**Author(s), date and country**
Investigated how research data has been generated, organised, shared, stored, preserved, accessed and re-used in Malawian public universities.	Pragmatism Mixed methods	Questionnaires Interviews	SPSS Thematic	Librarians Researchers Directors of research	Chawinga (2019)- Malawi
Examined Research Data Management (RDM) in Kenya's	Pragmatism Mixed methods	Interviews Questionnaires Documents	SPSS Thematic	Directors of institutes Heads of research Heads of IT	Ng'eno (2018)- Kenya

agricultural research institutes.		review		Librarians	
Establishing the data management practices of emerging researchers at the Council for Scientific and Industrial Research (CSIR).	Qualitative Quantitative	Survey Case study Questionnaires Interviews Spreadsheet Data documentation	Emerging researchers	Univariate analysis Levels of measurement Measures of central tendency Measures of dispersion Data correlation	Patterton (2016)- South Africa
Assessing the state of the art of RDM and institutional preparedness at the University of Ghana.	Qualitative	Case study Interviews Document analysis	Service providers Senior researchers	Thematic	Avuglah (2016)- Ghana
Determine how ready university and research libraries are in leading and participating in research data management services.		Survey Questionnaire	Librarians Library directors		Chiware and Becker (2018)- South Africa
Investigate how the library at the Cape Peninsula University of Technology is developing and integrating RDM services into institutional research workflows.		Survey Questionnaire	Librarians		Chiware and Mathe (2015)- South Africa

The following chapter critically examines the presentation, analysis, and interpretation of the data gathered from the current research's diverse respondents.

CHAPTER FIVE: DATA ANALYSIS AND PRESENTATION OF FINDINGS

5.1 Introduction

The previous chapter discussed the methods, procedures, and techniques used to carry out the entire study (research methodology). This chapter presents the analysis and interpretation of data gathered by using both quantitative and qualitative approaches. Qualitative data originated from interviews with seven (7) sampled population groups, namely librarians, technicians, research directors, deans/deputy deans research, deputy vice-chancellors research and heads of department. Quantitative data were obtained from NRF rated researchers by using questionnaires. The study participants were from the University of Zululand and Durban University of Technology. The data is presented in two parts. The first focuses on qualitative data based on responses from the sampled library, teaching and research staff. The second part focuses on quantitative data obtained from the NRF-rated researchers. Each part is organised by the themes from the research objectives.

5.2 Data Analysis- Qualitative Responses

5.2.1 Data from the library

The first and second groups of participants were both librarians and technicians from the selected universities, as explained in Chapter four (section 4.5.2.1) of the dissertation. An interview was used to collect data from these population groups (see appendix: D). Originally, it was intended to reach out to these populations physically; however, because the COVID-19 pandemic and its restrictions and regulations were still in effect during this study, all interviews were conducted online via Zoom and Microsoft Teams. Participating librarians oversaw the library's electronic resources and metadata management. It is the technicians' responsibility to collaborate closely with the RDM information librarians. This population of respondents (IT technicians) was chosen because of their active and significant roles in implementing the infrastructure to support research activities such as acquisition, storage, security, integration, preservation, archiving, searching and retrieval, use, and re-use, all of which are processes that are integral to RDM. A brief description of the study was

articulated at the start of each interview. Before each interview began, permission was obtained to record the interviews following the introduction and briefing. Previous related studies (Avuglah 2016; Patterson 2016; Chiware & Mathe 2016; Ng'eno 2018; Chiware & Becker 2018; Chawinga 2019) informed the use of semi-structured interviews in the current study. Careful notes were taken during each interview. The recorded data was transcribed to ensure that the details of each interviewee were correctly captured. Data are presented in this section based on biographical details; RDM practices; competency requirements of researchers and staff responsible for research data management; RDM development strategies; challenges faced by the selected public universities in managing research data; and an RDM framework for South African public universities. The collected data is organised into themes and subheadings derived from the study's research objectives and questions.

Table 5. 1: Response rate

Institution	**Population group**	**Total number**	**Targeted number**	**Responded**
University of Zululand (UZ)	Librarian	15	2	2
	Technician	1	1	1
Durban University of Technology (DUT)	Librarian	3	3	3
	Technician	1	1	1
Total		**20**	**7**	**7**

5.2.1.1 Biographical details

Table 5. 2: Biographical details

Participant	Gender	Age category	Institutional name	Academic qualifications	Job position	Work experience
P1	Male	25-34 years	UZ	Diploma in network engineering	Library Technician	4 years
P2	Female	45-54 years	UZ	Master's in information technology	Senior Librarian (Electronic resources)	5 years
P3	Female	55-64 years	UZ	Masters in Blbl	Senior Librarian (Metadata and electronic resources)	20 years
P12	Male	45-54 years	DUT	Master's in Information Technology	Library systems and resources manager	10 years
P13	Male	45-54 years	DUT	Master's in Arts	Digital services Librarian	1 year six months
P14	Female	45-54 years	DUT	B in Honours	Metadata Librarian	12 years
P15	Female	45-54 years	DUT	Master's in information technology	Metadata Librarian	14 years

Table 5.2 presents the biographical data of the study participants of the librarians and library technicians. Most were female participants falling in the 45-54 age category. The majority of the study participants held a Master's degree as the highest academic qualification and had more than five (5) years of work experience (5, 71%).

5.2.1.1.1 Research-related duties

This question sought to ascertain the research-related responsibilities of both librarians and library technicians, and their responses are as follows.

Participant 2: "*I am responsible for accessing and acquiring electronic resources and also ensuring that our library users can access relevant electronic resources wherever they are. I also promote open access".*

Participant 3: "*We provide support to researchers and postgraduate students, as a library we provide support in terms of facilities and resources and special services to students".*

Participant 12: "*Investigations into library systems and impact measures for academic libraries".*

Participant 13: "*My duty is to give a research support role".*

Participant 14: "*We upload the completed works of the researchers in the institutional repository. However, there is not that much, except for providing advisory support through discussions with students or faculty representative officers when they have to submit the item whenever necessary. As I have more knowledge of metadata, I also advise on how we can help or guide masters and doctoral students in submitting their research work like guiding them on providing the basic elements such as the titles, authors, and how they should name their files and file formats on their work; hence my research duty comes into play".*

Participant 15: "*I am the metadata librarian for the institutional repository and other repositories applications at DUT and I make information accessible to users".*

5.2.1.2 RDM services at the selected public universities

5.2.1.2.1 RDM services' provision and practices in support of the research data in the library

This question sought to determine whether or not participating libraries offer RDM services, as one of its goals is to support research in addition to teaching and learning. Aside from whether or not the libraries provide RDM services, it was necessary to determine whether the participating libraries provide any activities or

practices in support of the institution's research data activities. The verbatim narratives from participants are presented as follows.

Participant 1: "*No, I am not aware of any services of RDM".*

Participant 2: "*The library does not provide RDM services, apart from us depositing the already published research activities in our institutional repository, however, we are not involved right from the beginning of the research works of the researchers".*

Participant 3: "*We do not provide the services of RDM unfortunately".*

Participant 12: "*We have systems in place and processes, we haven't started, and we are still in the implementation stage of the process. However, the library supports data storage, sharing, preservation and archiving activities. We also support data creation by assisting researchers in doing data management plans or data planning".*

Participant 13: "*Yes, the institution does provide some services of RDM, but not the library per se. In terms of research-related activities, we offer preservation, archiving and sharing activities. We also offer data management plans".*

Participant 14: "*Not as yet, but it is something that is in the pipeline. We haven't started it full-time as the RDM policy is still in the draft state. However, the library assists researchers in terms of developing their data plans and how to preserve research data for ease of access and retrieval".*

Participant 15: "*My library offers data creation and storage activities. In terms of data creation, I add to the repository, whatever metadata is supplied by the research office".*

5.2.1.2.2 Library's storage, sharing and management of published research data

This question sought to determine where the library keeps, shares, and manages the published research data. The verbatim responses from the study participants are presented as follows.

All the study participants indicated using DSpace institutional repository to store, share and manage the published research data. Participants also indicated the

period of research data storage in the institutional repository (IR) and their responses are as follows.

Participant 1: "*In our institution, we do not remove data from the institutional repository as we have data that was uploaded a long time ago".*

Participant 15: "*For now, we are keeping the theses, dissertations and articles forever as we haven't started archiving their research data as yet".*

Participant 2: "*We do not delete the uploaded research data; it is kept permanently in our institutional repository".*

Participant 12: "*The university has the policy for five years, but that refers to printed works. In terms of the electronic, there is no limited time as the RDM policy is still in the pipeline, so I think it will be available as long as the life of the article or thesis".*

Participant 3: "*The published works of the researchers are kept permanently in the repository, and it is manageable as there are not many copies in there. There are about 1800 items".*

Participant 13: "*There is unlimited time for research data storage".*

Participant 14: "*In terms of the proposed draft policy, research data is going to be kept for ten years, unless otherwise there are some changes. However, for now, it is kept unlimited".*

5.2.1.2.3 Institutional repository in libraries

This question sought to determine if participating libraries have institutional repositories, measures put in place to enable researchers to self-archive their research data, institutional repository typology and preferred file format for research data to be captured in the system. The responses are presented as follows.

Participant 1: "*Yes, the library has DSpace institutional repository which is open access type. Researchers cannot self-archive their research data, instead, the staff generally does the upload. Generally, all the files in the institutional repository are in Portable Document Format (PDF)".*

Participant 2: "*The library has an institutional repository called DSpace, though researchers cannot self-archive their research data".*

Participant 3: "*The library has an open-access institutional repository. Students cannot self-archive their research data as we do not keep research data, instead, we keep publications in a form of journal articles, conference proceedings and theses and dissertations. However, we try to use open access events, for example, open access week and we try to invite the researchers, administrators, research office and so on to promote research in the institution. Unfortunately, it is like a roller coaster, we get good responses sometimes everything stays, so we decided to have the initiative and put open-access published articles on the institutional repository platform. PDF is preferred, however, there is a newly introduced system called Arkivum in our library which will accommodate all file formats like images, videos, text, and other file formats. Arkivum archives data stored in the institutional repository and serves as a backup, should we lose data stored in DSpace, then we won't lose it".*

Participant 12: "*There is an institutional repository, but we call it DUT Open Scholar".*

Participant 13: "*Yes, there is an institutional repository called DUT Scholar, but it is DSpace by the way".*

Participant 14: "*Yes, the library has an open-access institutional repository, called DUT Scholar and it is open. Students cannot self-archive their research data however, they get guidance from their respective faculty officers since faculty officers engage with the library in terms of research-related matters. Faculty representatives/officers play a huge role in facilitating research-related processes such as guiding students on required elements for them to upload their research work. So, as the library, we only upload the finished works of the researchers, and the library prefers a single PDF file in its system".*

Participant 15: "*Correct, we do have an institutional repository which is open for everyone. We are working on self-archiving, and it should be ready by next year, but for now, any research articles go via our research office, and they upload it on elements and our thesis and dissertations are uploaded to the repository by the faculty officer. Any file format is acceptable, but at the moment it is mostly PDF and Word documents".*

5.2.1.2.4 Amount of research data in gigabytes (GB) or terabytes (TB) kept in the library system

Participants in the study were asked to estimate the amount of research data stored in the library system. The responses came from the library's technical department, as technicians are in charge of implementing infrastructure to support research activities such as acquisition, storage, preservation, and archiving, among other essential research-related activities. The following responses were provided by participants.

Participant 1: "*Our system is currently accommodating about one (1) terabyte, but it can be updated it's just that currently, we do not have a lot of data".*

Participant 12: "*The amount of research data our library system can keep is terabytes as we are also planning to provide RDM".*

5.2.1.2.5 Requirements of the researchers to publish their research

It was necessary to identify the requirements of the researchers for their research works to be published. Some of the selected responses are presented below.

Participant 2: "*One of the requirements includes a researcher being a registered student in the University or a staff member within the university. We also consider if the research was done while the researcher was in the institution".*

Participant 12: "*Currently we have a research data management policy which is in the consultation phase at the moment which will stipulate the requirements of the researchers to publish their research data in the institution".*

Participant 14: "*To publish your thesis or dissertation you have to be enrolled at the DUT and in terms of articles, you still have to be a registered student or staff member. We do not publish the research works of students registered in other universities like UNISA, however, donations are allowed".*

Participants 15: "*Well, they must be students or staff in DUT and must be affiliated with the institution".*

Participant 13: "*In terms of the requirements, there is a policy that has not been implemented yet which will probably speak to that".*

5.2.1.2.6 Researchers' awareness of the publishing platforms and requirements

Participants in the study were asked to share their thoughts on whether they believe researchers, particularly postgraduates, are aware of where they need to publish their research data and what they need to do for their research work to be published. The following are the responses.

Participant 2: "*We normally have workshops which we arrange and organise with publishers for postgraduates or other researchers within the institution. The publishers train them on what to do when one needs to publish and also librarians have a role in providing awareness to researchers like informing researchers on some basics for publishing. Also, the research office and teaching and learning centre do provide some publishing training".*

Participant 3: "*Postgraduates students are aware of where they need to publish their research projects including master's and doctoral theses and dissertations as they are made aware through attending postgraduate induction postgrads are also provided with postgraduate booklets which contain procedures that need to be followed by students for the publication process. For example, the booklet contains information that stipulates the expectations of the postgraduate students in the completion of their studies, such as having to submit both electronic and printed copies of their research works. Librarians train and educate postgraduates and researchers on the self-archiving of research data".*

Some participants felt that students had no knowledge of publishing and what is required from them. The noted responses were in relation to the RDM policy which is still in the draft stage. The participants noted the following.

Participant 12: "*I don't think so, but we are on the marketing drive at the moment. We are presenting it on the faculty board to take to various forums, as I've just said from the marketing point of view we are there. However, on the actual ingestion of getting people to actual data activities, we are waiting for the policy to be approved, that is where we are sitting at the moment".*

Participant 13: "*No, they don't know because these forms part of RDM which it's a new service and there is no policy to enforce it yet. Regarding the research papers,*

there is a policy in our institution that specifies that all the research papers have to be published in the institutional repository".

Participant 14: "*I think some do, most of them have no idea though. This may relate to most students' habit of liking to be spoon-fed and being lazy to find out for themselves. For example, it is so surprising to find a postgraduate student who has not mastered some basic elements of research, such as proper referencing. But I am hoping that as soon as we kick start this project (RDM) we are going to liaise a lot with the research office to support postgraduates in terms of their research-related activities including publishing in accredited journals".*

Participant 15: "*No, at the moment I don't think they are very much aware of it unless it's stipulated in the funder's policy like NRF".*

5.2.1.2.7 Benefits of RDM for emerging researchers

It was deemed necessary to understand the benefits of RDM for emerging researchers in universities. The study participants felt as follows.

Participant 1: "*Benefits of research data management include research visibility and sharing of the research findings".*

Participant 2: "*One benefit is managing their data which is to prevent data loss and promote security and use and reuse purposes".*

Participant 3: "*It is a very useful practice as it will help researchers' data storage and availability. It also allows research data sharing and re-uses among researchers. It can also be useful in fighting plagiarism or academic dishonesty. It is a very interesting and useful practice by the way".*

Participant 12: "*RDM facilitates the entire process of managing research data such as collecting data, storing it, analysing data and so on for the researchers, particularly postgraduate students. It also helps in research data being available and discoverable".*

Participant 13: "*The availability of research data to other researchers avoids duplication of effort. It also keeps research open and transparent".*

Participant 14: "*One of the most important benefits of RDM is enhancing knowledge sharing and preventing data duplication and also making your data retrievable".*

Participant 15: "*The benefits would firstly be not to duplicate research and also to make the research more visible and allow for full citations".*

5.2.1.3 Competency requirements of staff responsible for research data management in selected public universities

5.2.1.3.1 Knowledge and skills required to support RDM services

Participants were asked to identify the knowledge and skills required to support RDM service delivery in their libraries. Their responses are interpreted as follows.

Participant 1: "*Well obviously one needs to have both knowledge and skills in data management and curation and technical skills, especially in my role as a technician".*

Participant 2: "*You need to understand research methodology, institutional repository and publishing. You also need to have computer and metadata skills".*

Participant 12: "*I think it is important to understand on a policy level how research takes place at the university, how different research-related policies work, how research is undertaken in the university context and high technical skills. One also needs to understand operating institutional repositories and lastly, skills in meta-creation".*

Participant 13: "*I think knowledge of metadata is required so that research data can be searchable and discoverable".*

Participant 14: "*A person needs to have more knowledge and skills in metadata, data curation and data storage".*

Participant 15: "*I think you need to have good communication skills and also interviewing skills so that when you speak to researchers you will be able to assure them that their data will be safe as they are not trusting and guard their research data".*

5.2.1.3.2 Need for new skills and other competencies of RDM and ways to acquire them

Participants in the study were asked if they needed new skills and other related competencies to support the provision of RDM services and how they think they would acquire them. Their responses were as follows.

Participant 1: "*Yes, there is a need for new skills, especially IT and technical skills as they are ever-changing, so I always need new skills to keep up with the latest trends taking place. I think I can acquire them through workshops, training, and attending conferences".*

Participant 2: "*Given that I must manage this RDM project, yes, I would need to go for some training or workshops"*.

Participant 3: "*Absolutely, I need new skills as well as other employed senior librarians. This requires human resources to provide relevant training in support of the library provision of RDM services and workshops and seminars could be good platforms".*

Participant 12: "*In terms of skills and knowledge I think we are covered as we have attended UCT short course and training on RDM. We are currently working on the university implementing new systems for research management to understand how the library can interface with the repository. We have done a short course to acquire new skills and competencies in RDM. Based practices workshops from institutions already offering services of RDM can also be used for new skills".*

Participant 14: "*Yes, I do need new skills which I believe I can acquire through workshops. I also think that I would need to equip myself through webinars, for example, the University of Pretoria RDM programme which is user-friendly and easy to understand with some available videos on RDM".*

Participant 15: "*I think yes to a certain degree for example, as metadata or cataloguing librarian you have some basic skills and knowledge but when it comes to metadata like how to store it, types of storage methods available and longevity of data. So, we should be skilled in that through training courses, webinars and learning from other institutions offering RDM services so that we can easily identify challenges and opportunities".*

5.2.1.4 RDM development strategies

5.2.1.4.1 RDM policy in the library

It was necessary to determine whether participating libraries currently have RDM policies in place to inform the institution's RDM agenda. The following are direct quotes from study participants.

Participant 3: "*There is currently no RDM policy in the library, however, we have been planning to have one as we have a quite close partnership with the research office. We had RDM training which we organised with the research office, unfortunately, everything stopped, and the library had to redefine its services due to covid-19 pandemic".*

Participant 12: "*Yes, the institution does have an RDM policy however, it has not been approved and it is called the DUT research data management policy. It basically will guide the implementation of research data management practices, it will set responsibilities in terms of the researchers, the library and the research office. Also, it will show relationships with our open access policy, institutional research policy and intellectual property policy".*

Participant 13: "*There is a draft policy that has not been approved yet. The name of the proposed policy is the DUT research data management policy. Well, it is expected to enforce the storage of research data and the implementation of research data management plans".*

Participant 14: "*The policy is still in the draft stage. The proposed policy name is the Durban University of Technology research data management policy, and it intends to provide a strategic framework that will facilitate the management of research data".*

Participant 15: "*The library, I think in conjunction with the research office drafted the policy which is still in the pipeline. I think it is called the research data management policy. It is aimed at highlighting what is expected of researchers in the institution".*

5.2.1.4.2 Internal and external capacity building programs for RDM staff skills development

It was necessary to find out about capacity-building programs available in participating institutional libraries because they are important in staff skilling and reskilling for better service provision. Below are some of the selected responses.

Participant 12: "*Not at the moment, but we will be offering capacity development in terms of data planning. There is a UCT short course on research data management for staff skills development".*

Participant 13: "*Nothing yet".*

Participant 14: "*There are currently no internal and external opportunities however, the DUT had previously invited UCT for a short course on RDM training and perhaps we can use this course as an external opportunity in future".*

Participant 15: "*There is nothing as yet internally but once we kick off with the RDM offering staff will be trained. Externally, we do have as we were a group of librarians who were allowed to attend the RDM course that UCT promotes".*

5.2.1.4.3 Resources or infrastructure for supporting RDM services and practices

The purpose of this question was to determine whether or not the participating institutional libraries have adequate resources or infrastructure to support RDM practices. The verbatim replies from respondents are presented as follows.

Participant 12: "*From human resources, currently I have one person as part of our planning, we also have an additional post that will be made available as research data management is rolled out in the institution".*

Participant 13: "*Looking at the RDM draft policy, I think we do have adequate resources it's just a matter of implementing. In terms of staffing, for now, we do have enough, however, the actual implementation of the RDM will determine if there is a need for new staff and skilling them. For example, there are two metadata/cataloguing Librarians responsible for the management of metadata for articles and thesis and are doing well, so if there will be a need for an additional staff member, it could just be one staff member".*

Participant 15: "*I think we do have sufficient resources, however, the actual implementation of RDM will act as a decider on whether there is a need for other resources to support RDM practices in the institution".*

Participant 2: "*We do not have the infrastructure and resources which makes us not to be ready to provide RDM services in our library".*

Participant 1: "*No, our library does not have the infrastructure and human resources for RDM services. The library still needs well-trained human resources who will be responsible for the support of RDM services provision and practices".*

5.2.1.5 Challenges facing public universities in managing research data

5.2.1.5.1 Challenges of offering RDM

Participants were asked to deliberate on the challenges they are facing regarding the provision of the services of RDM. Their responses are presented as follows.

Participant 1: "*As our library is currently not offering the RDM services, I would think one problem is technical infrastructure, also maybe the competent staff to start the process".*

Participant 2: "*We do not have the infrastructure and the resources, so we're still not yet ready. The whole process of RDM requires responsible personnel which is still problematic".*

Participant 3: "*We firstly need some sort of policy which is currently not in place. There is also no support from other stakeholders that is currently in place making it difficult to make any progress. A lack of funding and skilled labour is also problematic. I am also suspecting that it would be somehow difficult to convince the researchers to share their research data as they are having trust issues and don't understand open access correctly".*

Participant 12: "*I think from the department point of view we consolidated our infrastructure and human resources in terms of getting a single person to get research activities right across. Our current challenge is really in terms of getting our researchers on board. Our only interest is coming from persons who have NRF funding and wish to publish in a journal and the requirements of the finding include storing their research data in the institution. However, in the absence of the RDM policy, I think we have no point of direction. I think the major issue we have is adoption, getting people to adopt the culture of managing research data in the institution".*

Participant 13: "*I think the main challenge is that faculties seem to be resistant to it and inability to promote it nicely".*

Participant 14: "*I think the main problem that we could encounter is the lack of requisite skills to provide RDM services".*

Participant 15: "*The challenge we have encountered to not offering RDM services as yet is the reluctance of getting researchers on board".*

5.2.1.5.2 Possible ways to overcome the challenges of offering RDM

It was necessary to solicit suggestions from the participating institutions on possible solutions to the challenges that are impeding the provision of RDM services. The verbatim responses from study participants is as follows.

Participant 12: "*I think we must use research forums to engage postgraduates in capacity development programmes. We need to find ways of making this not a separate service but part of a complete research cycle".*

Participant 13: "*I think the policy will assist with the resistance once the policy has been approved and will therefore facilitate the entire process of RDM".*

Participant 14: "*I think training is much needed which will assist in informing and facilitating the roles and responsibilities of staff responsible in the provision of RDM services".*

Participant 15: "*Maybe we need to assure the researchers that their research data is going to be safe and show them the benefits of managing their research data".*

5.2.1.6 RDM framework for public universities in South Africa

5.2.1.6.1 Views on developing an RDM framework for public universities in South Africa

Participants in the study were asked for their opinions on the development of a standardised RDM framework for South African public universities. Their exact words are as follows.

Participant 1: "*I think the collaboration between universities across South Africa and relevant stakeholders".*

Participant 2: "*It would be much better if there was one common database where all the universities are depositing their research works. As all the universities in the*

country have varying goals and objectives regarding research practices, it is difficult to develop a uniform RDM framework. Therefore, for a uniform, RDM framework universities would have to discuss or collaborate on one common goal in developing a framework for public universities".

Participant 3: "*I am not sure who could take a lead on this if maybe we as higher education institutions could have one national policy or framework designed by the department of higher education which could guide the management of research data in public universities".*

Participant 12: "*As far as I know we have an open access kind of framework and institutions are collaborating at the moment for getting and creating a national database for theses and dissertations, so possibly could be an extension of that work which could perhaps help in the development of an RDM framework in the country".*

Participant 13: "*I think if we could have a body like Chelsea's research that could put together a list of resources or a group of people in universities who would curate a list of resources that could be the first step for the effective provision of RDM services in public universities other than anything else".*

Participant 15*: "I think public universities could create a learning plan for RDM. They could also form partnerships and collaborate in terms of sharing tools and skills needed for adopting RDM services, and also sharing knowledge of addressing encountered challenges in the provision of RDM. I believe this can then act as an initiative for the development of the RDM framework for our country's public universities".*

5.2.2 Data from teaching and research staff

As explained in Chapter four (section 4.5.2.1) of the dissertation, the fourth to sixth groups of participants were from teaching and research at the selected participating universities. The data from teaching and research, like the data from the library, was gathered by conducting online interviews via Zoom and Microsoft Teams with research directors, deputy deans research, deputy vice-chancellors research, and department heads, as was also done during the COVID-19 pandemic. NRF researchers are also part of the teaching staff, but their data is analysed in the following section (see section 5.3.1.) because it yields quantitative results. Research directors (RDs) were chosen for their active role in ensuring the ethical conduct of

research processes in universities by actively promoting and facilitating cutting-edge research, including collaborative interdisciplinary research that reflects the university's goals. Deans/deputy deans of research were chosen as operational heads of faculties to ensure that research matters are always on the right track by providing direction and resources, as well as ensuring that research is conducted in accordance with university policies and procedures or those of sponsoring agencies. Deputy vice-chancellors (DVC) research were chosen as research leaders from their respective institutions to guide research development. DVC research is well known for playing a critical role in providing leadership and guidance to universities to achieve their strategic priorities for research and innovation while focusing on quality assurance and enhancement. Heads of department (HODs) promote good practices while also ensuring the safety and well-being of both staff and students involved in research, knowledge use, and research exchange or sharing. Among other things, HODs ensure that decisions are made to strengthen the departmental research culture and philosophy. This section also presented data collected in lines of biographical details; RDM practices; competency requirements of researchers and staff responsible for research data management; RDM development strategies; challenges faced by the selected public universities in managing research data; and an RDM framework for South African public universities.

5.2.2.1 Biographical information

Table 5. 3: Biographical details

Participant	Gender	Age category	Institutional name	Academic qualifications	Job position	Work experience
P4	Male	45-54 years	UZ	PhD	HOD	3 years
P5	Male	55-64 years	UZ	Honours Degree	HOD	12 years
P6	Male	45-54 years	UZ	PhD	HOD	3 years
P7	Male	65 years and above	UZ	PhD	Senior Professor	3 years
P8	Female	25-34 years	UZ	PhD	Acting research director	5 months
P9	Female	55-64 years	UZ	PhD	HOD	4 years
P10	Male	35-44 years	UZ	PhD	HOD	5 years
P11	Female	45-54 years	UZ	Certificate in office management	Chief research administrator	15 years
P16	Female	45-54 years	DUT	PhD	Deputy Dean	2 years
P17	Male	55-64 years	DUT	PhD	HOD	1 year
P18	Female	45-54 years	DUT	PhD	HOD	3 years
P19	Female	45-54 years	DUT	PhD	HOD	5 years
P20	Male	45-54 years	DUT	PhD	HOD	9 years
P21	Female	45-54 years	DUT	PhD	HOD	2 years
P22	Male	35-44 years	DUT	PhD	HOD	4 months

Table 5.3 presents the biographical data of the study participants from teaching and research. Most (8, 53%) were male participants falling in the 45-54 age category (8, 53%). The African population group dominated the study participants. The majority (10, 67%) of the study participants held a Doctor of Philosophy as the highest academic qualification and had more than three (3) years of work experience (9, 60%).

5.2.2.1.1 Research-related duties

Participants were asked to identify the tasks they carry out in their respective positions that are related to research. Participants' responses are listed below.

Participant 4: "*One of my duties is to promote a culture of research by conducting workshops to assist novice researchers and supervise Master and Doctoral students. I also conduct workshops on article writing and also as a scholar I review research articles from other institutions".*

Participant 5: "*Well, it is to encourage and mentor mainly postgraduate students, but I don't supervise at all I never have. I also help in organising research workshops".*

Participant 6: "*As a HOD, I connect the research activities in the department, I try as much as possible to encourage staff in the department to publish research papers and I also supervise M and D students".*

Participant 7: "*Research Leader and Head of Centre".*

Participant 8: "*My duties are administrative, but as I am a researcher myself, so which means I also participate in conferences, publish research papers, supervise students, and also collaborate with colleagues within the faculty as well as other faculties".*

Participant 9: "*I mentor novice researchers and am also responsible for the supervision of postgraduate students".*

Participant 10: "*One thing that we are doing is to make sure that our academic members are capacitated and equipped with the necessary tools so that they can be actively involved in research-related duties. Lastly, we recruit students and ensure that they are on board as they are the main drivers of research".*

Participant 11: "*Well, when it comes to my duties, it is a wide range as it goes from the research claims that I deal with, to internal and external funding applications. I also deal with research publication output and all the incentives related to research publications, supervision incentives and NRF incentives. I am also part of the DHET UCDP project where we make funds available".*

Participant 16: "*In your HOD capacity, your research duties are to first of all be a driver for your department and to encourage staff in terms of their qualifications as the emphasis has moved to doctorate qualifications. The second is to be the quality control for your proposals and the implementation of the faculty research committee rules".*

Participant 17: "*So basically, I have to publish myself, attend conferences, inspire and motivate my department staff to publish, supervise masters and PhD students, encourage students to publish, and also increase the intake of masters and PhD students".*

Participant 18: "*I am supervising Masters and PhD students, I am also part of the research committee and responsible for reviewing research proposals".*

Participant 19: "*I am a research supervisor for postgraduate students, helping students to structure their topics and as a HOD I have to sign and allocate research supervisors to students".*

Participant 20: "*Admission of postgraduate students, supervision, examination, as well as organising publication workshops for students".*

Participant 21: "*Administrative duties, HR duties, financial duties".*

Participant 22: "*One, I am the departmental research committee chairperson, reviewing all research proposals in the department for undergrad and postgrad students".*

5.2.2.2 RDM services at the selected public universities

5.2.2.2.1 RDM services' provision and practices in support of the research data in the institution

This question sought to determine whether or not participating institutions offer RDM services. If not, participants were then asked if they provided any research-related activities or practices from their departments or institutions, and they responded as follows.

Participant 4: "*No, we don't provide RDM services, but we do data creation, data storage, sharing and mining".*

Participant 5: "*No, we do not offer RDM services as our department does not have one".*

Participant 6: "*For now, we don't, but we do data creation, preservation and sharing".*

Participant 7: "*Our department does not offer any RDM services or any research practices".*

Participant 8: "*None"*

Participant 9: "*None"*

Participant 10: "*I can mention data creation, obviously when we conduct research, we create data. We also do data sharing, and we share data in a form of conferences, publications and presentations. Data mining is also some of the things we do, and we do it in a form of extracting some valuable information using computer software. Access and re-use are also one of that we also do as we make information available for anyone and can be reused".*

Participant 11: "*At the moment we do not have RDM services, but we provide data creation, data storage, preservation and data sharing".*

Participant 16: "*The data creation, data mining, data access, use and re-use are what we provide other than the services of RDM".*

Participant 17: "*Data sharing, data preservation, data creation, data archiving and data security. We also have a process map which helps to track the lifecycle of a student. Through the use of the process maps, student supervisors, HODs, and deputy deans know exactly what process to follow in a lifecycle for example, if a student submits a proposal where does it go and when it is ready for exam submission what are the forms that need to be filled in, which systems does it flow through, and which entities does approval take place".*

Participant 18: "*We do practice data creation through our students who publish dissertations, and we also publish as well".*

Participant 19: *"We do store data for students who have passed which are their hardcopies dissertations in the department. However, the faculty is now in charge of*

the electronic and hard copies of dissertations. As the department, we use the library for data mining and data security".

Participant 20: "*Our department provides data creation and data sharing. The library is responsible for data storage".*

Participant 21: "*In terms of postgraduate research, we keep their data in the department by the research coordinator".*

Participant 22: "*We do have some storage for proposals that we have received, for example, we have got a file to say 2020 theses are all the files (proposals) we received, these are the ones that were approved and out of those these are the ones that registered. So, we do have some storage that we use but it is not automated. Each faculty in the institution has a faculty officer who receives all the completed students' theses and dissertations and captures them and then gets submitted to the library".*

Participant 23: "*Only one department 'planning' has started RDM in the institution (UNIZULU), however very few know about it at the moment".*

5.2.2.2.2 Institution's storage, sharing and management of published research data

Participants were asked where in the department published research works of researchers are kept and how long the research data is kept. The verbatim selected responses are as follows.

Participant 4: "*It is kept at the research office as it is a university's intellectual property and on the departmental side if you have collected data, we cannot save it for more than a year".*

Participant 5: "*There isn't a central register so the individuals who have published, the record is kept in the journals that they publish in".*

Participant 6: "*There is a database in the library where research data is kept".*

Participant 7: "*Electronic (Lib archive), Centre's Web Site and we don't keep research data, we archive publications on the Research Gate Website".*

Participant 8: "*In the research office there is storage where research data is kept, and administrators are the ones responsible for keeping data".*

Participant 9: "*Data is kept at the research office, and it is kept for 3 years".*

Participant 10: "*I cannot say we keep it ourselves per se in our department as the research office is mostly involved in keeping research data".*

Participant 11: "*In the years before covid-19, we only worked with paper and it was stored in the cabinet, but now since covid, we do everything electronically. So, what I do is electronic records of the published work. Research data in my section is kept for 10 years and more".*

Participant 16: "*All the published works itself is kept in the library in the research repository. It has got absolute open access to everyone and is not controlled by the department. The actual thesis or dissertation of a student is kept lifelong and the actual data that was used by the student or staff member must for 5 years that is as per the institutional research ethics committee".*

Participant 17: "*Published works in terms of students' dissertations are kept in the library and the institutional repository. The research work is there for all time and is not removed".*

Participant 18: "*It is kept in the institutional repository in the library, and it is kept for 5 years".*

Participant 19: "*The published works of researchers are kept in both the library and the faculty. Usually, they say data is kept for 5 years but the faculty keeps it longer than 5 years, so I would say it is kept forever".*

Participant 20: "*Research works are kept in the institutional repository, and I think there is no time for keeping data, it becomes available forever unless the organisation insists to keep it for 5 years, but it also depends on the nature of the research".*

Participant 21: "*We keep their documentation with us for five years and we keep it in the department by the research coordinator".*

Participant 22: "*The hardcopies, the department gets to keep one and one goes to the library. But then now, we have got soft copies, the library gets to keep all the soft and hard copies so if one needs to access it, they need to go to the library. The published work gets to be kept forever as hard copies and soft copies".*

5.2.2.2.3 Requirements of the researchers to publish their research

Participants were asked to state the requirements for the researchers' research to be published and they responded by articulating the following.

Participant 4: "*You must have a project and a project must have an ethical clearance".*

Participant 8: "*One needs to have research skills, have a mentor who will assist in writing a paper and attend workshops".*

Participant 10: "*I think when one publishes their work, they need to look at the quality of the work that is being conducted, its relevance, the target audience, whether has it ever been published before and is the work not plagiarised as this is what the publishing houses will require. By looking at these requirements, it makes it easier for your paper to be accepted by the auditors".*

Participant 16: "*I think that in the first part to publish the research data you need to make sure where the research data originates from. If you are using the DUT platform, then you need to get approval from the institutional research ethics committee (IREC) and you can only proceed with your study once the approval has been obtained. So, IREC is not just for students that are studying at DUT but for any project".*

Participant 18: "*The research work has to be aligned with the requirements of the university, and the journal which a researcher will publish".*

Participant 19: "*The research proposal has got to go to the institutional research ethics committee for ethical clearance because without the clearance you cannot be able to pursue whatever you want to do, whether you want to study or want to produce an article, it has to go through the IREC".*

Participant 20: "*At DUT you need ethics clearance especially if you are writing about the university or vulnerable groups you need ethics clearance. You also need a gatekeeper's letter from other institutions where you are collecting data so that you can publish that data".*

Participant 22: "*With us, you must have ethical approval, without any ethical approval you cannot publish that and align it with the institution. So, if you have gone*

and done data collection without ethical approval then the institution will not acknowledge the publication".

5.2.2.2.4 Researchers' awareness of the publishing platforms and requirements

Participants in the study were asked to express their opinions on whether they thought researchers, in particular postgraduates, were aware of the places they should publish their research data and the steps they should take to have their work published. The responses are listed below.

Participant 4: "*Yes, it is very clear for emerging researchers as on the workshops that one is conducting, we stress that firstly, if you want to publish you must look for accredited journals and the university provides academics with a list of the accredited journal each and every year. We also stress that you must study the requirements and guidelines of the journal you are publishing on".*

Participant 5: "*Accounting students at Unizulu do not know".*

Participant 6: "*From my understanding, I think they are upcoming and not there yet and still need to be fully assisted".*

Participant 7: "*Yes, they are aware".*

Participant 8: "*Most students rely on their supervisors to guide them as to where they can publish their data".*

Participant 9: "*No, they do not know".*

Participant 10: "*Some do, and some don't, especially with the emerging researchers. There are cases where you find a person completing up to PhD without a single paper and that is quite against the policy of the institution and that lies mostly with the supervisor concerned who is the main grooming of the postgrad student or the emerging researcher".*

Participant 11: "*Yes, they are aware that they need to publish especially when it comes to PhD. Students are supposed to be guided by their supervisors and in cases where the supervisor cannot, I then give guidance".*

Participant 16: "*I think in our faculty we have made progress in terms of that in the last two years. We are offering a whole number of webinars on how to write research*

papers every year. However, before that, there was a gap and I think that gap was institutional-wise where the emphasis was just on getting the doctorate or the master's and it was not on actually the publication part".

Participant 17: "*No, they wouldn't know if you ask them because they haven't dealt with it before publication and they need guidance from supervisors to identify possible journals where to publish".*

Participant 18: "*Yes, they are aware because there are workshops conducted".*

Participant 19: "*No, they are not aware because what I have realised is that if the student has struggled so much to attain the master itself the student does not want to look back to that information he has collected. I assume it is because of the uneasy journey they have endured so they don't want to go back and look at their work".*

Participant 20: "*Yes, they are aware, there are a lot of workshops conducted by our faculty to assist emerging researchers with information regarding where they can publish their research work".*

Participant 21: "*A lot of the time students are naïve and have absolutely no idea. So, it is the supervisor who makes the effort to help the student publish and together they source accredited journals and look at the guidelines together and work from there totally afresh".*

Participant 22: "*I would yes because the institution does host different seminars and workshops for postgraduate students where they get informed about all of this. They also have a session with the library as postgraduate students through workshops on how to use different things within the library which include the available repository, the requirements for them to publish and how to identify the journals that are accredited by the department of higher education and training. The supervisors would also share such information during the research process because we do not run coursework masters it is only pure research so the supervisors would assist the students in these elements".*

5.2.2.2.5 Benefits of RDM for emerging researchers

It was necessary to understand the benefits of RDM for emerging researchers in universities. The study participants felt as follows.

Participant 4: "*It helps in data sharing and preservation".*

Participant 5: "*I think it is probably good from the benchmarking perspective so that they know what they are expected to do and how they are expected to perform".*

Participant 7: "*Easy and Open retrieval process or mechanism".*

Participant 11: "*Students can retrieve data anytime they want it".*

Participant 17: "*Basically, RDM is there for tracking in terms of students' research areas and capturing the titles and the studies".*

Participant 18: "*It can encourage knowledge sharing".*

Participant 19: "*It is very important in terms of data sharing as other people will be able to see it since it will be in the public domain and use your information and build up on your information".*

Participant 20*: "It is very important because, to me, the benefit is that it serves as a reference point you can always go back to that data and refer to it and if you are doing a comparative study and if the data is available then you can always look at what happened in the past, what is happening now and maybe predict what may happen in the future".*

Participant 21: "*It does encourage knowledge sharing because it goes to the kind of work that has been done and the outcomes in terms of what the student has done".*

Participant 22: "*It becomes easy for students to access information if it is automated or available in a cloud space as it helps students to see what proposals have been approved similar to their topics and see if the area of interest has been researched. Also, if there are similar topics as someone else, they can see what recommendations and limitations are, and they can identify if there is an angle where they can also beat their study within that angle".*

5.2.2.3 Competency requirements of staff responsible for research data management in selected public universities

5.3.2.1 Knowledge and skills required to support RDM services

Participants were asked to identify the knowledge and skills required to support RDM service delivery in their institutions. Their responses are interpreted as follows.

Participant 5: "*I think IT skills are the most important skills to have, followed by database management skills and skills to search for information".*

Participant 11: "*In my position, office administration skills are very important in delivering the services of RDM".*

Participant 18: "*You need research skills; you need to know how to conduct research and you also need good communication skills to enable you to guide students".*

Participant 20: "*Well, I think it is research skills, and organising skills because there is no way you can manage research if you lack organisational skills, writing skills as well as communication skills".*

Participant 22: "*I think it depends on what system is being used, if it is an online system, you just need computer skills to understand and follow. When it is postgrad study then you need to know how you are going to upload the completed theses and dissertations into the library repository. But I think it is computer skills".*

5.2.2.3.1 Need for new skills and other competencies of RDM and ways to acquire them

Participants in the study were asked if they needed new skills and other related competencies to support the provision of RDM services and how they think they would acquire them. Their responses are as follows.

Participant 4: "*I would say I need new skills through workshops".*

Participant 5: "*As a person, you need to keep your knowledge and skills updated especially IT skills because of how technology is advancing. These can be acquired through attending workshops, webinars and training. However, the best way to do this is through practice because you can go on a 4-day workshop and come back you have forgotten all of those things within a week, so the best way to build capacity is through practice".*

Participant 8: "*Yes, I do need new skills because learning is an ongoing process, you know you never go wrong in getting some new skills and knowledge. I am always willing to learn, and I can acquire this through workshops, short courses and training".*

Participant 21: "*I would like to be able to retain or keep data at a departmental level and to make that data readily available to emerging researchers. I think I can acquire knowledge and skills to perform this through training, workshops and seminars and especially through conferences because these are some of the things that are discussed at conferences".*

Participant 22: "*Yes, I do need skills and be trained on how to use the RDM system. Training is required in the initial phases and then having a support unit that you could always go to when you run into clichés. YouTube videos could also support or assist in explaining the RDM system".*

5.2.2.4 RDM development strategies

5.2.2.4.1 RDM policy in a department or institution

To inform the institution's RDM agenda, it was necessary to ascertain whether participating institutions currently have RDM policies in place. The following quotes were provided by study participants.

Participant 4: "*No, we do not have an RDM policy".*

Participant 5: "*No, we don't have any".*

Participant 7: "*Now we don't".*

Participant 8: "*There are so much research-related policies, about 8 or 9 policies, but when it comes to data management, I am not sure if we have it".*

Participant 9: "*No".*

Participant 10: "*No, in our department we do not have an RDM policy".*

Participant 11: "*No, we just have research policies".*

Participant 16: "*As far as I know, the policy is still in the pipeline, and it is called the research data management policy".*

Participant 17: "*Currently there is no specific RDM policy but there are several research policies that we have".*

Participant 19: "*No, we do not have a departmental policy, but it is institutional, unfortunately, I do not know the name of the policy".*

Participant 20: "*No, our department does not have an RDM policy".*

Participant 23: "*No, the institution (UNIZULU) does not have RDM policy, but it defines its research practices through using related policies like research ethics policy".*

5.2.2.4.2 Available internal capacity building programmes for staff skills development in RDM

The study participants were asked if there are any capacity-building programmes within their institutions for staff skills development in RDM. The responses are as follows.

Participant 20: "*Yes, we do have a lot of these programmes and these programmes are run by the institutional research office and faculty research office. They address activities related to research in general and also highlight areas in relation to data management, but they are not specific to RDM".*

Participant 17: "*We do have internal writing and publication workshops in DUT related to research but the service provider for these could be internal or externa" l.*

Participant 16: "*We have writing retreats at the moment to encourage journal writing, we have webinars for each department throughout the year to encourage networking and different research perspectives, we have got workshops where the NRF is invited to inform staff of the process, and there are continuous workshops on ethics by the institutional research ethics committee".*

Participant 8: "*Training is mostly focused on those who are dealing with research for instance at the faculty level some people are responsible for capturing data and a chief research administrator they do get training for capacity building".*

Participant 5: "*Yes, we do have such programmes and they are offered all the time. The department and the faculty do run seminars on research tools that can be used but not RDM specifically".*

5.2.2.4.3 Resources or infrastructure for supporting RDM practices

This question was designed to ascertain whether the participating institutions have the necessary infrastructure or resources to support RDM practices. Following is a presentation of the respondents' verbatim responses.

Participant 4: "*No, we would need resources to offer research data management services".*

Participant 5: "*We don't have enough resources and infrastructure".*

Participant 6: "*For now, we don't have enough resources, we are not there yet".*

Participant 7: "*Yes, we do have enough resources".*

Participant 8: "*Well, I wouldn't say we have sufficient because even if it comes to human resources, we do have people at a faculty level and the research office but are not enough. They also need that capacity building now and again".*

Participant 9: "*I am not sure".*

Participant 11: *No, we don't, there is a lack of staff.*

Participant 16: "*I don't think so. At the moment we are still using a paper-based system in most instances where if it goes online completely it would be a very different environment. The electronic part is that it goes via email, and the rest of the documentation is done on paper".*

Participant 17: "*Not really, we do not have enough staff to offer RDM services".*

Participant 19: "*The department doesn't, but the library does so again the library is the centre for everything".*

Participant 20: "*Yes definitely, because we do have a research budget and can use that. On the human resource part, we are lacking we do not have sufficient staff".*

Participant 21: "*No, I don't think we do, we battle and are under a lot of pressure. We don't have enough supervisors in some departments for PhD which is a requirement at DUT. If you have a PhD and you supervise additional students, it is pressurizing when it comes to monitoring and follow-ups".*

Participant 22: "*We do not have adequate infrastructure. I think that staffing is where the issue might be. We need more staff in the research office to make sure*

that this is efficient, and it works out efficiently. Currently, it is handled by one person for the faculty and if all the proposals go to them it becomes problematic, so I do think we need more human resources or more hands-on with this portfolio".

5.2.2.5 Challenges facing the selected public universities in managing research data

5.2.2.5.1 Challenges of offering RDM

Participants were asked to identify challenges they face which prevent their departments or institutions to offer RDM services. Their responses are presented as follows.

Participant 4*: "I would say the research office at the top level is not coming down to people and there is no collaboration between the research office and the departments".*

Participant 5: "*It is purely a matter of staffing, and resource allocation in terms of additional costs involved and who would allocate that responsibility to and who would cover that cost".*

Participant 6: "*The major challenge is the lack of funds".*

Participant 9: "*Lack of competent staff in research positions at this institution (Unizulu)".*

Participant 11: "*It is mostly insufficient staff, lack of resources and lack of providing necessary information to the research office from the research community".*

Participant 19: "*There is not enough staff, and the capacity does not meet the requirements for offering RDM services."*

Participant 21: "*It is HR and staffing matters as we do not have the capacity for administrative additional roles and responsibilities. We are short-staffed and the current staff that are in the department are dealing with an extensive amount of work overload at the moment. If you had to look at research, then you would need someone with research skills to take over a position in the faculty and handle the research data collection and also to look at someone who will be ideal to oversee this kind of thing".*

Participant 22: "*The challenge that prevents us not to provide services of RDM in my department is that it is not offered at an institutional level and then within the institution, I think it is because of the lack of infrastructure and the human wheel".*

Participant 22: "*Lack of awareness, shortage of staff to support RDM and collaboration in the entire university are greatest hindrances of RDM service in the institution".*

5.2.2.6 RDM framework for public universities in South Africa

5.2.2.6.1 Views on developing an RDM framework for public universities in South Africa

Participants in the study were asked for their opinions on the development of a standardised RDM framework for South African public universities. Their exact words are as follows.

Participant 7: "*All University Research offices must have the means to archive research data and easy mechanisms of retrieval. They must also make sure it is an automated and intuitive software system to use".*

Participant 11: "*I think it is very important for institutions to collaborate and brainstorm together to see what each institution is doing. If we learn from each other, we can then agree upon a data management system framework so that all the institutions use the same system. If all institutions use the same system, it would be so much easier for everybody to track back the information that was provided and data management would be easy and accessible to everybody".*

Participant 16: "*The initiative to develop an RDM framework must be made by DHET so that they provide the funding to assist institutions. There should be one system that talks to everybody so that the data that you are receiving can be centralised and can be made available to all".*

Participant 17: "*I think universities would need to embrace technologies so that we have a robot system in place to do things in a few minutes which normally takes us many hours to do as we have a lot of students".*

Participant 18: "*I think collaboration is the key because even those universities that do not produce, can learn from those who have a high standard in terms of research".*

Participant 19: "*I think the DHET can develop this framework as they develop most of the frameworks for higher education. Therefore, the universities cannot do it on their own because they are leaning toward the DHET and DHET is in the custody of all major policies then institutions extract the main policies from the head (DHET). So, whatever policies we have as institutions are aligned with the DHET policies".*

Participant 21: "*I think that the universities should get together, the Deputy Vice Chancellors in charge of these across the universities in the country get together to look into how the best efforts can be made to develop a framework".*

Participant 22: "*One, it is human capital, we need more staff to make sure that the research area gets the adequate attention it requires. We need resources in terms of money to procure the software and the personnel to manage this software once they exist. We also need to change or shift the mental space that we are in and acknowledge that certain things are now done online and align it with what is currently happening in other places. RDM is something we would need for the institution or improve the current research activity and if we are to improve our outputs as universities in South Africa, we need a functional system to manage postgraduate activities or studies".*

Participant 22: "*Collaboration and discussions between departments and faculties and other relevant stakeholders is really required".*

5.3 Data presentation and analysis_Quantitative responses

5.3.1 Data from NRF-rated researchers

As explained in Chapter 4 (section 4.5), the last population of participants were NRF-rated researchers from both the University of Zululand and the Durban University of Technology. It was assumed that this group would interact with research data more than other university researchers and thus provide more reliable information for the study. The study included all NRF-rated researchers from the institutions, so there was no sampling. Originally, it was intended to physically reach this population

group; however, because the COVID-19 pandemic and its associated restrictions and regulations were still in effect during this study, research questionnaires were converted into Google forms and sent via email (see Appendix C). 21 survey questionnaires were emailed to all the NRF-rated researchers from the University of Zululand and 39 mailed questionnaires to Durban University of Technology NRF-rated researchers. The respondents had two weeks (10 working days) to complete the questionnaires. However, due to the pandemic, the majority of study participants worked from home, resulting in a very low response rate. Low response has also been noted by Shareia (2016: 3842) to be one of the main problems of using questionnaires. As a result, the period was extended, resulting in a significant delay in the analysis of the data gathered through questionnaires. Thirty (30) of the sixty (60) targeted participants responded, yielding a 50% response rate, which is considered adequate for analysis (Neuman, 2000). Neuman (2014: 269) suggests that the size of the sample in research is determined by population characteristics, the type of data analysis to be used, and the level of confidence in sample accuracy required for research purposes. Neuman (2014: 270) goes on to say that for small populations (under 500), a high sampling ratio (around 30%) is required. The number of mailed questionnaires, as well as the percentage of response rate per participating institution, are shown in Table 5.4.

Table 5. 4: Response rate

Institution	**No of targeted participants N=60**	**No of the participants responded N=30**
University of Zululand (UZ)	21	12
Durban University of Technology (DUT)	39	18
Total	**60**	**30**

5.3.1.1 Biographical data

Table 5. 5: Biographical details

Institution	Gender	No	%	Age category	No	%	Highest academic qualification	No	%
University of Zululand	Male	10	33%	18-24 years	0	0%	PhD	12	40%
	Female	2	7%	25-34 years	0	0%			
				35-44 years	5	17%			
				45-54 years	3	10%			
				55-64 years	2	7%			
				65 and over	2	7%			
Institution	**Gender**	**No**	**%**	**Age category**	**No**	**%**	**Highest academic qualification**	**No**	**%**
Durban University of Technology	Male	9	30%	18-24 years	0	0%	PhD	18	60%
	Female	9	30%	25-34 years	1	3%			
				35-44 years	5	17%			
				45-54 years	4	13%			
				55-64 years	4	13%			
				65 and over	4	13%			

The biographical data of the study participants of the NRF-rated researchers from both the University of Zululand (UZ) and Durban University of Technology (DUT) are presented in Table 5.5 above. The majority (19-63%) were male participants aged 35-44 from both participating institutions. The African population dominated the study's population. The highest academic qualification held by all (30-100%) of the study participants was a PhD.

5.3.1.2 RDM services

5.3.1.2.1 Awareness of the term ”Research Data management”(RDM)

Participants in the study were asked if they were familiar with the term "research data management." The majority of study participants (25-83%) agreed that they are aware of the term, while only a few (5-17%) were not.

5.3.1.2.2 RDM services offered in participating institutions

The study participants were asked to indicate if their institutions offer any research data management services. The results are presented in Figure 5.1.

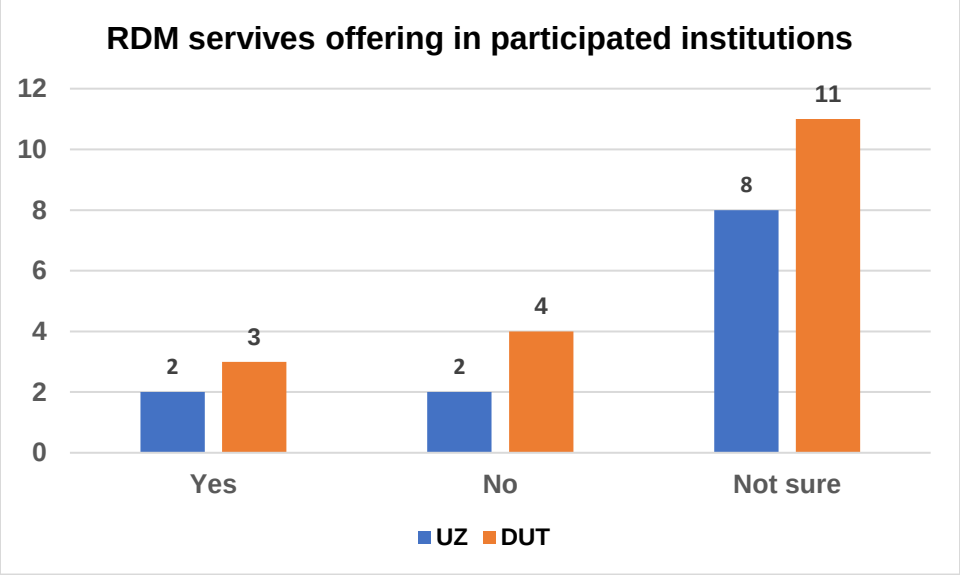

Figure 5. 1: RDM services offered in participating institutions

The majority of study participants, 19 (64%), from both participating institutions (UZ and DUT), were unsure whether their institutions provide RDM services or not, while 6 (20%) indicated that their institutions do not provide RDM services and only 5 (17%) agreed that their institutions do.

5.3.1.2.3 Submission of RDM plan for research activities

Participants in the study were asked if they had ever submitted an RDM plan for their research activities. Most study participants, 19 (64%), indicated that they had not submitted their RDM plans for their research activities, while only a few (11-36%) confirmed that they have submitted their RDM plans.

5.3.1.2.4 Types of research data generated in undertaking research-related activities

The study participants were asked to identify the type/s of research data they have generated in undertaking their research-related activities. The results are presented in Table 5.6.

Table 5. 6: Types of research data

Types of research data	Frequency	Percentage (%)
Documents (text, MS Word, PowerPoint, PDF, etc.)	24	80%
Spreadsheets	16	53%
Questionnaires	15	50%
Images	14	47%
Databases	13	43%
Dairies/ notebooks	10	33%
Observations	8	27%
Verbal context/ audios	7	23%
Websites	7	23%
Scanned documents	6	20%
Video recording	2	7%
Binary/histograms	1	3%
I no longer undertake research	1	3%
Musical compositions	0	0%

Documents (text, MS Word, PowerPoint, PDF, etc.) are the most commonly generated research data type (80%), followed by spreadsheets (53%), and questionnaires (50%).

5.3.1.2.5 The amount of research data generated from past research activities

Participants in the study were asked to estimate how much research data they generated while conducting research-related activities. The findings are summarised in Table 5.7.

Table 5. 7: Amount of research data

Amount of research data	Frequency	Percentage (%)
Not sure	16	53%
1-100 gigabytes (GB)	8	27%
1 Terabyte (TB)	5	17%
1-500 Megabytes (MB)	1	3%
Total	**30**	**100%**

The majority of study participants (16-53%) were unsure how much research data they generate while conducting research-related activities.

5.3.1.2.6 Personal storage of research data

The study participants were asked to identify their personal data storage facilities for their research data. The results are shown in Table 5.8.

Table 5. 8: Research data personal storage facilities

Storage facility	Frequency	Percentage (%)
Laptop/personal computer	28	93%
Emails	22	73%
External storage devices e.g, USB, external hard drive, CDs, etc.	19	63%
Clouds (e.g.Google Drive)	13	43%
Institutional repository	3	10%
Personal websites	1	3%

Laptops or personal computers are used by the majority (28- 93%) of study participants, followed by those using emails (73%), and external storage devices (63%).

5.3.1.2.7 Research data backup

Participants in the study were asked if they backed up their research data. The majority of the participants (26- 87%) stated that they back up their research data, while very few (4-13%) do not back up their research data.

5.3.1.2.8 Frequency of research data backup

This question sought to ascertain the frequency with which study participants back up their research data. Table 5.9 summarises the results.

Table 5. 9: Frequency of research data backup

Time span	Frequency	Percentage (%)
Monthly	7	23%
Never	7	23%
Annually	5	17%
Quarterly	4	13%
Daily	4	13%
Weekly	3	10%
Total	**30**	**100%**

The majority of study participants (7-23%) back up their research data monthly and a similar portion (23%) never back up their research data. Very few (3-10%) back up their research data on a weekly basis.

5.3.1.2.9 Back up storage facility for research data

This question sought to ascertain the research data storage facilities for researchers. Table 5.10 summarises the findings.

Table 5. 10: Back up storage facilities for research data

Storage facility	Frequency	Percentage (%)
External storage devices e.g. USB, external hard drive, CDs, etc.	23	77%
Laptop/personal computer	15	50%
Emails	12	40%
Clouds (e.g. Google Drive)	12	40%
Institutional repository	0	0%
Personal websites	0	0%

For research data backup, the majority of study participants (23-77%) use external storage devices such as USBs, external hard drives, and CDs, followed by half (15-50%) using laptops or personal computers. Institutional repositories and personal websites are never used.

5.3.1.2.10 Other responsible persons for research data management and its backup in the institution

Participants in the study were asked to identify other individuals within the institution who are responsible for the management of research data and its backup. The results are presented in Table 5.11.

Table 5. 11: Other responsible persons for research data backup

Responsible persons	Frequency	Percentage
Not sure	14	47%
My department	6	20%
Research office	6	20%
External research partners	4	13%
Technicians	4	13%
Librarians	3	10%
Third-party data centre	2	7%
I take the cloud to be my automatic back-up	1	3%
Nobody	0	0%
Postdocs	0	0%

The majority (14- 47%) of the study's participants reported being not sure of other persons responsible for research data management and its backup in the institution.

5.3.1.2.11 Period of research data storage within the institution

This question sought to ascertain the period of research data storage within the institution. Figure 5.2 summarises the results.

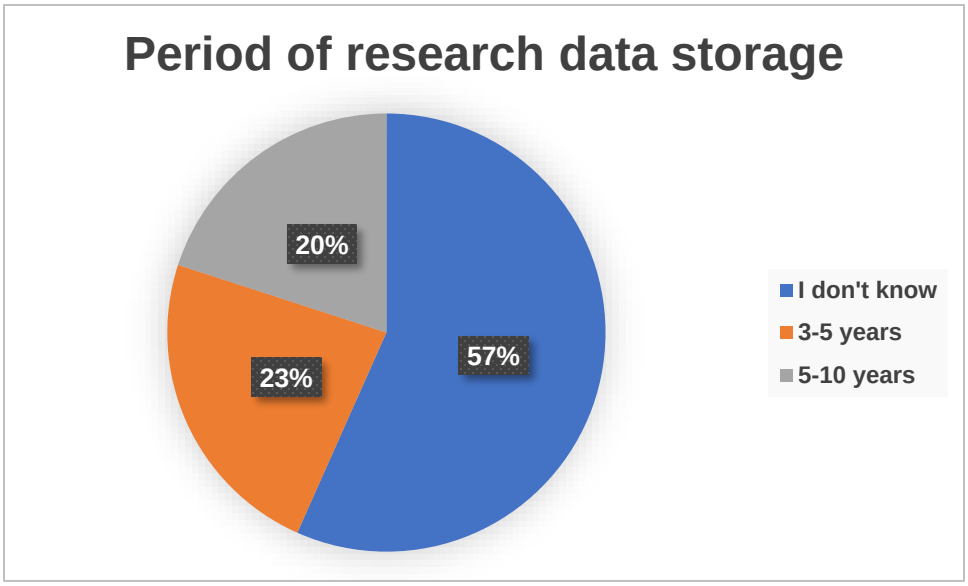

Figure 5. 2: Period of research data storage

Figure 4 shows that the majority (17-57%) of study participants are unsure of the duration of research data storage within their institutions.

5.3.1.2.12 RDM-related activities provided by the institution

Participants in the study were asked to identify RDM-related activities offered by their institution. Table 5.12 summarises the responses.

Table 5. 12: RDM-related activities

RDM related activities	Frequency	Percentage (%)
Grant applications	14	47%
Support for collaborative research	13	43%
Data storage	5	17%
Support for data analysis and computational analysis	4	13%
Support for storage and high speedy connectivity	4	13%
Guidance on RDM	2	7%
I don't know	2	7%
Data management planning	1	3%
Data preservation	1	3%
I am no longer involved in research	1	3%
Data curation	0	0%

The majority of study participants (14-47%) mentioned grant applications, followed by support for collaborative research (13- 43%) as RDM-related activities offered in their institutions.

5.3.1.2.13 Research data services offered by the institution

Participants were asked to identify RDM services provided by their institution. The responses are summarised in Table 5.13.

Table 5. 13: Research data services

Research data service	Frequency	Percentage (%)
Providing reference and advisory support to researchers on RDM	13	43%
Data citation	11	37%
Data sharing and access	8	27%
Providing technical support for research data service systems (repository, web portals, creating metadata, RDM system, access and discovery systems)	5	17%
None	4	13%
Data storage and backup planning	3	10%
Creating web guides and finding aids for data repositories	3	10%
Organising students' workshops on RDM-related activities	2	7%
Not sure	1	3%
Coordinating RDM advocacy programs between the researchers and the stakeholders	1	3%
Directly participating with Librarians and IT specialists on RDM	0	0%
Provision of RDM literacy programs	0	0%
Organising workshops, seminars, and direct training on RDM	0	0%

The most prominent (13-43%) identified RDM services provided by participating institutions were reference and advisory support to researchers on RDM and data citation (11- 37%).

5.3.1.2.14 Research data sharing

The study participants were asked if they have ever shared their research data and the majority (25- 83%) agreed that they have previously shared their research data, while very few (5-17%) never did.

5.3.1.2.15 The platform for research data sharing

This question sought to determine the platforms previously used for research data sharing. Table 5.14 shows the results.

Table 5. 14: Platforms for sharing research data

A platform for sharing research data	Frequency	Percentage (%)
Academic journals	23	77%
Conference presentations	19	63%
Book chapters	18	60%
Conference proceedings	17	57%
Email	9	30%
Hard copy	5	17%
External storage devices such as USB	3	10%
Media	3	10%
Other: I have an interactive web page that offers information to viewers	1	3%
Institutional repository	0	0%

The majority of participants (23-77%) use academic journals as research data-sharing platforms, followed by conference presentations with 19 (63%). On the options for other, one study participant indicated having an interactive web page that offers information to viewers.

5.3.1.2.16 Relevance of data sharing

The study participants were asked whether they thought data sharing was important or not. The majority of study participants (28-93%) felt data sharing to be important, while very few (2- 7%) perceived it not to be.

5.3.1.2.17 Motivations for data sharing

It was significant to obtain some motivations for data sharing from study participants. Table 5.15 summarises the findings.

Table 5. 15: Motivations for data sharing

Motivation	Frequency	Percentage (%)
Increases the impact and visibility of research	23	77%
Encourages scientific enquiry and debate	22	73%
This leads to new collaborations between data users and data creators	21	70%
Promotes innovations and potential new data uses	20	67%
Enables scrutiny of research findings	19	63%
Reduces the cost of duplicating data collection	14	47%
It creates awareness	1	3%
It could lead to plagiarism	1	3%

The majority of study participants (23-77%) agreed that data sharing is important because it increases the impact and visibility of research, followed by (22- 73%) vouching for scientific enquiry and debate and data sharing leading to new collaborations between data users and data creators by (21- 70%). Only a small percentage of study participants (1–3%) believed that data sharing could raise awareness and lead to plagiarism.

5.3.1.3 Competency requirement of researchers for research data management at the university of Zululand

This section deliberates on the researchers' competencies in terms of knowledge and skills in the management of research data.

5.3.1.3.1 Types of knowledge possessed by researchers in relation to the management of research data

The respondents were asked to determine the skills they have in relation to the management of research data. The results are presented in Table 5.16.

Table 5. 16: Knowledge of managing research data

Knowledge	Frequency	Percentage (%)
Ethics, including citation of data	20	67%
Data sharing	15	50%
Data management and organisation	11	37%
Discovery and acquisition of data	11	37%
Databases and data formats	10	33%
Quality assurance	9	30%
Data preservation	9	30%
Database development	6	20%
Data conversion and interoperability	4	13%
Cultures of practice	4	13%
Data curation	3	10%

Table 5.16 shows that most study participants (20- 67%) knew ethics, including citation of data, followed by knowledge of data sharing with 50%.

5.3.1.3.2 Knowledge gap in terms of RDM

Study participants were asked if they had ever noticed a knowledge gap in the management of their research data. Results are presented in Table 5.17.

Table 5. 17: Knowledge gap in managing research data

Knowledge gap	Frequency	Percentage (%)
Data management and organisation	16	53%
Database development	14	47%
Databases and data formats	13	43%
Quality assurance	10	33%
Data preservation	10	33%
Data curation	10	33%
Discovery and acquisition of data	9	30%
Cultures of practice	9	30%
Data conversion and interoperability	8	27%
Data sharing	5	17%
Ethics, including citation of data	2	7%

The majority of study participants (16- 53%) demonstrated a knowledge gap in data management and organisation, followed by a knowledge gap in database development (14- 47%).

5.3.1.3.3 Types of skills possessed by researchers in relation to the management of research data

The respondents were asked to determine the skills they have in relation to the management of research data. The results are presented in Table 5.18.

Table 5. 18: Skills in managing research data

Skills	Frequency	Percentage (%)
Data analysis skills	25	83%
Searching and retrieval skills	20	67%
Data capturing skills	18	60%
Collaboration and communication skills	16	53%
Sharing skills	13	43%
Software skills	8	27%
Data preservation skills	7	23%
Data visualisation skills	7	23%
Metadata skills	5	17%
Data curation and re-use skills	3	10%
Database development skills	2	7%

The majority of study participants (25- 83%) agreed to have data analysis skills, followed by searching and retrieval skills (67%) and data capturing skills (60%).

5.3.1.3.4 The skills gap in terms of the RDM

Respondents were asked if they had ever noticed a knowledge gap in the management of their research data. Table 5.19 displays the results.

Table 5. 19: Skills gap in managing research data

Skills	Frequency	Percentage (%)
Database development skills	16	53%
Data curation and re-use skills	13	43%
Software skills	13	43%
Metadata skills	12	40%
Data visualisation skills	11	37%
Data preservation skills	10	33%
Searching and retrieval skills	7	23%
Data analysis skills	6	20%
Sharing skills	6	20%
Data capturing skills	5	17%
Collaboration and communication skills	3	10%

The dominant skills gap was in database development (16-53%), followed by data curation and re-use skills as well as software skills by 43%.

5.3.1.3.5 Attendance of any research data management training in the institution

The respondents were asked if they had ever participated in RDM training at their institutions. The majority of study participants (18-60%) had never received RDM training at the institution and few (12- 40%) had.

5.3.1.3.6 Areas to receive RDM training

Study participants were asked to indicate the areas they feel they need training on concerning the management of research data in the institution. The results are presented in Table 5.20.

Table 5. 20: Areas to receive RDM training

Area	Frequency	Percentage (%)
Archiving data	12	40%
Data curation	12	40%
Data mining	10	33%
Developing RDM plan	10	33%
Sharing data	10	33%
Formatting data	9	30%
Storing data	9	30%
Ethical issues	8	27%
Documenting data	8	27%
None	8	27%
Copyright issues	7	23%
Intellectual Property Rights	7	23%

The majority (12-40%) of study participants indicated both a need for training on data archiving and data curation.

5.3.1.3.7 Practices that could be used in the skills development of RDM in the institution

Table 5.21 shows the results of a survey that asked study participants to identify practices for RDM skill development using a Likert scale.

Table 5. 21: Practices of skills development in RDM

Practice	Agree	Strongly agree	Moderate	Low
Attending workshops	13	11	5	1
Attending conferences	11	7	4	8
Provision of short courses on RDM	12	14	3	1
Attending training provided by professional bodies	12	9	6	3
Other, please specify ……..				

Most study participants vouched for the provision of short courses on RDM as the best method for RDM skills development.

5.3.1.4 RDM development strategies in universities

This section deliberates on participating institutional policies, infrastructure and capacity-building programmes supporting the delivery of RDM services.

5.3.1.4.1 Policies supporting practices of managing research data in universities

Study participants were asked whether there are any policies in place to support the proper management of research data for a better understanding of its lifecycle, availability, accessibility, sharing, use, reuse, and security from their respective institutions. Their responses are presented in Figure 5.3.

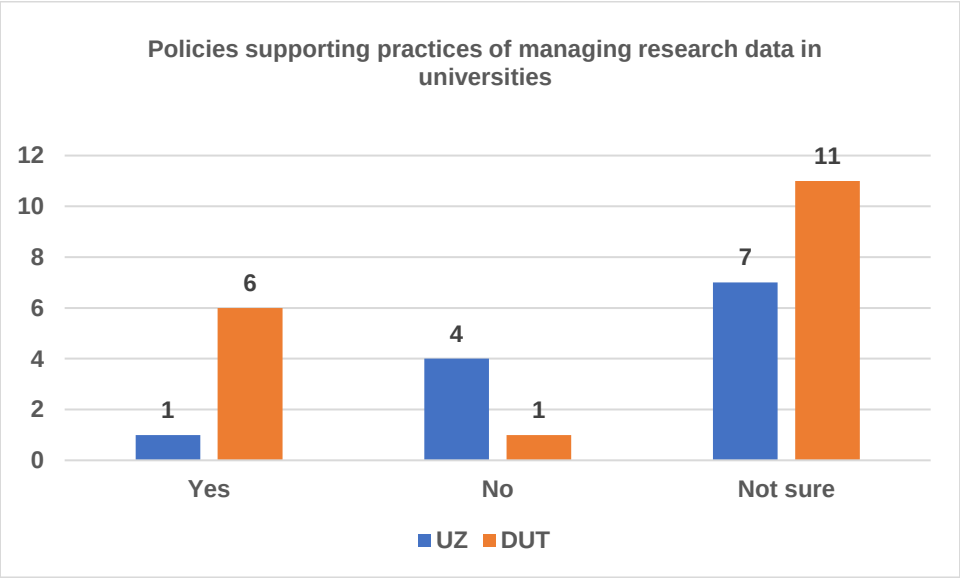

Figure 5. 3: Policies supporting the management of research data

The majority of study participants, 18 (60%), from both participating institutions (UZ and DUT), were unsure whether their institutions have policies supporting the proper management of research data, while 5 (17%) indicated that their institutions do not have any policies and 7 (23%) affirmeed that their institutions do.

5.3.1.4.2 RDM policies and procedures in facilitating and promoting RDM services and practices in universities

Figure 5.4 shows the results of a survey that asked study participants to identify whether policies and procedures are important in facilitating and promoting RDM services and practices in participating institutions using a Likert scale.

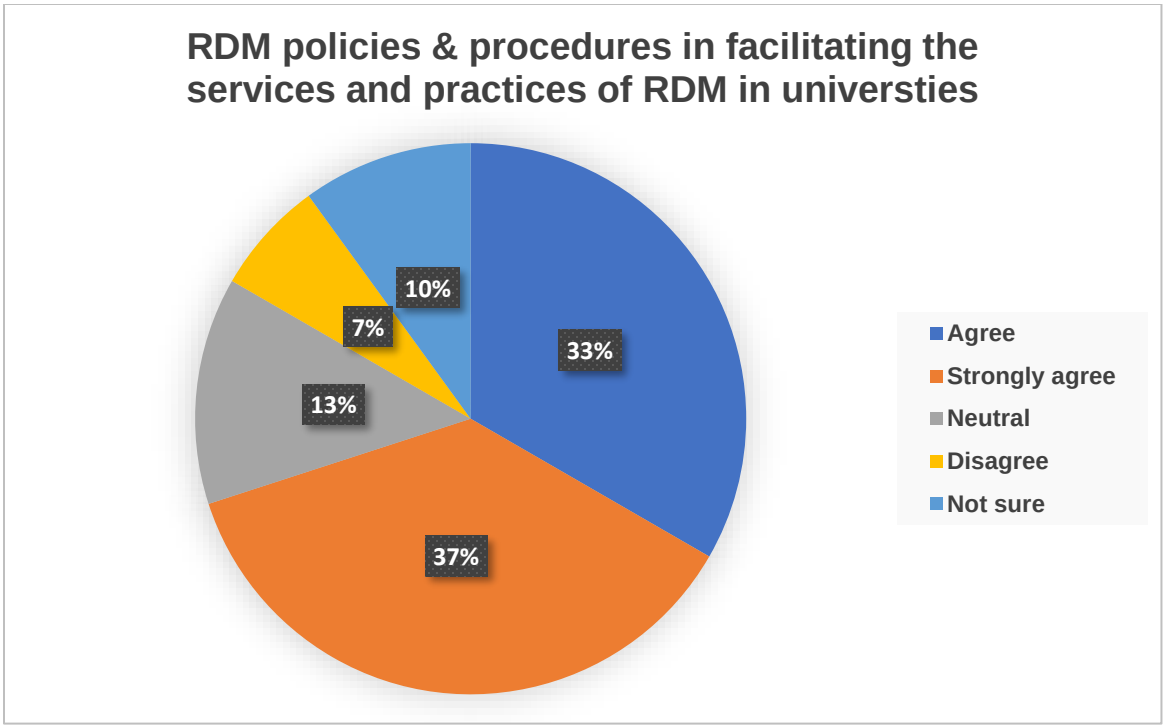

Figure 5. 4: Relevance of policies and procedures in managing research data

Most study participants 11 (37%) strongly agree with the statement that policies and procedures are important in facilitating and promoting RDM services and practices in universities.

5.3.1.4.3 RDM activity on awareness and capacity building and its role

Study participants were asked to identify any available RDM activity with regard to awareness and capacity building from their participating institutions. The results are presented in Figure 5.5.

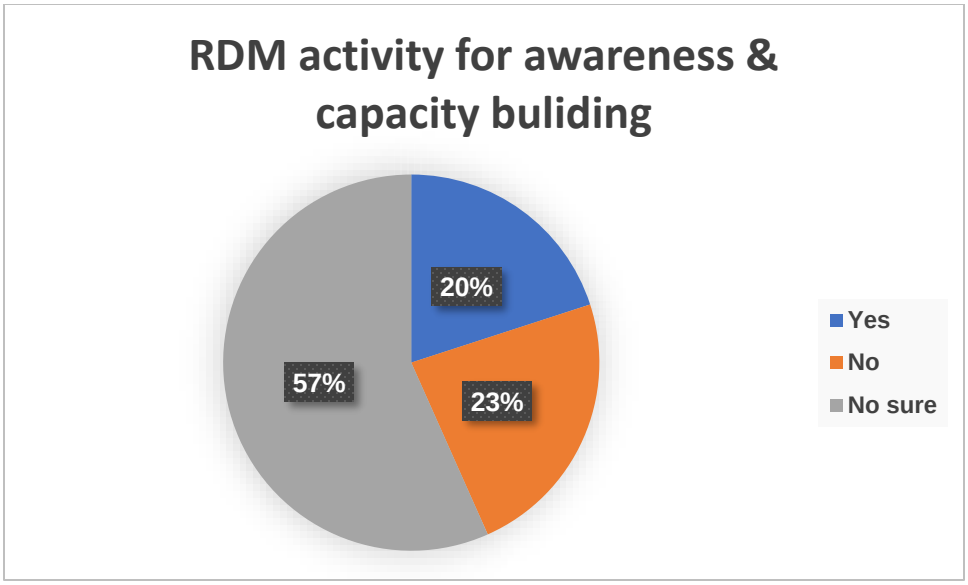

Figure 5. 5: RDM activity on awareness and capacity building

The majority (17-57%) of study participants are unsure of any available RDM activity with regard to awareness and capacity building at their participating institutions. Very few (6- 20%) of the study participants were aware of the RDM activity which they indicated to support in organising workshops and arranging seminars other than anything else.

5.3.1.5 RDM challenges

5.3.1.5.1 Challenges hindering the practices of managing research data in universities

Study participants were asked to identify challenges they are encountering in managing their research data and the results are presented in Table 5.22.

Table 5. 22: RDM challenges

Challenge	Frequency	Percentage (%)
Lack of awareness and understanding of the RDM activities and services provided by the institution	18	60%
Lack of proper knowledge and skills	17	57%
Lack of training workshops	14	47%
Lack of adequate policies and procedures	13	43%
Lack of support and guidance	13	43%
Lack of data resources/ infrastructure	12	40%
Shortage of staff responsible for RDM	9	30%
Lack of support from my faculty	8	27%
Lack of human resources with adequate skills	8	27%
Lack of support from the university library	7	23%
Lack of relevant policy or procedures	7	23%
No capacity-building activities in place to support RDM	6	20%
Lack of support from my department	5	17%
Lack of collaboration between the university and other stakeholders	4	13%

The majority of study participants (18 - 60%) lack awareness and understanding of the RDM activities and services provided by the institution, followed by a lack of proper knowledge and skills (57%).

5.3.1.5.2 Other challenges of managing research data in universities

Aside from the challenges identified in section 5.2.5.2, study participants were asked to identify any additional challenges they face when managing their research data. The following challenges have been identified:

- Lack of urgency and incentives for RDM
- Security
- We need cloud facilities and better archiving
- Data that has the potential for patents will be problematic to share
- Heavy teaching workloads limit available time for research management
- Lack of financial support from management
- Time constraints
- Data storage
- I have the challenge of adequate training in matters of data management

- Time to manage skills development
- Only my motivation and memory
- Lack of think tanks to generate simplicity of complexity [?] daily
- Ensuring that postgraduate students do not share classified information with patent potential

5.3.1.5.3 Suggestions to overcome identified challenges of managing research data in universities

It was appropriate to solicit suggestions from participating universities on some potential solutions to the challenges of managing research data. The following suggestions were made by study participants:

- The challenges should be turned into opportunities for RDM.
- Proper data management system should be put in place.
- Having a dedicated team for data.
- Exclusion of such data from sharing.
- Workshops.
- Reducing the teaching workload of academics that are active in research.
- Training on RDM.
- Rigorous training workshops on RDM to staff.
- The need for a university-wide policy.
- With the university and concerned department to make this visible and a way of common practice to all researchers.
- Functional data management system.
- Clear policies and procedures on RDM and access to data platforms to preserve data.
- Greater institutional assistance.
- Centre for Research Data Management.

From identified suggestions, training, workshops and institutional policies and procedures were most emphasised.

5.4 RDM Framework for public universities in South Africa

Participants in the study were asked to share their thoughts on what they thought could be done to develop an RDM framework for South African public universities. The outcomes are as follows:

- Develop something like a cloud
- Extensive consultation with researchers and research institutions
- Setting policies and guidelines, and providing adequate infrastructure to support its implementation
- Appoint RDM officers with mandates to develop frameworks within their institution
- Awareness and training
- Central database
- An extensive study must be conducted to identify the needs of the public universities
- Universities need to collaborate and share their RDM frameworks
- NRF to lead in the development of the RDM framework
- Providing sufficient resources
- Every university should have RDM policies and procedures in place.
- The university should have Data Governance Groups overseeing data platforms and data access processes. Similar to a National Ethics Committee there could be a Data Governance Body supporting universities for example StatsSA.
- Greater discussions on the topic
- More research to develop guidelines
- Comprehensive business plan.
- There is a need to consult all stakeholders involved in research and to form a coordination group that spearheads the development of the framework.

Among the identified suggestions, RDM policy formulation, awareness, collaboration, and a responsible group or body were cited by more than one study participant as ways to develop an RDM framework for South African public universities.

5.5 Chapter summary

This chapter presented an analysis of qualitative data derived from interviews with seven (7) sampled population groups, namely librarians, technicians, research directors, deans/deputy deans research, deputy vice-chancellors research, and department heads, as well as quantitative data derived from NRF-rated researchers via questionnaires. The findings were presented in accordance with the study's objectives.

The study's findings indicate that both participating institutions have a strong desire to make progress in providing RDM services; thus, DUT has an RDM policy in the draft stage, and UNIZULU has a planning department that has begun RDM. Though institutions do not currently provide RDM services, they do support intensive research data by engaging in research-related activities. Both participating institutions have an open institutional repository (IR) where research data is stored, managed, and archived. Most researchers store (93%) and backup (77%) their research data on personal computers, emails, and external storage devices. RDM services provide several benefits to researchers.

Knowledge and skills in data management, research methodology, data curation, metadata skills, and technical skills are perceived to be extremely important for RDM. Most researchers have shown promising ethical knowledge, including citation (67%), data sharing (50%), data analysis (83%), searching and retrieval (67%), and data capturing (60%). Data management and organisation (53%), database development (47%), data curation and re-use (43%), and software skills (43%), however, are lacking.

UNIZULU currently does not have an RDM policy that facilitates research data management, whereas DUT has a research data management policy that is still in the draft stage. There are insufficient resources and well-trained staff at UNIZULU to provide RDM services, whereas DUT, particularly the library, has adequate resources.

However, the findings revealed that no specific capacity-building programs or strategies for developing RDM skills are currently in place in the institutions studied.

The major challenges of providing RDM are a lack of resources or infrastructure, a lack of skilled personnel, getting researchers on board, and adoption. In light of the challenges, training, workshops, and institutional policies and procedures are suggested as possible solutions.

Collaboration between universities across South Africa and relevant stakeholders and development of a policy or framework by the department of higher education to guide the management of research data in public universities in public universities are all possible ways to develop an RDM framework for South African public universities. Additionally, increased awareness and training, as well as additional research, are regarded as significant.

The following chapter discusses the study's findings in relation to the reviewed literature in Chapter three and the employed theoretical framework in Chapter two of the dissertation.

CHAPTER SIX: DISCUSSION OF FINDINGS

6.1 Introduction

This chapter reflects on the key findings that emerged from the interpretation and analysis of data (Chapter 5) from the library, teaching and research, and NRF-rated researchers from both the University of Zululand (UNIZULU) and Durban University of Technology (DUT) in determining research data management services in selected public universities in KwaZulu-Natal, South Africa. The discussion of findings in this chapter is linked to the population groups of study, the literature reviewed, and theories used where applicable. The following research questions were set and investigated to gain a better understanding and a broader picture of the problem under investigation:

1. What are the RDM practices of the selected public universities?
2. How is the competency requirement of researchers and staff that are responsible for research data management in the selected public universities established?
3. Are there RDM development strategies in place in the selected public universities?
4. What are the challenges facing the selected public universities in managing research data?
5. How can an RDM framework for public universities in South Africa be developed?

6.2 What are the RDM services of the selected public universities?

Qualitative findings revealed that the DUT currently does not offer the services of RDM as it is something that is still in the pipeline, hence their RDM policy is still in the draft stage. The other participating institution (UNIZULU), on the other hand, has a planning department that has begun RDM in the institution, though it is not widely known and there is no specific RDM policy in place. This is consistent with the quantitative findings, which revealed that most study participants (64%), from both participating institutions (UNIZULU and DUT), were unsure whether RDM services were available at their

institutions. These submissions imply that both participating institutions are still in the early stages of providing RDM services. This study confirms the observation made by Nhendodzashe and Pasipamire (2017: 2) that "offering research data management services is an emerging trend in most academic institutions in developing countries," which includes South Africa. Similarly, Ng'eno and Mutula (2018: 28) note that while "RDM has been widely embraced in developed countries", Africa lags. Mushi, Pienaar, and van Deventer (2020: 1) agreed that "extensive RDM is not so common in developing countries," noting that most African higher education and research institutions have yet to implement data management services. On the same note, Patterton (2016) acknowledges that RDM practices used by emerging researchers at South Africa's Council for Scientific and Industrial Research (CSIR) needed to be improved. It is worth noting, however, that such initiatives from both participating institutions (DUT and UNIZULU) represent a significant step forward in offering RDM services, joining other South African public universities such as the University of Cape Town (UCT), University of Pretoria (UP), Cape Peninsula University of Technology (CPUT), Stellenbosch University (SU), University of South Africa (UNISA), and University of Western Cape (UWC), which already offer RDM services and have RDM policies in place.

It is however worth noting that participating institutions support data creation, storage, preservation, sharing, access, use, and re-use as some of the research activities. The relevance of active involvement in research activities is believed to be leading to intensive research data (Chawinga, 2019). This study, on the other hand, revealed grant applications and support for collaborative research from participating institutions as RDM-related activities, while reference and advisory support to researchers on RDM and data citation were revealed as research data services. The results concur with Avuglah's (2016) study which also acknowledges RDM-related activities, such as support for collaborative research; support for data analysis and computational sciences; RDM guidance, grant applications; and support for storage and high-speed connectivity to facilitate research at the University of Ghana. In addition to these research-related activities in the current research, DUT also assists researchers with

data management plans and data mining. Chiware and Becker (2018: 2) emphasise the role of academic libraries in improving RDM services such as data management planning, data storage, and preservation for researchers. Tang and Hu (2019: 9), also agreed to RDM planning (81.0%) and data sharing and dissemination (77.8%) are the most frequently offered services. Additionally, Flores *et al*., (2015: 91) study highlighted data management planning in addition to access control; data citation; intellectual property and copyright; privacy and confidentiality; RDM workshops; scholarly output as some of the RDM services that could be offered in an academic institution. On the other hand, Tang and Hu's (2019) study also found data preservation; data discovery and access; data visualisation; data organisation and curation; and protocol documentation as other services provided by participating institutions. Remarkably, participating institutions, like most other institutions around the world, now offer RDM-related activities and or services; however, their successful use by researchers cannot be guaranteed in the absence of the necessary set of knowledge and skills for carrying out RDM activities. In that regard, the current study has shed light on the competency requirements and training needs of academic researchers at public universities.

This study also revealed that, while neither participating institution has a specific RDM service, they each have an institutional repository (IR) where they store, manage, and archive research publications such as journal articles, conference proceedings, and theses and dissertations. DUT uses the DUT Open Scholar institutional repository, whereas UNIZULU uses DSpace, both of which are open IR types. Openness is one of eight capability factors in the community capability model framework (CCMF) (see section 2.2.1.3, Chapter two), which is one of the theoretical frameworks used in current research and is concerned with research communication in open platforms, including institutional repositories. Lyon *et al* (2012: 26) state that openness of research results and methods can be achieved both during the research study and during the publication stage in order to reach a larger audience. Opening research data is seen as adding value to research in terms of validation, reproducibility, and reusability (Lyon *et al*, 2012: 26). Higgins (2008), referring to the DCC Lifecycle model framework, notes that institutions typically practice ingestion by transferring data to a repository, archive, data

centre, or other curators. Shearer (2015: 30) has cited Edinburgh DataShare as an example of an institutional data repository at the universities of Edinburgh and Oxford, where both academics and PhD students are encouraged to deposit their data. Rather, Ganale and Akhoon (2019: 121) believe that many institutions in the United States are actively developing institutional repositories and a variety of field-specific online collections, particularly in the biological sciences, to store research data. Chigwada, Chiparausha, and Kasiroori (2017) recommend establishing research data repositories or using existing ones to ensure adherence to research data management standards when conducting research. Higgins *et al*. (2018: 5) identify institutional repository as one of the operations and professional practices under their data engagement and research data strategies, respectively. In this regard, it should be noted that institutional repositories exist to provide ready access to institutional research output via open access, thereby providing global visibility to an institution's research works. In the same vein, IRs support the management of scientific knowledge, enhancing knowledge generation, preservation, use, and sharing, and raising the scale of research performance in a research community.

The qualitative findings of this study also reveal that the published works of the researchers are kept in the IR of the participating institutions for a lifetime. Unfortunately, the findings reveal that most study participants were unsure of the duration of research data storage within their institutions. As both participating institutions are still in the development stage of offering RDM services, most researchers store and back up their research data on personal computers, emails, and external storage devices such as USBs, external hard drives, and CDs rather than an institutional repository and personal websites. Ng'eno and Mutula (2018), Chawinga (2019) and Maurya and Subaveerapandiyan (2022) reported similar findings. Mushi (2021: 3) also reported a large portion with more than 88% of researchers storing and handling research data in their devices. However, an institutional repository should be the primary platform supporting institutional research. Though the participating institutions are still in the development stages of the RDM services, RDM services are perceived to be relevant for researchers, as the current study's findings reveal research

visibility, data sharing, researchers' data storage and availability, data use and re-use as RDM benefits for emerging researchers. Although novice researchers are aware of where they can publish their research data, the role of research supervisors in assisting them to find and publish in accredited journals is a priority. According to quantitative findings, data sharing has a significant impact on increasing the impact and visibility of research, encouraging scientific inquiry and debate, and leading to new collaborations between data users and data creators. Findings of similar nature have also been found in the studies of Masinde, Chen and Muthee (2021) and Tenopir *et al*. (2020).

6.3 How is the competency requirement of researchers and staff that are responsible for research data management in the selected public universities established?

This study's qualitative findings reveal a wide range of skills and knowledge required for proper research data management; however, knowledge and skills in data management, research methodology, data curation, metadata skills, and technical skills were overemphasised as truly necessary. Similarly, Tang and Hu (2019: 14) have discovered that metadata skills and data management plan skills are the most important knowledge and skills for librarians and researchers to have when it comes to RDM. Promisingly, quantitative findings reveal that most researchers know ethics, including citation, data sharing, data analysis skills, searching and retrieval skills, and data capturing skills. Hamad, Al-Fadel and Al-Soub (2021: 77) too, highlight that researchers should have the ability to collect, analyse, share, and effectively manage and preserve their research data. Mavodza (2022) has also found data citing skills to be relevant. Unfortunately, most researchers lack knowledge of data management and organisation, database development, data curation and reuse, and software skills. Similar findings were made by Kahn *et al*. (2014: 302), who identify database development and software skills as the most frequent knowledge and skill gaps in RDM. The Chawinga (2019) study also discovered RDM gaps in capacities of librarians and researchers as a result of universities' failure to hold training workshops. Specifically, most authors (Kahn *et al*, 2014; Cox & Penfield, 2014; Pinfield, Cox & Smith, 2014; Ng'eno & Mutula, 2018; Cox *et al*, 2019; Avuglah & Underwood, 2019; Chawinga & Zinn, 2020; Masinde, Chen &

Muthee, 2021) agree that RDM staff and researchers lack certain relevant skills and knowledge.

The study also revealed that most study participants had not received RDM training at their institutions, and they felt a pressing need for training in data archiving and curation. Cox *et al*. (2019: 1434), concur with ongoing staff skill limitations, highlighting curation skills as well as other skills required to properly support RDM. Masinde, Chen and Muthee (2021) also reveal that most respondents had poor knowledge of specific data curation activities, while this study suggested immediate training on data curation, among other training needs. A need for data curation skills is also evident in the study by Tripathi, Shukla and Sonkar (2017: 419) who advise that there is a need to develop digital curation skills in library professionals and researchers. Woeber (2017: 8), too, recognised a need for data curation capabilities and the background of the research environment. Chiware and Becker (2018: 9) agree that digital/data curation and IT technology skills are the most important for the library workforce responsible for providing RDM services in libraries. Similarly, Nhendodzashe and Pasipamire (2017: 7) reveal that digital curation skills are required for librarians to support RDM services. Cox *et al* (2019: 1443) also mention some skills and knowledge that need to be developed in staff to support RDM, including data curation skills among other skill and knowledge types. Though most study participants vouched for data curation and data archiving skills as areas for training, the findings of the Unal *et al* (2019) study indicated that the researchers needed training in areas such as data management planning, metadata, version control of datasets, consistent file naming, and data citation. Clearly, metadata skills, skills on data management plans, data curation skills, data archiving and technical skills are urgently required as they correlate the findings of the current research with most related study findings. Higgins (2008: 137) commenting on the employed DCC Lifecycle model framework suggests that there should be awareness and the implementation of management and administrative actions that are planned to help promote curation and preservation activities in an organisation.

Short courses on RDM, attending workshops, and attending training provided by professional bodies were also identified as practices that could be used in the development of RDM skills on the institutions studied in the current study. Masinde, Chen and Muthee (2021: 12) agree that RDM training may be offered through workshops, online courses, lectures, and tutorials. However, providing short RDM courses was found to be the most effective way of developing RDM skills for both RDM service providers and researchers. Reflecting on formal education to supplement informal training, Mosha and Ngulube (2022) observe that teaching RDM in universities is regarded as an essential component of academic education, including online guides, classroom presentations, and workshops. Wiljes and Cimiano (2019: 7) also recommend that other universities which consider offering a course on research data management, integrate it as an optional module in their degree programs. Lyon *et al* (2012: 25) on the employed CCMF framework, considered training to be highly significant and that it can be provided through early education and professional development as a way of providing skills and other competencies to perform data-intensive research in an institution, in this context, a university. Naros and Simionescu (2019) emphasise the importance of continuous professional training as a means of gaining new knowledge and experience. Workshops and independent studies were viewed as effective methods of promoting RDM training services for librarians in Chiware and Becker's (2018) study. In addition to this observation, Chiware and Becker (2018) identified conference attendance, professional body training, short courses in digital technology, hired consultations, vendor training, and training from internal and external colleagues as platforms for promoting RDM skill development. Tenopir *et al*, (2016) and Tenopir *et al*, (2012), cited by Chigwada, Chiparausha, and Kasiroori (2017: 3), also commended skills development through workshops, some courses on research data, professional development working groups, attending conferences and internal workshops, as well as presentations to equip staff responsible for RDM services and activities. It is worth noting that the University of Edinburgh has an RDM training program called MANTRA, which is a free online course designed for those who manage digital data as part of their research project at the University (University of Edinburgh website, 2019). Van Wyk (2017) also stated that MANTRA is a course offered at the

University of Pretoria (UP) in South Africa. On the other hand, Chiware and Mathe (2016: 9) have observed the Cape Peninsula University of Technology (CPUT) libraries as having established a skills development plan to support faculty or subject librarians in participating in RDM services, of which the plan is facilitated by internal staff and external service providers. Necessary skills and knowledge can be incorporated into the school curriculum through formal education or initiated through training, workshops, and other skill development programs, projects, and strategies through information education including Continuous Professional Development (CPD)

6.4 Are there RDM development strategies in place in the selected public universities?

One participating institution (UNIZULU) does not currently have an RDM policy that facilitates research data management. However, another participating institution (DUT) has a research data management policy that is still in the draft stage and is titled “Durban University of Technology research data management policy”. The absence of an RDM policy is consistent with previous relevant literature findings, which revealed that most higher education institutions, particularly those in Africa, do not have RDM policies because they are still in the implementation stage of RDM services. Unal *et al*. (2019) acknowledges that, among other countries, the United Kingdom is the most advanced in terms of RDM. In South Africa, only a few universities, including the University of Cape Town, the University of Pretoria, the University of the Western Cape, the University of South Africa, Stellenbosch University, and the Cape Peninsula University of Technology, have RDM policies. Despite the absence of RDM policies from participating institutions and most African nations, policies and procedures are regarded as the most important in facilitating and promoting RDM services and practices in universities. This is consistent with Chiware and Becker's (2018) findings, which acknowledge the development and implementation of RDM policies in organisations, particularly universities, as they aid in the delivery of RDM services. Legal and ethical issues are addressed in the employed CCMF as the model's seventh capability factor, which Lyon *et al*. (2012) regard as exclusively addressing the ethical issues that are typically embedded in data-sharing practices, which could then aid in the

provision of restrictions for data-intensive research in institutions. Shearer (2015), Chigwada, Chiparausha, and Kasiroori (2017) agree on the importance of having an RDM policy in place to respond to claims of misconduct in research studies and to assist in the protection of intellectual property. Though RDM policies are regarded as important for providing RDM services in universities, they have some flaws. For example, Ng'eno's (2018: 28) study indicates that "RDM legal framework did not exist in the institutes surveyed; RDM policies and regulations were outdated; there was limited RDM awareness and advocacy; the institutes lacked RDM security systems; the institutes suffered from a lack of inadequate RDM guidelines on standardisation and technical infrastructure." In this regard, Kahn *et al*. (2014) emphasise the importance of multi-stakeholder RDM policy development. Essentially, having a policy alone in place is not enough. The policy needs to be functional.

In terms of infrastructure or resources, the findings from both library and teaching and research reveal UNIZULU to lack resources and well-equipped staff for the provision of RDM services. On the other hand, findings from the DUT library revealed adequate resources; most participants noted the actual implementation of RDM to determine whether additional resources are required to support RDM practices in their institution. This contradicts the findings from teaching and research in DUT as most study participants noted the inadequacy of infrastructure and a lack of human resources or capital. Such findings reveal the library to be the centre of RDM service provision in DUT. The narration from UNIZULU confirms the findings of Nhendodzashe and Pasipamire (2017); Mushi, Pienaar, and van Deventer (2020) report that RDM practice in developing countries is very limited due to several deficiencies, including infrastructure, resources, financial constraints, stakeholder collaboration, and human capabilities; all of which are significant threats to RDM service provision. Similarly, Shearer (2015) observes a fragmented infrastructure for research data management in Canada, with some fields receiving extensive coverage while others receive little. It could be argued that a proper infrastructure should support all processes of the research lifecycle, beginning with the research data management plan and ending with the publication of the results. On this note, universities are encouraged to invest in existing and new infrastructure to support their research data management practices or

services, as good infrastructure in a university is critical for providing RDM services. For example, Chiware and Mathe (2016), pointed out that the Cape Peninsula University of Technology libraries in South Africa invested in a pilot project before providing full-service RDM infrastructure. Similarly, Lyon *et al*. (2012) confirmed from the used CCMF that technical infrastructure supporting research is marked with necessary tools and services following different stages of the research life cycle. In the context of the current study, good infrastructure can support RDM service provision in the university setting by being friendly to its users, including both researchers and RDM staff.

There are currently no specific capacity-building programs or strategies for developing RDM skills in the institutions studied. This is consistent with quantitative findings from the current study, which show that most study participants are unaware of any available RDM awareness and capacity-building activity. UNIZULU, on the other hand, has several programmes run by the institutional research office and faculty research office that are said to address activities related to research in general and to highlight areas related to data management, but they are not RDM specific. DUT had previously invited UCT to conduct a short course on RDM training for them, and the institution may be able to use this course in the future as an external opportunity. Furthermore, DUT supports journal writing retreats, webinars for networking and different research perspectives, and ongoing ethics workshops by the institutional research ethics committee, all of which are part of the process of preparing staff for research-related activities. However, it is worth noting that some universities in South Africa and other parts of the world have capacity-building programmes and projects to support RDM. For example, Kahn *et al*. (2014: 298), highlight the Network of Data and Information Curation Communities (NeDICC) in South Africa, which "arranges seminars, workshops, and conferences to promote awareness around digital (including data) curation aimed at practitioners and managers involved with digital object management and encourage the growth of knowledge in this area." The UCT Library and Information Studies Centre (LISC), which offers a master's degree module in Digital Curation, is another example of a capacity-building platform. Kahn *et al*. (2014), on the other hand, believe that because there has been a growing interest in RDM in South Africa, with the attendant need for capacity building in this area, the LIASA's Higher Education Libraries Interest Group

hosted a workshop on RDM in Cape Town in March 2014. Again, Chiware and Mathe (2016) point out that the Cape Peninsula University of Technology (CPUT) has established several collaborative and institutional working groups to assist in the implementation of RDM. Van Wyk (2017) discusses five (5) pilot projects implemented at the University of Pretoria between 2013 and 2015 to provide insight into researchers' needs for RDM services and activities. Gunjal and Gaitanou (2017) identify the Leeds Building Capacity Project (LBC) as one of the global projects in which libraries engage in RDM practices. As a result, capacity-building programs are important platforms for improving collaboration across departments, institutions, or external bodies during research processes.

6.5 What are the challenges facing the selected public universities in managing research data?

One participating institution (UNIZULU) identified a lack of resources or infrastructure, as well as skilled personnel, as a major challenge in providing the RDM service. A lack of infrastructure to support extensive RDM is also confirmed by related studies (Avuglah & Underwood 2019, Cox & Penfield 2014, and Qin 2013). Several studies (Chawinga & Zinn, 2020; Cox *et al* 2019, Chawinga 2019, Avuglah & Underwood 2019, Ng'eno & Mutula 2018, Cox *et al* 2017, Pinfield, Chiware & Mathe 2016, Cox & Penfield 2014, Cox & Smith 2014, Tenopir *et al*. 2014) which agree that staff skills in RDM support are still limited. The UNIZULU research findings also reveal a lack of funding, a lack of relevant policy, a lack of collaboration between the research office and the departments, and researchers' reluctance as challenges impeding the provision of RDM services in the institution. Similarly, Anduvare (2019: 64) observes that South Africa faces financial challenges in supporting and delivering RDM services. In addition, Hamad, Al-Fadel and Al-Soub (2021) had findings of a similar nature. However, financial constraints are a problem for developed countries as well (Anduvare 2019). A lack of funding is also confirmed by the Council on Library and Information Resources CLIR (2013: 4-7). Lötter's (2014) study also reveals a lack of financial resources and policies. Nge'no and Mutula (2018), on the other hand, confirm a lack of collaboration and partnerships between institutions to improve RDM activities. In this regard, the CCMF framework

employed in this study emphasises the types of relationships established during research as a major determinant of the types of research that can be performed in an organisation, where collaboration between a section or department, across departments, within institutions, or with external bodies is significant (Lyon *et al*. 2012: 21).

Another participating institution (DUT) also revealed getting researchers on board as problematic in addition to adoption, which is about getting people to adopt the institution's culture of managing research data. Halbert in Gunjal and Gaitanou (2017: 10), too, observed a lack of a sense of priority among researchers. Additionally, issues relating to culture are also evident in previous studies (Anduvare 2019, Avuglah 2016, Molly *et al* 2012, Morgan, Duffield & Walkley Hall 2017, Penfield, Cox & Smith 2014, Cox *et al* 2017). Anduvare's (2019) study, for example, revealed that most universities appear to face a cultural challenge when adopting RDM, in the sense that not every stakeholder will find it easy to migrate from one form of managing and handling data to another. Furthermore, Avuglah (2016: 30) identified culture as a serious aspect of the RDM challenge, ranging from the institutional to the disciplinary levels. Acceptance and cultures are also mentioned by Pinfield, Cox, and Smith (2014: 28) as factors influencing the development of RDM in libraries. Molloy and Snow (2012: 249) denote "changes to the way research data is managed imply cultural change in the way research is practised". Similarly, Penfield, Cox, and Smith (2014: 28) see varying cultures as a barrier to the implementation of RDM initiatives. As a result, Avuglah (2016: 22), citing Van Deventer and Pienaar, suggests that institutions should consider the cultural environment when implementing RDM initiatives. Furthermore, the findings from DUT reveal faculty resistance to RDM and an inability to promote it properly, as well as a lack of policy, which is still in the draft stage, which becomes a stumbling block as the policy should facilitate the entire RDM process in the institution. Quantitative findings from both participating institutions, on the other hand, confirm a lack of proper knowledge and skills, awareness among researchers, as well as a lack of understanding of the institution's RDM activities and services and a lack of training workshops. Senft, Stahl and Svoboda (2022: 14) also find a need for awareness regarding the need for substantial RDM. Among other challenges in providing RDM,

Chawinga's (2019) study discovered a lack of training workshops for librarians. Similarly, Wachtler *et al* (2021: 12) observed that that training in RDM is either not provided at all or carried out in an *ad hoc* manner in NFDI-Neuro, Germany. Hamad, Al-Fadel and Al-Soub (2021: 81) also confirm a lack of training to be problematic. Xu *et al* (2022: 7) confirm a limited RDM-related training as well. Considering the challenges, study participants frequently suggest training, workshops, and institutional policies and procedures as possible solutions. As this is the case, Wachtler *et al.* (2021: 12) suggest educating researchers as the key element for improving RDM, and they observed a major task in promoting knowledge and competence in RDM throughout the neuroscience community would be to develop and implement a coordinated training concept. Kwanya (2021: 9) also suggests appropriate training and publicity programmes for researchers in Sub-Saharan Africa to adopt the concept of RDM. Bunkar and Butt (2020: 144) also recommend intensive training for library professionals. Relevance of training is also evident in the studies of (Masinde, Chen & Muthee, 2021; Maurya & Subaveerapandiyan, 2022; Birkbeck, Nagle & Sammon, 2022).

6.6 How can an RDM framework for public universities in South Africa be developed?

Several propositions on the development of an RDM framework for public universities in the country, South Africa, were highlighted in qualitative findings, including collaboration between universities across South Africa and relevant stakeholders; one common database where all universities deposit their research; and the development of a policy or framework by the department of higher education that could guide the management of research data in public universities; having a body like Chelsea's research, NRF, and StatsSA; public universities could develop an RDM learning plan; and human capital and resources. This is consistent with the quantitative findings, which have also highlighted RDM policy formulation, awareness, a central database, collaboration, and a responsible group. Furthermore, quantitative findings reveal the importance of awareness and training, as well as more research to develop guidelines and other recommendations. According to the qualitative and quantitative findings of the participating institutions, RDM policy formulation, a collaboration between universities

and relevant stakeholders, and a responsible group or body are all viable options for developing an RDM framework for South African public universities. A detailed framework is proposed in the next chapter (see section 7.4.5), Figure 7.1.

6.7 Chapter Summary

This chapter discussed the findings considering the research questions that were established, the literature that was reviewed, and the theoretical frameworks that were used (community capability model framework (CCMF) and digital curation DCC lifecycle model framework).

Participating institutions are still in the early stages of fully providing RDM services. One participating institution (UNIZULU) has one department that has started RDM, though it is not widely known. Another participating institution (DUT) currently only has a draft RDM policy. Some of the research activities that support research practices in both participating institutions include data creation, storage, preservation, sharing, access, use, and re-use. DSpace and DUT Open Scholar are institutional repositories (IRs) that are used to permanently store, manage, and archive research data such as journal articles, conference proceedings, theses, and dissertations.

Most researchers keep their research data on personal computers, emails, and external storage devices like USBs, external hard drives, and CDs. The RDM service is useful for new researchers because it promotes research visibility, data sharing, data storage and availability for researchers, data use and re-use. Researchers benefit from data sharing in several ways, including increasing the impact and visibility of their research, promoting scientific inquiry and debate, and fostering new collaborations between data users and data creators.

RDM requires both staff and researchers to have knowledge and skills in data management, research methodology, data curation, metadata skills, and technical skills. Fortunately, most researchers understand data citation, data sharing, data analysis, searching and retrieval. They are, however, lacking in data management and organisation, database development, data curation and re-use, and software skills.

Short RDM courses, workshops, and professional body training are all practices that could be used in institutions to develop RDM skills.

DUT has a research data management policy, titled Durban University of Technology research data management policy, which is currently in draft form, whereas UNIZULU does not. UNIZULU lacks resources and well-equipped staff to provide RDM services, whereas DUT library has adequate resources; however, whether additional resources are required to support RDM practices in their institution will be determined by the actual implementation of RDM. Participating institutions currently have no specific capacity-building programs or strategies for developing RDM skills, but their institutional research office runs several programmes addressing general research.

The chapter identifies a lack of resources or infrastructure, skilled personnel, getting researchers on board, and adoption as major challenges in providing the RDM service. RDM challenges may be addressed through training, workshops, and institutional policies and procedures.

Finally, this chapter identifies RDM policy formulation, collaboration between universities and relevant stakeholders, and a responsible group or body such as NRF and StatsSA as viable options for developing an RDM framework for South African public universities.

The next chapter presents a summary of the findings, conclusions, and recommendations.

CHAPTER SEVEN: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

7.1 Introduction

The previous chapter discussed the findings of the study. This chapter summarises, concludes, and suggests recommendations based on research data management services in selected public universities in KwaZulu-Natal, South Africa. This study used the community capability maturity model framework (CCMF) and the digital curation centre (DCC) lifecycle model to gain insight into the problem under investigation and collected both qualitative and quantitative data from its sampled population groups, which included librarians, technicians, research directors, deans/deputy deans research, deputy vice-chancellors research and heads of departments, and NRF rated researchers. As data was collected during the COVID-19 pandemic, qualitative data was collected through 23 semi-structured interviews conducted online via Zoom and Microsoft Teams, and 60 research questionnaires were converted into Google forms and emailed to all NRF rated researchers from the University of Zululand (UNIZULU) and the Durban University of Technology (DUT). The study's conclusions and recommendations have been made in relation to the following research objectives that were set:

1. To determine the RDM practices of the selected public universities.
2. To establish the competency requirements of researchers and staff that are responsible for research data management in the selected public universities.
3. To determine the RDM development strategies in the selected public universities.
4. To establish the challenges facing the selected public universities in managing research data.
5. To develop an RDM framework for public universities in South Africa.

7.2 Summary of findings

This section provides a summary of the findings aligning it with the study's objectives. Following that, the summary is used to draw conclusions and make recommendations.

7.2.1 To determine the RDM services of the selected public universities

The purpose of this section was to answer the main research question: What are the RDM practices of the selected public universities? The findings revealed that DUT is currently not providing RDM services due to their RDM policy, which is still in the draft stage. UNIZULU has a planning department that has begun RDM in the institution, which is encouraging, though it is not widely known and there is no specific RDM policy in place. These findings support quantitative findings that most study participants from both participating institutions were unsure whether their institutions provided RDM services or not. In this regard, both participating institutions are still in the early stages of RDM service provision. However, such initiatives from both participating institutions represent a significant step forward in offering RDM services, joining other South African public universities that already offer RDM services and have RDM policies, such as the University of Cape Town, University of Pretoria, Cape Peninsula University of Technology, Stellenbosch University, University of South Africa, and the University of Western Cape. Detailed discussion on RDM activities, practices and services in libraries and higher learning institutions has been extensively covered in Chapter three of the dissertation (see section 3.2). Even though participating institutions are still in the early stages of fully providing RDM services, it is important to note that they support data creation, storage, preservation, sharing, access, use, and re-use as research activities of which are the activities that significantly support intensive research data. Grant applications and support for collaborative research from participating institutions, on the other hand, were identified as RDM-related activities, while reference and advisory support to researchers on RDM and data citation were acknowledged as research data services. DUT, on the other hand, helps researchers with data management plans and data mining. Both participating institutions have an institutional repository (IR) in which they store, manage, and archive research data such as journal articles, conference proceedings, and theses and dissertations for the duration of their existence. Unfortunately, most researchers from both participating institutions are unsure how long research data is stored at their institutions. DUT employs the DUT Open Scholar institutional repository, whereas UNIZULU employs DSpace; both are open IR types.

IRs aid in the management of scientific knowledge by improving knowledge generation, preservation, use, and sharing, as well as raising the research performance scale in a research community and these form basic components of RDM. Extensive discussion on the IRs in support of RDM has been broadly covered in Chapter three of the dissertation (see section 3.6.2.4). Since both participating institutions are still in the process of developing RDM services, most researchers store and backup their research data on personal computers, emails, and external storage devices like USBs, external hard drives, and CDs rather than an institutional repository and personal websites. An IR, on the other hand, should be the primary platform for institutional research. Other persons responsible for research data backup from participating institutions are unknown. RDM services are thought to be relevant for researchers, as the current study's findings reveal that RDM benefits for emerging researchers include research visibility, data sharing, researchers' data storage and availability, data use and re-use. Although novice researchers are aware of where they can publish their research data, research supervisors must assist them in finding and publishing in accredited journals. Furthermore, data sharing has a significant impact on increasing the impact and visibility of research, promoting scientific inquiry and debate, and fostering new collaborations between data users and data creators.

7.2.2 To establish the competency requirements of researchers and staff that are responsible for research data management in the selected public universities

This objective sought to answer the main research question: how is the competency requirement of researchers and staff that are responsible for research data management in the selected public universities established? Competencies are frequently used as the foundation for skill standards, which specify the level of knowledge, skills, and abilities required for workplace success, as well as potential measurement criteria for assessing competency accomplishment. The findings of research find a wide range of skills and knowledge as required for staff responsible for managing research data in an institution as well as researchers to properly manage their research data; however, knowledge and skills in data management, research

methodology, data curation, metadata skills, and technical skills are perceived to be extremely important. Detailed discussion on the competency requirements of researchers and staff responsible for RDM has been broadly covered in Chapter three of the dissertation (see section 3.5). Most researchers from both participating institutions have demonstrated ethical knowledge, including citation, data sharing, data analysis skills, searching and retrieval skills, and data capturing skills. They are, however, lacking in data management and organisation, database development, data curation and re-use, and software skills. This is because most study participants had not received RDM training at their institutions and felt a strong need for data archiving and curation training. Both data archiving and curation skills have been identified as lacking in most of the literature, and a section addressing the skills and knowledge deficits of researchers and staff responsible for providing RDM services is included in section 3.7 of the dissertation, Chapter three. Short courses on RDM, attending workshops, and attending professional body training are thought to be practices that could be used in the development of RDM skills in institutions. Short RDM courses, on the other hand, are perceived as the most effective way of developing RDM skills for both RDM staff and researchers. Furthermore, required skills and knowledge can be integrated into the school curriculum or initiated through training, workshops, and other skill development programs, projects, and strategies.

7.2.3 To determine the RDM development strategies of the selected public universities

This objective addresses the research question: are there RDM development strategies (e.g. policy, investment, infrastructure, capacity building) in place in the selected public universities? In this context, developing strategies allow institutions to focus on how they will support the provision of RDM services through, among other things, the development and implementation of relevant policies, infrastructure, investment, the human wheel, and capacity building. The findings reveal UNIZULU as not currently having an RDM policy that facilitates research data management, whereas DUT has a research data management policy that is still in the draft stage called Durban University

of Technology research data management policy. Section 3.6.1 of Chapter three of the dissertation discusses RDM policies and procedures in a broader context.

The findings also reveal that UNIZULU lacks resources and well-trained staff to provide RDM services, whereas DUT, particularly the library, possessed adequate resources; most participants noted the actual implementation of RDM to determine whether additional resources are required to support RDM practices in their institution. Unfortunately, this contradicts the findings of DUT teaching and research, as most study participants reported a lack of adequate infrastructure and skills in workforce. In this regard, the library is regarded as the hub of RDM service provision at DUT. All processes of the research lifecycle, beginning with the research data management plan and ending with the publication of the results, should be supported by a proper infrastructure. A good infrastructure in this context, for example, should support RDM service provision in the university setting by being welcoming to its users, who include both researchers and RDM staff. Section 3.6.2 (Chapter three) of the dissertation discusses infrastructure and investment as some development strategies that support RDM services.

The study's findings also reveal that there are currently no specific capacity-building programmes or strategies in place in the institutions studied for developing RDM skills. UNIZULU, on the other hand, has several programmes run by the institutional research office and faculty research office that are said to address general research activities and to highlight areas related to data management, but they are not RDM-specific. DUT had previously invited UCT to participate in a short course on RDM training, and the institution may be able to use this course as an external opportunity in the future. Furthermore, as part of the process of preparing staff for research-related activities, DUT supports journal writing retreats, webinars for networking and different research perspectives, and ongoing ethics workshops by the institutional research ethics committee. Capacity-building programmes, in this vein, become critical platforms for improving collaboration across departments, institutions, or external bodies during research processes. Capacity-building has been extensively discussed in Chapter three as another development strategy for RDM (section 3.6.3).

7.2.4 To establish the challenges facing the selected public universities in managing research data

This section answered the research question: what are the challenges facing the selected public universities in managing research data? The findings revealed numerous challenges hindering, and encountered in managing, research data management, with participants citing a lack of resources or infrastructure, a lack of skilled personnel, getting researchers on board, and adoption, which is about adopting the institution's culture of managing research data, as major challenges. A lack of funding, a lack of relevant policy, a lack of collaboration between the research office and the departments, and researchers' reluctance, faculty resistance to RDM and an inability to promote it attractively, awareness among researchers, as well as a lack of understanding of the institution's RDM activities and services, and a lack of training workshops are also mentioned as challenges. Training, workshops, and institutional policies and procedures were suggested as possible solutions considering the challenges. Other challenges were cited in Chapter three of the dissertation (section 3.5).

7.2.5 To develop an RDM framework for public universities in South Africa

This objective intended to answer the main research question: how can an RDM framework for public universities in South Africa be developed? In qualitative findings, several propositions on the development of an RDM framework for public universities South Africa were highlighted, including collaboration between universities across South Africa and relevant stakeholders; one common database where all universities deposit their research; and the development of a policy or framework by the Department of Higher Education that could guide the management of research data in public universities. This corresponds to the quantitative findings, which also highlight RDM policy formulation, awareness, a central database, collaboration, and a responsible group. Furthermore, quantitative findings demonstrated the importance of increased awareness and training, as well as additional research to develop guidelines and other recommendations. The qualitative and quantitative findings of the participating institutions confirm RDM policy formulation, collaboration between universities and

relevant stakeholders, and the establishment of a responsible group or body as the best options for developing an RDM framework for South African public universities.

7.3 Conclusions

From the findings of the study, it can be concluded that this study has met and successfully answered its research objectives and questions. The conclusions are arranged according to the sub-themes derived from the set objectives of the study.

7.3.1 RDM services of the selected public universities

The study established that both participating institutions are still in the early stages of fully providing RDM services. There is a glimpse of an RDM initiative within the planning unit at UNIZULU that is still not widely known within the institution, and DUT RDM policy is still in the draft stage. However, both institutions support data creation, storage, preservation, sharing, access, use, and re-use as research activities. In addition to these activities, both institutions also support grant applications and collaborative research, as well as RDM and data citation as research-related activities and data services. Institutional repositories are used to store, manage, and archive research data for a lifetime. Unfortunately, most researchers are unaware of the duration of research data storage at their institutions. Most researchers store and backup their research data on personal computers, emails, and external storage devices. RDM service is relevant for emerging researchers because it encourages research visibility, data sharing, data storage and availability for researchers, data use and re-use. Research supervisors have the potential to help emerging researchers find and publish in accredited journals. The study further notes data sharing as significant in increasing the impact and visibility of research, promoting scientific inquiry and debate, and fostering new collaborations between data users and data creators.

7.3.2 Competency requirements of researchers and staff that are responsible for research data management in the selected public universities

There are plenty of knowledge and skills required by researchers and staff responsible for the provision of RDM, however, data management, research methodology, data curation, metadata skills, and technical skills are really a requirement. Most researchers

have been noted to understand ethics, including citation, data sharing, data analysis, searching and retrieval, and data capturing skills. Unfortunately, they lack a knowledge of data management and organisation, database development, data curation and re-use, and software skills. Short RDM courses, workshops, and professional body training are all practices that could be used in the development of RDM skills in institutions. It is significant that both researchers and staff responsible for RDM receive training in data archiving and curation, which is a necessity, since most study participants had not received any RDM training at their institutions.

7.3.3 RDM development strategies in the selected public universities

The study showed that UNIZULU does not currently have an RDM policy that facilitates research data management, whereas DUT has a research data management policy that is still in the draft stage. In terms of infrastructure or resources, the study concludes that UNIZULU lacks resources and well-equipped staff to provide RDM services, whereas DUT library has adequate resources; though the actual implementation of RDM will determine whether additional resources are required to support RDM practices in their institution. However, there is a contradiction as results from teaching and research showed a lack of adequate infrastructure as well as a lack of human resources or capital. These results bring into recognition that the library is the hub for RDM service delivery at DUT. There are currently no known capacity-building programmes in the institutions that participated. Some capacity-building programmes in the institutions mentioned in the study are not specifically designed to aid RDM. In addition, DUT previously invited UCT for a short course on RDM training, and the institution may be able to use this course as an external opportunity in the future.

7.3.4 Challenges facing the selected public universities in managing research data

Both researchers and staff responsible for RDM face several challenges in managing research data, however a lack of resources or infrastructure, skilled personnel, getting researchers on board, and adoption, which is about getting people to adopt the institution's culture of managing research data are most prevalent. In this sense, the study showed training, workshops, and institutional policies and procedures as potential solutions to RDM challenges.

7.3.5 RDM framework for public universities in South Africa

Finally, as viable options for developing an RDM framework for South African public universities, this study concluded RDM policy formulation, collaboration between universities and relevant stakeholders, and a responsible group or body such as NRF and StatsSA.

7.3.6 Study limitations

The empirical results reported in this study should be viewed considering some limitations, as is the case with some other research studies. First, because the study used qualitative research, the results cannot be generalised. This gap was filled, however, by including quantitative research in a single study to allow for generalisation of the findings. Second, the study only focused on two public universities in the KwaZulu-Natal province. However, the study initially targeted three institutions in this province, but there was no response other than receiving a letter from the gatekeepers, so the third institution was removed. Third, because the study was conducted while COVID-19 and its rules and regulations were still in effect, there was a very slow response rate, which delayed data collection and analysis; as a result, the data collection period was extended beyond the initial planned timeframe.

7.4 Recommendations

7.4.1 RDM services of the selected public universities

Given that participating institutions are still in the early stages of providing RDM services, the primary platform supporting institutional research should be an institutional repository. The study suggests collaboration between the institutional research office and faculties or departments to raise awareness because the study's findings revealed a lack of awareness about whether the institution provides RDM services. Again, the study suggests that RDM training and workshops are greatly needed to support the institution's RDM services. Furthermore, data sharing among researchers is recommended by the study because it has a significant impact on increasing the impact and visibility of research, promoting scientific inquiry and debate, and fostering new collaborations between data users and data creators.

7.4.2 Competency requirements of researchers and staff that are responsible for research data management in the selected public universities

Given the significant knowledge and skill gap for both RDM staff and researchers, short RDM courses, workshops, and professional body training are highly recommended for institutional RDM skill development. Some training and workshops could include all staff members in the institution who are responsible for the support of RDM services, such as those in the research office, library, faculty research officers, and departments, so that they can gain a common understanding of the expected roles and responsibilities they must perform in providing RDM services, as well as understanding collaborative partnerships within institutions. The study also suggests that DVC research provide a mandatory training workshop at the departmental, faculty, and institutional levels. Again, the study strongly suggests functional partnerships between department heads, research deans, and the research office to establish a culture of training workshops for staff skill development. Furthermore, because most study participants had not received RDM training at their institutions, data archiving and curation training is required.

7.4.3 RDM development strategies in the selected public universities

Given the lack of RDM policies in participating institutions and most universities, the study recommends that more than one stakeholder, such as publishers, research officers, information professionals, and researchers, develop an RDM policy in an organisation. The policy is important in guiding the implementation of research data management practices and establishing responsibilities for researchers, libraries, and research offices. The study also suggests that before selecting an infrastructure to support the provision of RDM services, institutions of higher learning conduct thorough planning. They should also invest in both existing infrastructure (e.g. institutional repository) and new infrastructure to avoid a duplication of efforts. Again, the study recommends a strong need for RDM training for existing employees to provide and support RDM services in their departments. Furthermore, the study recommends that capacity-building programmes be offered at universities as short courses, workshops, training, and/or seminars to support RDM services.

7.4.4 Challenges facing the selected public universities in managing research data

Given the difficulties that universities face in managing research data, such as a lack of resources or infrastructure, skilled personnel, getting researchers on board, and adoption, to name a few, the study strongly recommends training, workshops, and institutional policies and procedures. The study also suggests that research forums be used to engage postgraduates in capacity development programs. Furthermore, the study strongly suggests that RDM be integrated into the research cycle rather than as a separate service.

7.4.5 RDM framework for public universities in South Africa

The study strongly recommends that the Department of Higher Education and Training (DHET) develop an RDM framework, as they are responsible for developing most higher education frameworks. As a result, universities cannot do it on their own because they are leaning toward the DHET, and the DHET oversees all major policies, which institutions extract from the head (DHET). The study also suggests RDM policy formulation, collaboration between universities and relevant stakeholders, and a responsible group or body such as the NRF and StatsSA for the development of an RDM framework for South African public universities. The study also suggests that all universities deposit their research papers in a single database. The study also suggests forming partnerships and collaborating in terms of sharing tools and skills required for adopting RDM services, as well as knowledge of addressing challenges encountered in the provision of RDM. Figure 7.1 is the proposed RDM framework for South African public universities which has been developed through the recommendations made by study participants in the current research.

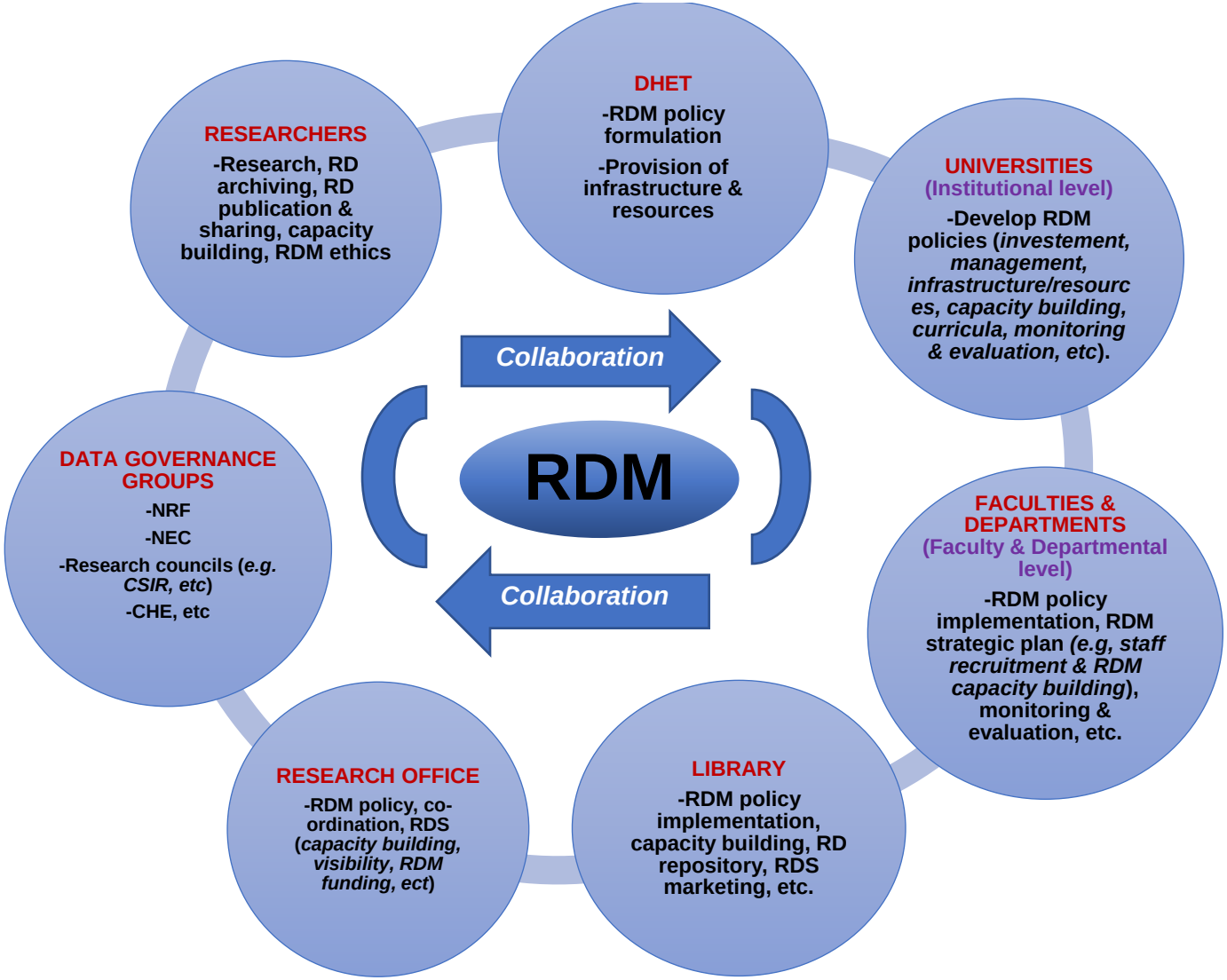

Figure 7. 1: Proposed RDM framework for SA public universities

Figure 7.1 highlights several key elements that should be in place for proper research data management in a higher learning institution/university. These elements are thought to be united through collaboration rather than anything else. Collaboration is one of the eight capabilities of the community capability model framework that Lyon *et al* (2012: 12) perceive as having a strong bearing on the types of research that can be performed in an organisation. Lyon *et al*. (2012), indicate that collaborations can be informal or semi-formal, or they can be strictly controlled, managed, and enforced through contracts and agreements. In this context, collaboration can take place between the Department of Higher Education and Training (DHET), universities/institutions, faculties and departments, institutional library, research office, data governance groups, and researchers to support good research practice and culture in a university setting. The narrative of each element and their critical roles in promoting RDM service endorsement is provided below.

Department of Higher Education and Training (DHET)

The department of higher education and training (DHET) is the university's custodian, in charge of all major policies from which institutions draw theirs. DHET could significantly assist universities by providing resources/infrastructure to support institutional RDM services and practices, in addition to developing the majority of the country's policies. It is critical for organisations to have RDM policies in place or to develop and implement RDM policies (Chiware & Becker, 2018: 6). Chiware and Becker (2018) go on to say that RDM policies help with service delivery by ensuring consistency among researchers and the provision of uniform services by responsible personnel such as library professionals. Collaboration between the DHET, universities, and data governance groups, on the other hand, is critical to the successful formulation of RDM policy. Open discussions can help to foster a collaborative environment by providing more opportunities to share experiences, challenges, tool adoption, infrastructure, and other necessary elements of providing RDM services in a university setting.

Universities (Institutional level)

Universities, rather than DHET, could take responsibility and develop their own RDM policies. Some South African universities, for example, already have RDM policies in

place that they developed themselves by referring to research-related policies. Among these are the University of Cape Town, the University of Pretoria, the Cape Peninsula University of Technology, Stellenbosch University, the University of South Africa, and the University of Western Cape. Then, in order to put RDM provision into practice, universities must adopt and implement formulated RDM policies. It should also be acknowledged that RDM policy alone will not result in the provision of RDM services. As a result, infrastructure, among other things, becomes necessary. Lyon *et al* (2012: 31) believe that infrastructure should meet the needs of users in their various research activities such as data capture and processing, data storage, curation and preservation, discovery and access, and so on. On this note, universities can invest in both new and existing infrastructure, such as institutional repositories (IRs). This is because most universities have institutional repositories in place that they use as research hubs. Universities, on the other hand, are in a better position to raise awareness of the concept of RDM by offering short RDM courses. RDM is infused in the curricula offered in this note. Accordingly, Wiljes and Cimiano (2019: 7) consider research data management training to be an essential component of academic education. Some universities already offer some RDM short courses. For example, Van Wyk (2017) notes MANTRA to be an RDM course offered at the University of Pretoria. RDM short courses will provide significant combined content or specific skill training because they are more practical in nature and contain less theory than a university course. In this sense, short courses will provide staff with more practical experience that will help them deliver RDM services. Universities also play an important role in capacity-building programmes, which can be facilitated through RDM training and workshops, as well as internal and external capacity programmes for staff who support RDM provision. Capacity-building is also mentioned in a list of sustainable development goals (SDG) 17, which is about revitalising the global partnership for sustainable development employers. This goal emphasises how universities, in particular, can serve as capacity-building hubs through research, innovation, and data collection and analysis. Universities can offer training and workshops to help employees gain the knowledge and skills they need. Employees can benefit from training and workshops to learn new skills or update existing ones, increasing productivity. Capacity development assists universities in developing and

strengthening the skills, instincts, abilities, processes, and resources required to survive, adapt, and thrive in a rapidly changing world. In addition, universities should have a management team in place to oversee monitoring and evaluation (M&E) of RDM service delivery. According to Hobson, Mayne, and Hamilton (2013: 6), conducting M&E in an organisation can improve internal learning and decision making about project design, how the group operates, and implementation, such as success factors, barriers, working approaches, and others. In this sense, M&E aids in the efficient use of allocated resources while also assisting in good resource management. Moreover, M&E contributes to novel approaches in the successful delivery of RDM services.

Faculties and departments (Faculty and departmental level)

Universities' faculties and departments play an important role in implementing institutional RDM policies. Universities must develop a proper strategic plan for RDM policy implementation, which includes recruiting skilled personnel and organising RDM capacity building programmes. Monitoring and evaluation are also critical in this regard to ensure a successful implementation of RDM policy as well as allocation and management of resources. To achieve this, collaboration between faculties and departments should be maintained at all times. Their collaboration, as well as that of the research office and library, will significantly contribute to the improvement of research integrity and culture. Similarly, faculties and departments could make use of team leaders who are at the forefront of promoting a positive research culture, such as research officers, deans/deputy deans research, and department heads. Different departments within faculties could highlight their practices that promote research culture and integrity such as conference planning and organisation.

Library

The role of the institutional library in research is as important as it is in teaching and learning. Faculty and departments, as well as institutional libraries, are responsible for RDM policy implementation. University libraries, on the other hand, should offer RDM capacity building programmes. As Devan (2020) emphasises, one role of academic libraries is training researchers to manage data for better results, reuse, and long-term

success. An institutional library, once again, manages research data repositories that serve as institutional research hubs. Libraries play an important role in research because they have extensive resources in the form of books, journals, magazines, articles, and bibliographies. Furthermore, institutional repositories allow researchers to self-archive their research output, which can improve the visibility, use, and impact of research conducted at a university. Miller (2016: 7) recognises the importance of libraries in facilitating data sharing and use processes via institutional repositories. Knowledge management, research assessment, and open access to scholarly research are some of the other functions of an institutional repository. Furthermore, institutional libraries may be able to market research data services (RDS) available in their respective libraries via their library/institutional websites, social media, pamphlets, and other relevant promotion platforms. Collaboration is also required between the library, faculties, departments, and the research office to promote research culture within the institution.

Research office

The role of the research office is to ensure that RDM policy is implemented in order to facilitate good RDM practices within institutions. In contrast, the research office plays an important role in coordinating with external bodies such as data governance groups such as the National Research Foundation (NRF). The research office, on the other hand, assists in ensuring that applications are submitted and funding for research activities in institutions is secured. Another important role of a university's research office is to ensure that research data services are provided by encouraging RDM capacity building, research visibility, and funding. In terms of capacity building, the research office could organise some training courses and workshops for the institution's researchers and RDM staff. Conferences, research indabas, publications in accredited journals, and other research-related activities that promote research culture in an institution are commonly used to increase research visibility. The research office, on the other hand, could develop some initiatives to promote innovation in the research culture. It is believed that research culture is critical to an institution's research excellence.

Data governance groups

Data governance groups or research units, such as the National Research Foundation (NRF), are critical in developing innovative funding instruments, advancing research career development, increasing public science engagement, and establishing cutting-edge research platforms that will transform the scientific landscape and inspire a diverse research community to strive for institutional global competitiveness. Chiware and Becker (2018: 12) acknowledge the NRF as an example of an accredited institutional repository. The Council for Scientific and Industrial Research (CSIR), on the other hand, is a leading scientific and technological research organisation that assists researchers in developing, localising, and disseminating technologies in order to accelerate South Africa's socioeconomic prosperity. The higher education council (CHE) plays an important role in research by monitoring trends and development. Another function of the CHE is to oversee the implementation of higher education policies. In this regard, collaboration between CHE and DHET, as well as other data governance groups, is absolutely necessary for the successful provision of RDM services in universities.

Researchers

Researchers, including postgraduate students and academic researchers, are the primary drivers of research in an institution, and their research data must always be securely and properly managed. In this sense, researchers are actively engaged in various research practices, such as archiving their research data, publishing in accredited journals, and sharing their preliminary and final research findings through conference presentations. However, successful participation in several research activities necessitates the knowledge and skills of researchers. As a result, an institutional research office, library, faculties, and departments should work together to organise training and workshops to help researchers develop their capabilities. Lyon *et al*. (2012: 25) identify skills and training as important determinants of a community's ability to conduct data-intensive research. In the current study's findings and reviewed

literature, knowledge and skills in data management, research methodology, data curation, metadata skills, and technical skills are overemphasised as truly necessary for RDM. Similarly, researchers can feel more valued if their skill needs are reviewed individually and as a group, ensuring that they all have the necessary skills for their role, such as data collection, data handling, data analysis, proposal writing, data citation, and resource management, among others. Researchers are encouraged to publish and share research data in order to increase research visibility, scrutinise research findings, and understand new developments in a particular field, among other things. Tang and Hu's (2019: 9) study, as well as Unal *et al's* (2019) study, both confirm the importance of data sharing in increasing research visibility.

7.5 Recommendations for further research

As previously stated in section 1.9 (Chapter one) of the dissertation, South Africa has 26 public universities, making it difficult to sample them all along with their employed staff responsible for the provision of RDM and academic researchers. As a result, the study was limited to two institutions in the province of KwaZulu-Natal, namely the Durban University of Technology (DUT) and the University of Zululand (UZ), on how they manage their research data. However, initially, the study targeted three universities, including the University of KwaZulu-Natal, unfortunately there was no response from this institution other than a gatekeepers' letter to the researcher as the study was conducted while COVID-19 and its rules and regulations were still in the effect. Given that the research was limited to the KZN Province, it is suggested that future research be more extensive and include most universities throughout the country with, larger populations. Further research could include private universities and research councils, government entities, national research facilities, data centres, and libraries as research units in the country, which were excluded from the current study due to financial constraints and COVID-19 pandemic rules and regulations in place at the time the study was conducted. Investigation into practical skill development programmes for universities in the country could be more important for training RDM staff and researchers. The study also suggests using the proposed RDM framework for public universities as the first step in offering RDM services, as detailed in Figure 7.1.

7.6 Originality of the study

The scope, subject matter, research methods used, and application of this study all contribute to its uniqueness and novelty. This is the first study in South Africa focusing on research data management services in the country's public universities. According to the reviewed literature search in Chapter three of the dissertation, the subject of RDM has received the most attention in Western nations, as evidenced by most European publications dominating the reviewed literature compared to African ones. RDM literature and publications are still scarce in African countries. This study adds to previous related studies in the country and elsewhere in this sense by providing theoretical background for the development of research, particularly in the Library and Information Science (LIS) domain. Again, contextually, research data management in KwaZulu-Natal universities has not received adequate attention, necessitating the conduct of this study. Indeed, general research on RDM is still in its early stages.

In practice, the study's findings provide an understanding of current research data management services, activities, and tool adoption endorsements, as well as shedding light on areas of support and development in selected public universities. Since both participating institutions are still in the early stages of providing RDM, they could use their existing research-related activities to strengthen their respective institutions' full provision of RDM services. The study has necessitated RDM knowledge and skills for both researchers and staff responsible for RDM provision. As a result, institutions can incorporate those prerequisite skills into their training workshops or RDM short courses. The study revealed that neither participating institution had an RDM policy; however, it is worth noting that one institution's RDM policy is in draft. As a result, the importance of institutions having RDM policies in place to facilitate the management of their research data is highlighted in this study. Institutions can also invest in pre-existing infrastructure, such as institutional repositories, to support RDM. Existing research programmes at the selected institutions could be modified to address RDM issues. The current study's RDM challenges could be turned into opportunities.

On the other hand, this study filled some gaps in previous related studies on RDM in South Africa. Patterton (2016), for example, did not investigate policy and procedure

issues, which the current study did by investigating RDM development strategies such as policy, investment, infrastructure, and capacity building. The current study, on the other hand, proposed an RDM framework for South African public universities, which Chiware and Becker (2018) did not, nor did Patterton and Kahn *et al* (2014). Investigating the current landscape of research data management services in higher education institutions is important and timely for universities that want to improve their data management services as well as those that already provide such services to their researcher user community.

References

Adom, D. & Hussein, E.K. (2018). A.--agyem Joe, “Theoritical and Conceptual Framework: Mandatory Ingrredients Engineering,”. *Int. J. Sci. Res*, *7*(1), 438-441.

Anduvare, E. (2019). eResearch support: an exploratory study of private university libraries in Nairobi County, Kenya. PhD Thesis, University of KwaZulu-Natal, Pietermaritzburg.

Avuglah, B. K. & Underwood, P. G. (2019). Research data management (RDM) capabilities at the University of Ghana, Legon. *Library, Philosophy and Practice,* [Online]. https://digitalcommons.unl.edu/libphilprac/2258/ (15 February 2022).

Avuglah, B. K. (2016). Developing an implementation plan for research data management (RDM) at the University of Ghana. PhD Thesis, University of Pretoria, Pretoria.

Barton, A., Bracke, P. J. & Clark, A. M. (2016). Digitization, data curation, and human rights documents: Case study of a library-researcher-practitioner collaboration. *IaSSIST QUARTERLy*, *40*(1), 27-27.

Birkbeck, G., Nagle, T. & Sammon, D. (2022). Challenges in research data management practices: a literature analysis. *Journal of Decision Systems*, 1-15.

Blanche, M.T., Durrheim, K. & Painter, D. (Editors). (2006) Research in practice: applied methods for the social sciences. Cape Town: UCT Press.

Borghi, J., Abrams, S., Lowenberg, D., Simms, S. & Chodacki, J. (2018). Support your data: A research data management guide for researchers. *Research Ideas and Outcomes*, *4*, e26439.

Bryant, R., Lavoie, B. & Malpas, C. (2017). A Tour of the Research Data Management (RDM) Service Space. The Realities of Research Data Management, Part 1. Dublin, Ohio: OCLC Research. https://doi.org/10.25333/C3PG8J

Bryman, A. (2012). Social research methods. Fifth edition. Oxford: OUP.

Bunkar, A. R. & Bhatt, D. D. (2020). Perception of researchers & academicians of Parul University towards research data management system & role of library: A study. *DESIDOC Journal of Library & Information Technology*, *40*(3).

Cape Peninsula University of Technology Research Data Management (RDM) Policy. (2016). https://www.cput.ac.za/storage/library/policies/CPUT_RDM_Policy_-_final.pdf

Carlson, J., Fosmire, M., Miller, C.C., & Nelson, M.S. (2011). Determining data information literacy needs: A study of students and research faculty. portal: Libraries and the Academy, 11(2), 629-657.

Charles Darwin University Research Data Management Procedures. (2019).

Chawinga, W. D. & Zinn, S. (2019). Global perspectives of research data sharing: A systematic literature review. *Library & Information Science Research*, *41*(2), 109-122.

Chawinga, W. D. (2019). Research data management in public universities in Malawi. PhD Thesis, University of Western Cape, Cape Town.

Chawinga, W. D., & Zinn, S. (2020). Research data management at an African medical university: Implications for academic librarianship. *The Journal of Academic Librarianship*, *46*(4), 102161.

Chigwada, J., Chiparausha, B. & Kasiroori, J. (2017). Research Data Management in Research Institutions in Zimbabwe. *Data Science Journal*, 16(31), 1- 9.

Chiware, E. R., & Becker, D. A. (2018). Research data management services in South Africa: a readiness survey of academic and research libraries*. African Journal of Library Archives and Information Science,* 28(1), 1-16.

Chiware, E., & Mathe, Z. (2016). Academic libraries' role in research data management services: a South African perspective. *South African Journal of Libraries and Information Science*, *81*(2), 1-10.

Cloete, N., Bunting, I., & Van Schalkwyk, F. (2018). *Research universities in Africa* (p. 314). African Minds.

Constantopoulos, P., Dallas, C., Androutsopoulos, I., Angelis, S., Deligiannakis, A., Gavrilis, D. & Papatheodorou, C. (2009). DCC&U: An extended digital curation lifecycle model. *International Journal of Digital Curation*, *4*(1), 34-35.

Corti, L. (2014). *Managing and Sharing Research Data*: *Research Data Management Team*. UK Data Service.

Council on Library and Information Resources. (2013). Research Data Management: Principles, Practices, and Prospects. https://www.clir.org/wp-content/uploads/sites/6/pub160.pdf

Cox, A. M., & Verbaan, E. (2016). How academic librarians, IT staff, and research administrators perceive and relate to research. *Library & Information Science Research*, *38*(4), 319-326.

Cox, A., Kennan, M., Lyon, L., Pinfield, S., & Sbaffi, L. (2019). Maturing research data services and the transformation of academic libraries, *Journal of Documentation*, 75(6). DOI:10.1108/JD-12-2018-0211

Cox, A.M., & Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. J*ournal of Librarianship and Information Science*, 46(4), 299-316.

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 and Technology*, 69(9), 2182-2200.

Crane, D. (2018). Practical strategies for research data management. The Open University. Available at: https://www.slideshare.net/dancrane_open/practical-strategies-for-rdm

Cresswell, J.W. and Cresswell, J.D. (2018). Research design: qualitative, quantitative and mixed methods approaches: Fifth edition. Los Angeles: SAGE.

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Los Angeles: Sage Publications.

Davidson, J., Jones, S., Molloy, L. & Kejser, U.B. (2014). Emerging good practice in managing research data and research information within UK Universities. *Procedia computer science*, *33*, 215-222.

DCC. (2015). 'Five Things You Need to Know About Research Data Management and the Law: DCC Checklist on Legal Aspects of RDM'. DCC Publications. Edinburgh: Digital Curation Centre. Available online: www.dcc.ac.uk/resources

Department of Science and Technology. (2016). South African Research Infrastructure Roadmap. Pretoria.

Devan, N.V. (2020). Role of Academic Libraries in Research. Available at: https://www.researchgate.net/publication/339473798_Role_of_Academic_Libraries_in_Research/citations

Fischer, F. (1998). Beyond empiricism: policy inquiry in post positivist perspective. *Policy studies journal*, *26*(1), 129-146.

Flores, J.R., Brodeur, J. J., Daniels, M. G., Nicholls, N., & Turnator, E. (2015). Libraries and the research data management landscape. *The process of discovery: The CLIR postdoctoral fellowship program and the future of the academy*, *2010*, 82-102.

Fox, N. J. (2008). Post-positivism. *The SAGE encyclopedia of qualitative research methods*. London: Sage. *2*, 659-664.

Giarlo, M. J. (2013). Academic Libraries as Data Quality Hubs. *Journal of Librarianship & Scholarly Communication*, *1*(3).

Gordon, A. S., Millman, D. S., Steiger, L., Adolph, K. E., & Gilmore, R. O. (2015). Researcher-Library Collaborations: Data Repositories as a Service for Researchers. *Journal of Librarianship and Scholarly Communication,* 3(2), eP1238. http://dx.doi.org/10.7710/2162-3309.1238

Green, A., Macdonald, S., & Rice, R. (2009). *Policy-making for Research Data in Respositories: A Guide*. EDINA.

Gunjal, B., & Gaitanou, P. (2017). Research Data Management: A proposed framework to boost research in Higher Educational Institutes.: A proposed framework to boost research in Higher Educational Institutes. *IASSIST Quarterly*, 41(1-4), 12-12.

Gustafsson, J. (2017). Single case studies vs. multiple case studies: A comparative study. Available at: https://www.diva-portal.org/smash/get/diva2:1064378/FULLTEXT01.pdf

Halbert, M. (2013). Prospects for Research Data Management. In: Council on Library and Information Resources. Research Data Management: Principles, Practices and Prospects CLIR Report, 160, November 2013, pp.1-15. Available at: http://www.clir.org/pubs/reports/pub160/pub160.pdf

Hamad, F., Al-Fadel, M., & Al-Soub, A. (2021). Awareness of research data management services at academic libraries in Jordan: Roles, responsibilities and challenges. *New Review of Academic Librarianship*, *27*(1), 76-96.

Higgins, S. (2008). The DCC curation lifecycle model. *The International Journal of Digital Curation*, 1(3), 134-140.

https://datascience.codata.org/articles/10.5334/dsj-2019-038/

Igwenagu, C. (2016). *Fundamentals of research methodology and data collection*. LAP Lambert Academic Publishing.

Jones, S. (2013). Bringing it all together: a case study on the improvement of research data management at Monash University. DCC RDM Services Case Studies.

Jones, S., Ball, A., & Ekmekcioglu, Ç. (2008). The data audit framework: A first step in the data management challenge.

Jones, S., Pryor, G., & Whyte, A. (2013). Develop Research Data Management Services - a guide for HEIs'. *DCC How-to Guides*. Edinburgh: Digital Curation Centre. Available at: http://www.dcc.ac.uk/resources/how-guides

Jones, S., Ross, S., & Ruusalepp, R. (2009). The Data Audit Framework: a toolkit to identify research assets and improve data management in research led institutions.

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.

Kirub, A. (2017). Research Data Management Policy. Ethiopian Institute of Agricultural Research.

Knight, G. (2012). A digital curate's egg: a risk management approach to enhancing data management practices. *Journal of web librarianship*, *6*(4), 228-250.

Koopman, M.M., & De Jager, K. (2016). Archiving South African digital research data: How ready are we?. *South African Journal of Science*, 112(7-8), pp.1-7.

Kumar, R. (2011). *Research methodology: a step-by-step guide for beginners*. 4th ed. Los Angeles: Sage.

Kwanya, T. (2021). Publishing trends on research data management in Sub-Saharan Africa: A bibliometrics analysis. *IASSIST Quarterly*, *45*(3-4).

Lee, D. J., & Stvilia, B. (2017). Practices of research data curation in institutional repositories: A qualitative view from repository staff. *PloS one*, *12*(3), e0173987.

Leedy, P.D. & Ormrod, J.E. (2013). Practical research: Planning and design: Tenth edition. Boston: Pearson.

Lincoln, Y., & Guba, E. G. (1985). Naturalistic inquiry. NewburyPark, CA: Sage.

Liu, G., Zotoo, I. K., & Su, W. (2020). Research data management policies in USA, UK and Australia universities: An online survey. *Malaysian Journal of Library & Information Science*, *25*(2), 21-42.

Lötter, L. (2014). *Reflections on the RDM Position in South Africa: Development Research Data Management Services*. Cape Town, South Africa: Research Methodology Centre.

Lyon, L., Ball, A., Duke, M., & Day, M. (2012). Developing a Community Capability Model Framework for data-intensive research. In *Proceedings of the 9th International Conference on the Preservation of Digital Objects* (pp. 9-16).

Lyon, L., Patel, M., & Takeda, K. (2014). Assessing requirements for research data management support in academic libraries: introducing a new multi-faceted capability tool. *Libraries in the Digital Age*,131.

Macanda, M., Rammutloa, M., & Bezuidenhout, R. (2015). Research data management at Unisa. http://hdl.handle.net/10500/13907

Macdonald, S., & Martinez-Uribe, L. (2010). Collaboration to data curation: Harnessing institutional expertise. *New Review of Academic Librarianship*, *16*(S1), 4-16. doi:10.1080/13614533.2010.505823.

Mamba, J. T. (2019). Enhancing career decision making in senior phase in one of secondary schools in Umlazi district. Masters Thesis, University of KwaZulu-Natal, Durban.

Maree, K. (ed) (2016). First steps in research 2. Pretoria: Van Schaik publishers.

Masinde, J. M., Chen, J., & Muthee, D. W. (2021). Researchers' Perceptions of Research Data Management Activities at an Academic Library in a Developing Country. *International Journal of Library and Information Services (IJLIS)*, *10*(2), 1-17.

Maurya, A. and Subaveerapandiyan, A. (2022). Research Data Preservation Practices of Library and Information Science Faculties. *Journal of Library & Information Technology*, 42(4), 259-264.

Mavodza, J. (2022). Research Data Management: A review of UAE academic library experience. *Open Information Science*, *6*(1), 16-27.

Mayne, R., Hamilton, J., & Hobson, K. (2013). Step by Step Guide to Monitoring and Evaluation. Resource.

Michener, W. K. (2018). Project Data Management Planning. In Ecological Informatics (pp. 13-26). Springer, Cham.

Miller, M., (2016). Research Data Management in Canada Independent Study LIS 9410. https://lib.fims.uwo.ca/wp-content/uploads/2016/04/LIS9410_Miller.pdf

Molloy, L., & Snow, K. (2012). The data management skills support initiative: Synthesising postgraduate training in research data management.

Morgan, A., Duffield, N., & Walkley Hall, L. (2017). Research data management support: sharing our experiences. *Journal of the Australian Library and Information Association*, 66(3): 299-305.

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods*, *1*(2), 13-22.

Mosha, N. F., & Ngulube, P. (2022). Teaching research data management courses in higher learning institutions in Tanzania. Library Management, (ahead-of-print).

Mushi, G. E. (2021). Research data management and services: Resources for different data practitioners. Available at: https://iassistquarterly.com/index.php/iassist/article/view/995

Mushi, G. E., Pienaar, H., & van Deventer, M. (2020). Identifying and Implementing Relevant Research Data Management Services for the Library at the University of Dodoma. *Tanzania. Data Science Journal*, 19(1), 1- 9.

Naroş, M. S., & Simionescu, M. (2019). The role of education in ensuring skilled human capital for companies. Theoretical and Empirical Researches in Urban Management, 14(1), 75-84.

Neuman, W. (2014). Pearson new international edition: social research methods: qualitative and quantitative approaches. Seventh edition. Pearson: London.

Neuman, W. L. (2000). Social Research Methods: Qualitative and Quantitative Approaches. Boston, MA: Allyn & Bacon.

Ng'eno, E., & Mutula, S. (2018). Research Data Management (RDM) in agricultural research institutes: a literature review. *Inkanyiso: Journal of Humanities and Social Sciences*, *10*(1), 28-50.

Ng'eno, E. J. (2018). Research Data Management in Kenya's Agricultural Research Institutes. PhD Thesis, University of KwaZulu-Natal, Pietermaritzburg.

Ngulube, P. (2015). Trends in research methodological procedures used in knowledge management studies. *African Journal of Library, Archives and Information Science*, *25*(2), 125-143.

Nhendodzashe, N., & Pasipamire, N. (2017). Research data management services: are academic libraries in Zimbabwe ready? The case of the University of Zimbabwe library. [Online]. http://library.ifla.org/id/eprint/1728/1/S06-nhendodzashe-en.pdf (13 August 2021).

Nwabugwu, M. J., & Godwin, L. S. (2020). Research data management (RDM) services in libraries: lessons for academic libraries in Nigeria. *Library Philosophy and Practice (e-journal)*, *4238*.

Ocholla, D. (2020). Decolonizing higher education in Africa: Implications and possibilities for university libraries. *College & Research Libraries News*, *81*(6), 289.

Ocholla, D. N., & Ocholla, L. (2020). Readiness of academic libraries in South Africa to research, teaching and learning support in the Fourth Industrial Revolution. *Library Management*.

Organisation for Economic Co-Operation and Development. (2007). Principles and Guidelines for Access to Research Data from Public Funding. Available at: https://www.oecd.org/sti/inno/38500813.pdf

Osanloo, A., & Grant, C. (2016). Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for your "house". *Administrative issues journal: connecting education, practice, and research*, *4*(2), 7.

Palavitsinis, N., Manouselis, N., & Sanchez-Alonso, S. (2010). Preliminary Discussion on a Digital Curation Framework for Learning Repositories. In *Ceur Workshop Proceedings* (pp. 46-50).

Pandey, P., & Pandey, M. M. (2015). *Research methodology tools and techniques*. Bridge Center.

Panhwar, A. H., Ansari, S., & Shah, A. A. (2017). Post-positivism: An effective paradigm for social and educational research. *International Research Journal of Arts & Humanities (IRJAH)*, *45*(45).

Patterton, L. H. (2016). *Research data management practices of emerging researchers at a South African research council*. PhD Thesis, University of Pretoria, Pretoria.

Phillips, D. C., & Burbules, N. C. (2000). *Postpositivism and educational research*. Rowman & Littlefield.

Pinfield, S., Cox, A. M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS One*, *9*(12), e114734.

Prathapan, K. (2014). Research methodology for scientific research. I.K. International publishing house: New Delhi.

Qin, J. (2013). Infrastructure, Standards, and Policies for Research Data Management. In: Sharing of Scientific and Technical Resources in the Era of Big Data: The Proceedings of COINFO 2013, pp. 214-219. Beijing: Science Press.

Rice, R. (2022). A Decade of Research Data Management at the University of Edinburgh: Looking Back, Looking Forward. In *Handbook of Research on Academic Libraries as Partners in Data Science Ecosystems* (pp. 308-333). IGI Global.

Sedgwick, P. (2015). Multistage sampling. *Bmj*, *351*. Available at: https://www.researchgate.net/profile/Philip-Sedgwick/publication/281107209_Multistage_sampling/links/55d5bfe008aeb38e8a8036a0/Multistage-sampling.pdf

Senft, M., Stahl, U., & Svoboda, N. (2022). Research data management in agricultural sciences in Germany: We are not yet where we want to be. *Plos one*, *17*(9), e0274677.

Shareia, B. F. (2016). Qualitative and quantitative case study research method on social science: Accounting perspective. *International Journal of Economics and Management Engineering*, *10*(12), 3849-3854.

Shearer, K. (2015). A comprehensive brief on research data management policies. [Online]. http://ozk.unizd.hr/proceedings/index.php/lida/article/viewFile/121/123 (13 September 2021).

Shen, Y. (2016). Strategic planning for a data-driven, shared-access research enterprise: Virginia Tech research data assessment and landscape study. *College & Research Libraries*, *77*(4), 500-519.

Shen, Y., & Varvel Jr, V. E. (2013). Developing data management services at the Johns Hopkins University. *The Journal of Academic Librarianship*, *39*(6), 552-557.

Stahl, N. A., & King, J. R. (2020). Expanding approaches for research: Understanding and using trustworthiness in qualitative research. *Journal of Developmental Education*, *44*(1), 26-28.

Stellenbosch University Research Data Management Regulations. (2020). Available at: https://www.sun.ac.za/english/Documents/Terms_and_conditions/Current/RDM_Regulations_English.pdf

Struwig, F. W., & Stead, G. B. (2001). Understanding reliability and validity. *Andrea Nattrass.(3 rd ed. 2004). Planning, designing and reporting research. Cape Town: Hanli Venter*, 130-142.

Tang, R., & Hu, Z. (2019). Providing research data management (RDM) services in libraries: Preparedness, roles, challenges, and training for RDM practice. *Data and information management*, *3*(2), 84-101.

Tenopir, C., Birch, B., & Allard, S. (2012). Academic libraries and research data services: Current practices and plans for the future. https://trace.tennessee.edu/cgi/viewcontent.cgi?article=1019&context=utk_datao ne

Tenopir, C., Rice, N., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. PLoS One, 15(3), 15. doi:10.1371/journal.pone.0229003 PMID:32160189.

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(2), 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.J., & Allard, S. (2016). Research Data Services in European Academic Research Libraries. Submitted to LIBER QUARTERLY. Available at: http://libereurope.eu/wpcontent/uploads/2016/10/LIBER_RDM_SURVEY_FIN AL_MANUSCRIPT.pdf [Last accessed 04 April 2017]

The Australian National Data Service (ANDS). (2018). Creating a data management framework. Available at: www.ands.org.au/guides/creating-a-data-management-framework

The University of Cape Town Research Data Management Policy. (2018). Available at: http://www.uct.ac.za/sites/default/files/image_tool/images/328/about/policies/TG O_Policy_Research_Data_Management_2018.pdf

The University of Southampton Research Data Management Policy. (2019). Available at: https://www.southampton.ac.uk/~assets/doc/calendar/Research%20Data%20Ma nagement%20Policy.pdf

Tripathi, M., Shukla, A., & Sonkar, S.K. (2017). Research Data Management practices in university libraries: A study. *DESIDOC Journal of Library & Information Technology*, *37*(6), p.417.

Trochim, W. M. K. (2020). Research methods knowledge base: Conclusion validity.

Unal, Y., Chowdhury, G., Kurbanoglu, S., Boustany, J., & Walton, G. (2019). Research data management and data sharing behaviour of university researchers. *Information Research,* 24(1). [Online]. http://informationr.net/ir/24-1/isic2018/isic1818.html (12 September 2021).

Unit, C. I. A. (2017). CGIAR Internal Audit Unit-Good Practice Notes. Available at: https://cgspace.cgiar.org/bitstream/handle/10947/4664/GPN-Research-Data-Management-Sept-2017.pdf?sequence=4&isAllowed=y

University of Edinburgh website. (2019). Data Management Training. Available at: https://www.ed.ac.uk/institute-academic-development/research-roles/research-only-staff/research-good-practice/data-management

University of Pretoria Research Data Management Policy. (2017). Available at: https://www.up.ac.za/media/shared/12/ZP_Files/research-data-management-policy_august-2018.zp161094.pdf

University of the Western Cape Research Data Management Policy. (2021). Available at: https://eresearch.uwc.ac.za/wp-content/uploads/2021/04/UWC-Research-Policy_Section-13-C2021.01_RDM-Policy.pdf

University of Twente Research Data Management Policy. (2018). Available at: https://www.utwente.nl/onderzoek/onderzoekorganisatie/wetenschappelijke-integriteit/downloads/research-data-management-policy-2018.pdf

Van den Eynden, V., Corti, L., Woollard, M., Bishop, L., & Horton, L. (2011). Managing and sharing data; a best practice guide for researchers.

Van Deventer, M., & Pienaar, H. (2015). Research data management in a developing country: a personal journey.

Van Wyk, B. J. (2018). The relationship between research data management and virtual research environments. PhD Thesis, University of Pretoria, Pretoria.

Van Wyk, J. (2017). Facilitate Research Data Management (RDM) – Project. Available at:https://repository.up.ac.za/bitstream/handle/2263/61130/VanWyk_RDM_2017.pdf?sequence=2&isAllowed=y

Wachtler, T., Bauer, P., Denker, M., Grün, S., Hanke, M., Klein, J., ... & Witte, O. W. (2021). NFDI-Neuro: building a community for neuroscience research data management in Germany. *Neuroforum*, *27*(1), 3-15.

Welman, C., Kruger, F., & Mitchell, B. (2005). *Research methodology*. Oxford University Press.

Whyte, A., & Allard, S. (2014). How to discover requirements for research data management services.

Whyte, A., & Tedds, J. (2011). *Making the case for research data management*. Digital Curation Centre.

Wiljes, C., & Cimiano, P. (2019). Teaching research data management for students. *Data Science Journal*, *18*, 38-38.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., & Bouwman, J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific data, 3(1), pp.1-9.

Williman, N. (2011). Research methods: The basics. Routledge.

Woeber, C.A. (2017). Towards best practice in research data management in the humanities. https://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/6620/paper_access.pdf?sequence=4 (15 September 2021).

Xu, Z., Zhou, X., Kogut, A., & Clough, M. (2022). Effect of online research data management instruction on social science graduate students' RDM skills. *Library & Information Science Research*, *44*(4), 101190.

Yoon, A., & Schultz, T. (2017). Research data management services in academic libraries in the US: A content analysis of libraries' websites. [Online]. https://scholarworks.iupui.edu/bitstream/handle/1805/14535/16788-20911-1-PB.pdf?sequence=1(12 September 2021).

Appendices

Appendix A: Letter of request to public universities

University of Zululand
Private Bag X1001
KwaDlangezwa
3886

Letter of request to selected public universities

To The Research Office

RE: Introducing Mthembu Mpilo Siphamandla (201256633) – PhD Student at the University of Zululand

This letter serves to confirm that Mr Mthembu Mpilo Siphamandla is registered for a Doctor of Philosophy in Information Science qualification in the Department of Information Studies at the University of Zululand. The title of his research is ***Research Data Management Services in Selected Public Universities in KwaZulu-Natal, South Africa***. We note that such an investigation has implications on many stakeholders as it will provide theoretical background for the development of research in the public universities, staff responsible for RDM, LIS researchers and LIS curriculum developers. The study outcome will also be used widely for research, teaching and learning. Additionally, the study will inform RDM strategy and policy in South Africa, and perhaps, elsewhere.

As part of the requirements for the award of a Doctor of Philosophy degree, the researcher is expected to undertake original research in an environment and place of his choice. The UNIZULU ethical compliance regulations require him to provide proof (letter or email) that the relevant authority where the research is to be undertaken has given approval. We thus request you as his chosen organisation to grant him the required permission to conduct his research.

We appreciate your support and understanding in this regard. Should you need any further clarification, do not hesitate to contact us:

Thanks,

Supervisor: Prof. D.N. Ocholla
Email: OchollaD@unizulu.ac.za
Mobile: 082 372 4638

Researcher: Mr. M.S. Mthembu
Email: MthembuSM@unizulu.ac.za
Mobile: 071 170 2832

Appendix B: Gatekeepers Letter from DUT

Directorate for Research and Postgraduate Support
Durban University of Technology
Tromso Annexe, Steve Biko Campus
P.O. Box 1334, Durban 4000
Tel.: 031-3732576/7
Fax: 031-3732946

8th June 2021
Mr Mpilo S Mthembu
c/o Department of Library Information Science
Faculty of Arts
University of Zululand

Dear Mr Mthembu

PERMISSION TO CONDUCT RESEARCH AT THE DUT

Your email correspondence in respect of the above refers. I am pleased to inform you that the Institutional Research and Innovation Committee (IRIC) has granted **Gatekeeper Permission** for you to conduct your research "Research Data Management Services in Selected Public Universities in KwaZulu-Natal, South Africa" at the Durban University of Technology. **Kindly note that this letter must be issued to the IREC for approval before you commence data collection.**

The DUT may impose any other condition it deems appropriate in the circumstances having regard to nature and extent of access to and use of information requested.

We would be grateful if a summary of your key research findings would be submitted to the IRIC on completion of your studies.

Kindest regards.
Yours sincerely

DR LINDA ZIKHONA LINGANISO
DIRECTOR: RESEARCH AND POSTGRADUATE SUPPORT DIRECTORATE

Appendix C: Gatekeepers Letter from DUT library

PERMISSION LETTER

28 September 2021

Mr Mpilo Siphamandla Mthembu
University of Zululand

PERMISSION TO CONDUCT RESEARCH AT THE DUT LIBRARY

Dear Mr Mthembu

This letter serves as authorisation for you to conduct the research study entitled, ***"Research Data Management services in selected public universities in KwaZulu-Natal, South Africa"***, at the DUT Library.

Upon review of the associated documentation submitted to us by yourself, we are glad to offer you an opportunity to conduct the said study in the Library. The administering of the online questionnaire and interviews is approved and will be duly supervised by the Manager: Academic Services. If you have any concerns or require additional information, feel free to contact Mr David Thomas (davidt@dut.ac.za).

Thank you

Yours faithfully,

Dr M. P. Phaladi (Library Director)

Appendix D: Ethical Clearance Certificate

UNIVERSITY OF ZULULAND RESEARCH ETHICS COMMITTEE
(Reg No: UZREC 171110-030)

UNIVERSITY OF ZULULAND

ETHICAL CLEARANCE CERTIFICATE - RENEWAL

Certificate Number	**UZREC 171110-030 PGD 2020/41**								
Project Title	**Research data management services in selected public universities in KwaZulu-Natal, South Africa**								
Principal Researcher/ Investigator	M S Mthembu								
Supervisor and Co-supervisor	Prof DN Ocholla								
Department	Information Studies								
Faculty	Humanities and Social Sciences								
Type of Risk	Medium Risk – Data collection from people								
Nature of Project	Honours/4th Year		Master's		**Doctoral**	X	Departmental		

The University of Zululand's Research Ethics Committee (UZREC) hereby gives ethical **renewal** approval in respect of the undertakings contained in the above-mentioned project. This approval is extended for another 1 year. The Researcher may therefore continue with data collection as from the date of this Certificate, using the certificate number indicated above.

SPECIAL CONDITIONS: **(1) This certificate is valid for 1 year from the date of issue.**
(2) Principal researcher must provide an annual report to the UZREC in the prescribed format [due date- 28 November 2023]
(3) The UZREC must be informed immediately of any material change in the conditions or undertakings mentioned in the documents that were presented to the meeting.
(4) Under the Protection of Personal Information Act, 04 of 2013 ("POPIA"), researchers have a general legal duty to protect information they process. They must ensure the security and protection of any personal information processed through the research and provide a compliant and consistent approach to data protection. The information collected via interviews must be for research purposes only. No personal information such as opinions, views and academic background may be linked to the respondents' identity or shared with anyone for marketing purposes or otherwise.

The UZREC wishes the researcher well in conducting research.

Prof. Nokuthula Kunene
Chairperson: University Research Ethics Committee
Deputy Vice-Chancellor: Research & Innovation
28 November 2022

CHAIRPERSON
UNIVERSITY OF ZULULAND RESEARCH ETHICS COMMITTEE (UZREC)
REG NO: UZREC 171110-30
2022 -11- 28
RESEARCH & INNOVATION OFFICE

RESEARCH & INNOVATION OFFICE
Website: http://www.unizulu.ac.za
Private Bag X1001
KwaDlangezwa, 3886
Tel: 035 902 6374/6324
Email: MthembuNL@unizulu.ac.za/MangeleS@unizulu.ac.za

Appendix E: Proof of editing

Declaration

Dr Ian Robin Raper
rosslyni@lantic.net
078 761 2241
17 Sabi Road, Selcourt, Springs 1559

To whom it may concern

I hereby declare that I have comprehensively edited the use of language in the thesis to be submitted by Mpilo Siphamandla Mthembu in fulfilment of the requirements for the degree of Doctor of Philosophy.

The supervisor for the candidate's research is Professor Dennis Ocholla.

Title: Research Data Management Services in Selected Public Universities in KwaZulu-Natal

I received the manuscripts in MS-Word format, and have performed all the language and style editing required. The edited thesis was returned to the candidate in MS-Word format with the tracking details indicated.

Ian Robin Raper

Ian Raper

D Litt et Phil (RAU), MA (WITS), THED (JCE)

Appendix F: Interview guide for Librarians

University of Zululand

Faculty of Humanities and Social Science (HSS)

Department of Information Studies

Interview guide

Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa

Dear réspondent,

You are invited to participate in a research study titled ***"Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa"***. This study is being conducted by Mpilo Siphamandla Mthembu, à PhD student in the Department of Information Studies (DIS) at the University of Zululand. It is noted that such an investigation will be used widely for research, teaching and learning, inform RDM strategy and policy in South Africa and perhaps elsewhere.

As part of the requirements for the fulfilment of the Doctor of Philosophy (Information Science) degree, I kindly ask for your participation in my research project by taking part in this interview. Your participation in this study is voluntary and you are free to withdraw your participation from this study at any time. There are no risks associated with participating in this study. Please be assured that this survey collects no identifying information of any respondent. All of the responses in the survey questionnaire will be recorded anonymously. By completing and submitting this survey questionnaire, you are indicating your consent to participate in the study. Your participation is appreciated and looking forward to sharing the results of the study with you. For any queries, please do not hesitate to contact us:

1. **Mpilo Siphamandla Mthembu**, Doctoral candidate, Department of Information Studies, University of Zululand
   Email: mpilomthembu55@gmail.com
   Cell: 071 170 2832

2. **Prof Dennis Ngong Ocholla**, Research supervisor, Department of Information Studies, University of Zululand
   Email: OchollaD@unizulu.ac.za
   Cell: 082 372 4638

Thank you

Instructions

This interview is divided into six (6) sections and should take about 20 minutes.

SECTION A: Biographical data
SECTION B: RDM practices at the selected public universities
SECTION C: Competency requirements of researchers and staff responsible for research data management in selected public universities.
SECTION D: RDM development strategies at the selected public universities.
SECTION E: Challenges facing the selected public universities in managing research data.
SECTION F: RDM framework for public universities in South Africa.

Questions

SECTION A: Biographical data

1. What is your gender?
   Male [ ]
   Female [ ]

2. How old are you?
   25-30 [ ]
   35-40 [ ]
   45-50 [ ]
   55-60 [ ]
   60-above [ ]

3. What is your population group?
   African [ ]
   White [ ]
   Coloured [ ]
   Indian/ Asian [ ]

4. What is the name of your institution?
   University of Zululand (UZ) [ ]
   University of KwaZulu-Natal (UKZN) [ ]
   Durban University of Technology (DUT) [ ]

5. What is your highest academic qualification?
6. What is your current position?
7. How long have you been working in your position?
8. What are your research related duties in your position?

SECTION B: RDM practices at the selected public universities

9. What activities or practices do your department/section provide in support for the following research data activities:
    a) Data creation
    b) Data storage
    c) Data sharing
    d) Data preservation
    e) Data archiving
    f) Data mining
    g) Data access, use and reuse
    h) Data security
10. Where do you currently keep, share and manage the published research data in your library?
11. Do you have an institutional repository in your library?
12. Are there any measures you have put to enable researchers to self-archive their research data in the repositories?
13. Is your institutional repository open, institutional or non-open access?
14. What format do you prefer for data to be captured in your system?
15. What is the amount of research data in gigabytes (GB) or terabytes (TB) can you keep in your system?
16. What are the requirements of a researcher to publish his/her work?
17. For how long do you keep research data in your library?
18. Do you think most researchers, particularly emerging researchers (post graduates) are aware of where they need to publish their research work and what needs to be done?
19. What do you think are the benefits of RDM on emerging researchers?

SECTION C: Competency requirements of staff responsible for research data management in selected public universities.

20. What knowledge and skills do you require to support RDM services?
21. Do you think you need some new skills and knowledge in your current position?
22. How do you think you can acquire needed competencies?
23. Do you offer any programmes or activities like workshops to support researchers on their RDM practices?

SECTION D: RDM development strategies at the selected public universities

24. What policies does your library have to support RDM?
25. If **yes**, how long is the policy or procedure?
26. How do you prevent unauthorised access from your research data?

27. Does your department have any collaborative partnerships within the institution or other third parties in support for RDM?
28. Do you think you have adequate resources or infrastructure for RDM practices in your section?
29. Do you think your department have sufficient human resources to provide RDM services to provide good RDM services?
30. Does your department have RDM activity with regards to awareness and capacity building? E.g. for organising workshops, seminars, conferences?

SECTION E: Challenges facing the selected public universities in managing research data

31. What challenges do you face when providing research data management services in your library?
32. What suggestions would you recommend in overcoming the challenges of RDM?

SECTION F: RDM framework for public universities in South Africa

33. In your opinion, what do you think can be done to develop a RDM framework for public universities in South Africa?

Thank you so much for your cooperation

Appendix G: Interview guide for Technicians

University of Zululand

Faculty of Humanities and Social Science (HSS)

Department of Information Studies

Interview guide

Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa

Dear réspondent,

You are invited to participate in a research study titled ***"Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa".*** This study is being conducted by Mpilo Siphamandla Mthembu, à PhD student in the Department of Information Studies (DIS) at the University of Zululand. It is noted that such an investigation will be used widely for research, teaching and learning, inform RDM strategy and policy in South Africa and perhaps elsewhere.

As part of the requirements for the fulfilment of the Doctor of Philosophy (Information Science) degree**,** I kindly ask for your participation in my research project by taking part in this interview. Your participation in this study is voluntary and you are free to withdraw your participation from this study at any time. There are no risks associated with participating in this study. Please be assured that this interview collects no identifying information of any respondent. All of the responses in the interview will be recorded anonymously. Your participation is appreciated and looking forward to sharing the results of the study with you.

For any queries pertaining this research, please do not hesitate to contact us:

3. **Mpilo Siphamandla Mthembu**, Doctoral candidate, Department of Information Studies, University of Zululand
   Email: mpilomthembu55@gmail.com
   Cell: 071 170 2832

4. **Prof Dennis Ngong Ocholla**, Research supervisor, Department of Information Studies, University of Zululand
   Email: OchollaD@unizulu.ac.za
   Cell: 082 372 4638

Thank you

Instructions

This interview is divided into six (6) sections and should take about 20 minutes.

SECTION A: Biographical data
SECTION B: RDM practices at the selected public universities
SECTION C: Competency requirements of researchers and staff responsible for research data management in selected public universities.
SECTION D: RDM development strategies at the selected public universities.
SECTION E: Challenges facing the selected public universities in managing research data.
SECTION F: RDM framework for public universities in South Africa.

Questions

SECTION A: Biographical data

34. What is your gender?
    Male [ ]
    Female [ ]

35. How old are you?
    25-30 [ ]
    35-40 [ ]
    45-50 [ ]
    55-60 [ ]
    60-above [ ]

36. What is your population group?
    African [ ]
    White [ ]
    Coloured [ ]
    Indian/ Asian [ ]

37. What is the name of your institution?
    University of Zululand (UZ) [ ]
    University of KwaZulu-Natal (UKZN) [ ]
    Durban University of Technology (DUT) [ ]

38. What is your highest academic qualification?

39. What is your current position?
40. How long have you been working in your position?

SECTION B: RDM practices at the selected public universities

41. Could you please describe your role in IT section in support for RDM undertakings?
42. What activities or practices does your library provide in support for the following research data activities:
    i) Data creation
    j) Data storage
    k) Data sharing
    l) Data preservation
    m) Data archiving
    n) Data mining
    o) Data access, use and reuse
    p) Data security
43. Where do you currently keep, share and manage the published research data in your library?
44. How adequate is your IT infrastructure in support for RDM services in terms of storage and network bandwidth?
45. Do you have an institutional repository in your library?
46. Who is managing the institutional repository in your library?
47. Are there any measures you have put to enable researchers to self-archive their research data in the repositories?
48. Is your institutional repository open, institutional or non-open access?
49. What format do you prefer for data to be captured in your system?
50. What challenges do you encounter to store data in the institutional repository?
51. What is the amount of research data in gigabytes (GB), or terabytes (TB) can you keep in your system?
52. What are the requirements of a researcher to publish his/her work?
53. For how long do you keep research data in your library?
54. Do you think most researchers, particularly emerging researchers (postgraduates) are aware of where they need to publish their research work and what needs to be done?
55. What do you think are the benefits of RDM on emerging researchers?
56. Is there any plan in your library for implementing services to support RDM in the future?

SECTION C: Competency requirements of staff responsible for research data management in selected public universities.

57. What knowledge and skills do you require to support RDM services?
58. Do you think you need some new skills and knowledge in your current position?
59. How do you think you can acquire needed competencies?
60. Do you offer any programmes or activities like workshops to support researchers on their RDM practices?

SECTION D: RDM development strategies at the selected public universities

61. What policies does your library have to support RDM?
62. If **yes**, how long is the policy or procedure?
63. How do you prevent unauthorised access from your research data?
64. What internal capacity building programmes and strategies are available to develop skills for providing RDM services?
65. Are there any external opportunities for developing staff in terms of RDM have been availed?
66. Do you think you have adequate resources or infrastructure for RDM practices in your section to provide worthy RDM services?
67. Do you think your department has sufficient human resources to provide RDM services?

SECTION E: Challenges facing the selected public universities in managing research data

68. What challenges do you face when providing research data management services in your library?
69. What suggestions would you recommend in overcoming the challenges of RDM?

SECTION F: RDM framework for public universities in South Africa

70. In your opinion, what do you think can be done to develop a RDM framework for public universities in South Africa?

Thank you so much for your cooperation

Appendix H: Interview guide for Research Directors, Deans/Deputy Deans Research, Deputy Vice-Chancellors Research and Heads of Departments

University of Zululand

Faculty of Humanities and Social Science (HSS)

Department of Information Studies

Interview guide

Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa

Dear répondent,

You are invited to participate in a research study titled ***"Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa".*** This study is being conducted by Mpilo Siphamandla Mthembu, à PhD student in the Department of Information Studies (DIS) at the University of Zululand. It is noted that such an investigation will be used widely for research, teaching and learning, inform RDM strategy and policy in South Africa and perhaps elsewhere.

As part of the requirements for the fulfilment of the Doctor of Philosophy (Information Science) degree**,** I kindly ask for your participation in my research project by taking part in this interview. Your participation in this study is voluntary and you are free to withdraw your participation from this study at any time. There are no risks associated with participating in this study. Please be assured that this interview collects no identifying information of any respondent. All of the responses in the interview will be recorded anonymously. Your participation is appreciated and looking forward to sharing the results of the study with you.

For any queries pertaining this research, please do not hesitate to contact us:

5. **Mpilo Siphamandla Mthembu**, Doctoral candidate, Department of Information Studies, University of Zululand

Email: mpilomthembu55@gmail.com
Cell: 071 170 2832

6. **Prof Dennis Ngong Ocholla**, Research supervisor, Department of Information Studies, University of Zululand
   Email: OchollaD@unizulu.ac.za
   Cell: 082 372 4638

Thank you

Instructions

This interview is divided into six (6) sections and should take about 20 minutes.

SECTION A: Biographical data
SECTION B: RDM practices at the selected public universities
SECTION C: Competency requirements of researchers and staff responsible for research data management in selected public universities.
SECTION D: RDM development strategies at the selected public universities.
SECTION E: Challenges facing the selected public universities in managing research data.
SECTION F: RDM framework for public universities in South Africa.

Questions

SECTION A: Biographical data

71. What is your gender?
    Male [ ]
    Female [ ]

72. How old are you?
    25-30 [ ]
    35-40 [ ]
    45-50 [ ]
    55-60 [ ]
    60-above [ ]

73. What is your population group?
    African [ ]
    White [ ]
    Coloured [ ]
    Indian/ Asian [ ]

74. What is the name of your institution?

University of Zululand (UZ) [ ]
University of KwaZulu-Natal (UKZN) [ ]
Durban University of Technology (DUT) [ ]

75. What is your highest academic qualification?
76. What is your current position?
77. How long have you been working in your position?
78. What is your niche area in research?
79. What are your research duties in your position?

SECTION B: RDM practices at the selected public universities

80. What activities or practices do your department/section provide in support for the following research data activities:
    q) Data creation
    r) Data storage
    s) Data sharing
    t) Data preservation
    u) Data archiving
    v) Data mining
    w) Data access, use and reuse
    x) Data security
81. Where do you keep published works of the researchers in your department?
82. What are the requirements of a researcher to publish his/her work?
83. For how long do you keep researchers' data in your section?
84. Do you think most researchers, particularly emerging researchers (postgraduates) are aware of where they need to publish their research work and what needs to be done?
85. What do you think are the benefits of RDM on emerging researchers?

SECTION C: Competency requirements of staff responsible for research data management in selected public universities.

86. What knowledge and skills do you need to perform RDM activities in your section?
87. Who are you working with within your department, faculty and institution in terms of delivering RDM services in the organisation?
88. Do you think you need some new skills and knowledge in your current position?
89. How do you think you can acquire needed competencies?
90. Do you offer any programmes or activities like workshops to support researchers on their RDM practices?

SECTION D: RDM development strategies at the selected public universities

91. What policies does your university have to support RDM?
92. If **yes**, what is the name of the policy and how long is that policy or procedure?
93. What exactly does it do?
94. Who owns the intellectual property rights for research data in the institution?

95. What internal capacity building programmes and strategies are available to develop skills for providing RDM services?
96. Are there any external opportunities for developing staff in terms of RDM have been availed? Do you think you have adequate resources or infrastructure for RDM practices in your section to provide good RDM services?
97. Do you think your department has sufficient human resources to provide RDM services?
98. Does your department have RDM activity with regards to awareness and capacity building? E.g., for organising workshops, seminars, conferences?

SECTION E: Challenges facing the selected public universities in managing research data

99. What challenges of RDM do you encounter in your section?
100. What suggestions would you recommend in overcoming the challenges of RDM?

SECTION F: RDM framework for public universities in South Africa

101. In your opinion, what do you think can be done to develop a RDM framework for public universities in South Africa?

Thank you so much for your cooperation

Appendix I: Research Questionnaire for Researchers (NRF Rated)

University of Zululand

Faculty of Humanities and Social Science (HSS)

Department of Information Studies

Research Questionnaire

Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa

Dear répondent,

You are invited to participate in a research study titled ***"Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa".*** This study is being conducted by Mpilo Siphamandla Mthembu, à PhD student in the Department of Information Studies (DIS) at the University of Zululand. It is noted that such an investigation will be used widely for research, teaching and learning, inform RDM strategy and policy in South Africa and perhaps elsewhere.

As part of the requirements for the fulfilment of the Doctor of Philosophy (Information Science) degree**,** I kindly ask for your participation in my research project by completing this questionnaire. Your participation in this study is voluntary and you are free to withdraw your participation from this study at any time. There are no risks associated with participating in this study. Please be assured that this survey collects no identifying information of any respondent. All of the responses in the survey questionnaire will be recorded anonymously. By completing and submitting this survey questionnaire, you are indicating your consent to participate in the study. Your participation is appreciated and looking forward to sharing the results of the study with you.

For any queries pertaining this research, please do not hesitate to contact us:

7. **Mpilo Siphamandla Mthembu**, Doctoral candidate, Department of Information Studies, University of Zululand
   Email: mpilomthembu55@gmail.com
   Cell: 071 170 2832

8. **Prof Dennis Ngong Ocholla**, Research supervisor, Department of Information Studies, University of Zululand
   Email: OchollaD@unizulu.ac.za
   Cell: 082 372 4638

Thank you

Instructions

This document is divided into six (6) sections.

SECTION A: Biographical data
SECTION B: RDM practices at the selected public universities
SECTION C: Competency requirements of researchers and staff responsible for research data management in selected public universities.
SECTION D: RDM development strategies at the selected public universities.
SECTION E: Challenges facing the selected public universities in managing research data.
SECTION F: RDM framework for public universities in South Africa.

Please go through the questions and where relevant, mark your choice with an "X" or a "tick." (✓)

- Please note that some questions require a single response while others may require more than one response.
- The information you provide will be treated confidentially and only be used for the completion of the afore-mentioned qualification.

SECTION A: BIOGRAPHICAL DATA

1. What is your sex?

Male	
Female	

2. What is your population group?

African	
White	
Coloured	
Indian/ Asian	

3. What is your age category?

25-30 [ ]
35-40 [ ]
45-50 [ ]
55-60 [ ]

60-above [ ]

4. Please mark the name of your institution

University of Zululand (UZ)	
University of KwaZulu-Natal (UKZN)	
Durban University of Technology (DUT)	

5. What is your highest academic qualification?

Diploma	
Higher Diploma	
Bachelor's Degree	
Master's Degree	
PhD	
Other, please specify ……..	

SECTION B: RDM PRACTICES

6. Are you aware of the term 'Research Data management' (RDM)?
   a) Yes ☐
   b) No ☐

7. Have you submitted any RDM plan for your research activities?
   c) Yes ☐
   d) No ☐

8. Please indicate the type(s) of research data that you have generated as you undertake any research activities. Please select all applicable.

Spread sheets	
Images	
Documents (text, MS Word, PowerPoint, PDF, etc)	
Verbal context/ Audios	
Observations	
Dairies/ notebooks	
Questionnaires	
Databases	
Websites	
Scanned documents	
Musical compositions	
Video recording	
Other, please specify ……	

9. Approximate the amount of research data you have generated from your past research activities.

1-500 Kilobytes (KB)	
1-500 Megabytes (MB)	
1-100 gigabytes (GB)	
1 Terabyte (TB)	
I don't know	

10. Where do you store your research data/work? Please select all applicable

Laptop	
External storage devices (e.g., USB, external hard drive, CDs, etc)	
Emails	
Clouds (e.g., google drive)	
Personal website	
Institutional repository	
Other, please specify ……	

11. Do you back up your research data?
    a) Yes ☐
    b) No ☐

12. How often do you back up your research data?

Hourly	
Daily	
Weekly	
Monthly	
Quarterly	
Annually	
Never	

13. If you back up your research data, where do you back it up?

Laptop	
External storage devices (e.g., USB, external hard drive, CDs, etc)	
Emails	
Clouds (e.g., google drive)	
Personal website	
Sever unit	
Institutional repository	
Other, please specify ……	

14. Other than yourself managing and backing up your research data, who is responsible for its management and backup in your institution?

Librarians	
IT specialists	
My department	

Research office	
Research supervisor	
External research partners	
Third-party data centre	
Not sure	
Other, please specify ……..	

15. For how long does your institution keep research data?

Less than a year	
1 year	
1-3 years	
3-5 years	
5-10 years	
More than 10 years	
I don't know	

16. Which RDM related activities provided by your institution from the following are you aware of? Please select all applicable

Support for collaborative research	
Grant applications	
Guidance on RDM	
Support for data analysis and computational sciences	
Support for storage and high-speed connectivity	
Data management planning tool	
Data storage	
Data preservation	
Data curation	
Data use and reuse	
Data mining	
Other, please specify ……	

17. Which research data services offered by the library are you using or have used?

Directly participating with Librarians and IT specialist on RDM	
Providing reference and advisory support to researchers on RDM	
Creating web guides and finding aids for data repositories	
Providing technical support for research data service systems (repository, web portals, creating metadata, RDM system, access and discovery systems)	
Provision of RDM literacy programs	
Coordinating RDM advocacy programs between the researchers and the stakeholders	
Organising students' workshops on RDM related activities	
Organising workshop, seminars and direct training on RDM	

Data citation	
Data sharing and access	
Data storage and backup planning	

18. Have you shared your research data before?
    a) Yes ☐
    b) No ☐

19. If yes, which platform did you use?

Institutional repository	
Academic journals	
Conference proceedings	
Conference presentations	
Book chapters	
Email	
Hard copy	
External storage devices such as USB	
Media	
Other, please specify ……	

20. Do you think data sharing is important?
    a) Yes ☐
    b) No ☐

21. Motivations for data sharing

Promote innovations and potential new data uses	
Encourages scientific enquiry and debate	
Reduce the cost of duplicating data collection	
Enables scrutiny of research findings	
Increases the impact and visibility of research	
Leads to new collaborations between data users and data creators	
Other, please specify ……	

SECTION C: COMPETENCY REQUIREMENTS

22. Which types of knowledge do you possess in relation to RDM? Please, select all that apply.

Databases and data formats	
Data management and organization	
Data conversion and interoperability	
Cultures of practice	
Discovery and acquisition of data	
Quality assurance	
Data curation	

Data preservation	
Data sharing	
Database development	
Ethics, including citation of data	
Other, please specify ……	

23. From identified knowledge, where do you feel there is knowledge gap in terms of the RDM?

Databases and data formats [ ]
Data management and organization [ ]
Data conversion and interoperability [ ]
Cultures of practice [ ]
Discovery and acquisition of data [ ]
Quality assurance [ ]
Data curation [ ]
Data preservation [ ]
Data sharing [ ]
Database development [ ]
Ethics, including citation of data [ ]

24. Which types of skills do you possess in relation to RDM? Please, select all that apply.

Metadata skills	
Data curation and re-use skills	
Searching and retrieval skills	
Data preservation skills	
Data analysis skills	
Data capturing skills	
Sharing skills	
Database development skills	
Software skills	
Collaboration and communication skills	
Data visualization skills	
Other, please specify ……	

25. From identified skills, where do you feel there is skills gap in terms of the RDM?

Metadata skills [ ]
Data curation and re-use skills [ ]
Searching and retrieval skills [ ]
Data preservation skills [ ]
Data analysis skills [ ]
Data capturing skills [ ]
Sharing skills [ ]
Database development skills [ ]

Software skills [ ]
Collaboration and communication skills [ ]
Data visualization skills [ ]

26. Have you ever attended any research data management training in your institution?
    a) Yes ☐
    b) No ☐

27. If **yes**, who organised the training?

My university library	
My department	
My faculty	
By government agency	
Funding agency (e.g., NRF)	
Not sure	
Other, please specify …… ……..	

28. If **no**, which areas would you like receiving RDM training from the following? Please select all that apply.

Documenting data	
Formatting data	
Storing data	
Sharing data	
Archiving data	
Data curation	
Data mining	
Ethical issues	
Developing RDM plan	
Copyright issues	
Intellectual Property Rights	
Other, please specify ……	

29. Please rate the practices that could be used in the skills development of RDM in your department?

Attending workshops	Agree	Strongly agree	Moderate	Low
Attending conferences				
Provision of short courses on RDM				
Attending training provided by professional bodies				
Other, please specify				

……..				

SECTION D: RDM DEVELOPMENT STRATEGIES

30. Are there any policies in place to properly manage research data for a better understanding of its lifecycle, availability, accessibility, sharing, use, reuse and security in your institution?
    a) Yes ☐
    b) No ☐
    c) Not sure ☐

31. Do you think RDM policies and procedures are important in facilitating and promoting RDM services and practices in your institution?
    a) Agree ☐
    b) Strongly agree ☐
    c) Neutral ☐
    d) Not sure ☐

32. Is there any available RDM activity with regards to awareness and capacity building in your institution?
    a) Yes ☐
    b) No ☐
    c) Not sure ☐

33. If **yes**, what is that activity responsible for?

Arranging seminars	
Organising workshops	
Organising conferences	
Organising webinars	
Other, please specify ……	

34. Do you have your research data management plan for your current research activities?
    a) Yes ☐
    b) No ☐

35. What are the challenges in your institution that hinder your practices on managing your research data? Please select all that apply.

Lack of proper knowledge and skills	
Lack of awareness and understanding of the RDM activities and services provided by the institution	
lack of training workshops	
Lack of data resources/ infrastructure	
Lack of adequate policies and procedures	
Lack of support and guidance	

Lack of support from my department	
Lack of support from my faculty	
Lack of support from university library	
Shortage of staff responsible for RDM	
Lack of relevant policy or procedures	
Lack of human resource with adequate skills	
Lack of collaboration between the university and other stakeholders	
No capacity building activities in place to support RDM	

36. Except the identified challenges in question 35, what are other challenges of RDM do you encounter?
    .....
    .....
    .....
    .....
37. What suggestions would you recommend in overcoming the challenges of RDM?
    .....
    .....
    .....
    .....
    .....

SECTION F: RDM FRAMEWORK FOR PUBLIC UNIVERSITIES IN SOUTH AFRICA

38. In your opinion, what do you think can be done to develop a RDM framework for public universities in South Africa?
    .....
    .....
    .....
    .....

Thank you for your time and participating in this research

Appendix J: Consent Form

University of Zululand

Faculty of Humanities and Social Science (HSS)

Department of Information Studies

Consent Form

Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa

Dear respondent,

You have been invited to take part in a research project called "**Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa.**" The study is being conducted by Mpilo Siphamandla Mthembu, a PhD student in the Department of Information Studies at the University of Zululand. The results of the investigation will be widely used for research, teaching, and learning purposes, as well as to inform RDM strategy and policy in South Africa and beyond.

As a requirement for my Doctor of Philosophy (Information Science) degree, I am requesting your participation in this research project. Your participation is voluntary, and you are free to withdraw at any time without penalty. There are no risks associated with participating in this study, and all responses will be recorded anonymously to ensure confidentiality. I appreciate your participation and look forward to sharing the study's results with you.For any queries pertaining this research, please do not hesitate to contact us:

1. **Mpilo Siphamandla Mthembu**, Doctoral candidate, Department of Information Studies, University of Zululand
   Email: mpilomthembu55@gmail.com
   Cell: 071 170 2832

2. **Prof Dennis Ngong Ocholla**, Research supervisor, Department of Information Studies, University of Zululand

Email: OchollaD@unizulu.ac.za
Cell: 082 372 4638

CONSENT

I voluntarily agree to participate in the research project by (……) on the study entitled: **"Research Data Management services in the selected public universities in KwaZulu-Natal, South Africa"**.

Yes ☐

No ☐

Signature of participant …….

Date …….