DEVELOPING A FRAMEWORK FOR RESEARCH DATA MANAGEMENT CAPABILITIES AT THE UNIVERSITY OF NAMIBIA

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS IN INFORMATION SCIENCE

OF

THE UNIVERSITY OF NAMIBIA

BY

RONALD NGUATUMO KANGUTI

200746065

APRIL 2023

SUPERVISOR: PROF. TRYWELL KALUSOPA (UNIVERSITY OF NAMIBIA)

ABSTRACT

This study sought to investigate the University of Namibia's (UNAM) Research Data Management (RDM) capabilities with a view to developing an RDM framework for the institution. The study was informed by a pragmatic research paradigm and it adopted an explanatory sequential mixed method research design. In line with this research design, the study deployed a two-phased data collection process. First, during the quantitative phase, a bibliometric analysis on UNAM's research productivity, visibility and impact was conducted using Scopus (SciVal) and the Web of Science. Second, in the qualitative phase, 22 UNAM top researchers as derived from Scopus (SciVal) and the Web of Science were purposefully interviewed based on their relevance to the study. The population of the study consisted of nine hundred and thirty-four (934) researchers. The sample comprised of 40 UNAM top researchers as indexed in Scopus (Scival) and the Web of Science. The study found that RDM is a new concept at UNAM and that it is not integrated in the research processes. It was also discovered that although the concept is new, some researchers practice RDM services using their own means. The study further revealed that there is a lack of institutional support to enhance the management of research data. The study concluded that RDM do exist at UNAM even though different researchers understands and practices the concept differently. The study recommends that UNAM develops and implements the RDM framework as part of the organisation's structure to support staff members. The study proposes that UNAM undertake an RDM pilot study with leading universities in the region, and document some lessons learnt that can advance the implementation of RDM at UNAM.

Keywords: Bibliometric analysis, Research Data Management (RDM), RDM capabilities, Scopus (SciVal), Web of Science

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

CRP	Centre for Research and Publications
DMP	Data Management Plan
DIRISA	Data Intensive Research Infrastructure for South Africa
DOI	Digital Object Identifier
ICT	Information Communication Technology
MRC	Multidisciplinary Research Centre
OA	Open Access
RDM	Research Data Management
RIMS	Research Information Management Systems
SADA	South Africa Data Archive
SADC	Southern African Development Community
UREC	UNAM Research and Ethics Committee
UK	United Kingdom
USA	United States of America
UCT	University of Cape Town
UNAM	University of Namibia
UP	University of Pretoria

ACKNOWLEDGEMENTS

First and foremost, I am highly grateful to The Creator of all and my ancestors for giving me the strength to cross the line against all odds. I highly appreciate such never-ending blessings. In life, you cross paths with different people and amongst the many, there are those who really try to get the best out of you as an individual. To Prof. Trywell Kalusopa, my honoured supervisor, words will not be enough to appreciate your never-ending support, the seasoned words of encouragement and the wisdom. It was a great privilege and honour to work under your guidance. I take this journey as an introduction to academia.

I am also grateful to my fellow colleagues from the UNAM Library: Anna Leonard; had it not been for your credentials, data analysis in Web of Science was not going to be possible; Jacobina Mwiiyale for always reminding me that my Master's degree was long overdue even when I was in my first year; Astridah Samupwa for always pushing me to complete my studies; and the Technical Services team members and everyone in the UNAM Library not mentioned directly. I am truly thankful.

My acknowledgement would be incomplete without thanking Paulus (farm worker) for understanding that my frequent visit to the farm was no more possible and for ensuring that everything was under control.

GRATITUDE IS A MUST!!!

DEDICATION

I dedicate this thesis to my mother, Ms Hilaria Kanguti, my wife Caroline Kanguti and my joyful daughters Mbatjandangi and Uzuvira Kanguti. Words would not be enough to express my appreciation. Thank you for your words of encouragement and prayers.

DECLARATION

I, Ronald N. Kanguti hereby declare that this study is a true reflection of my own research and that this work, or part thereof, has not been submitted for a degree in any other institution of higher learning.

No part of this research maybe reproduced, stored in any system or transmitted in any form or by any means (e.eg electronic, mechanical, photocopying, recording or otherwise without prior permission from the author or the University of Namibia.

I grant the University of Namibia the right to reproduce this research as a whole or in part, in any format, which the University of Namibia deem fit, for any person or institution requiring it for studies and research, provided that the University of Namibia shall waive this right if the whole research has been published in a manner satisfactory to the University

……..…… **APRIL 2023**

CHAPTER ONE

INTRODUCTION AND BACKGROUND OF THE STUDY

1.1 Introduction

This chapter begins by presenting the orientation of the study and underlining the importance of the research, while giving background information. Recent research data management (RDM) developments are reviewed in this chapter, while the research objectives and significance of the research are outlined. In addition, the statement of the problem is presented followed by the significance of the study. The chapter also provides enlightenment about the limitations and delimitations encountered during the process of data collection and a review of the findings. Finally, operational definitions and meanings of key concepts used are given.

1.2 Orientation of the study

The terms data curation, data stewardship, data governance, data archiving and data management are often used interchangeably to imply Research Data Management (RDM) practices (van Wyk, 2018, p. 138). RDM refers to the process that involves storing and curating data that is generated from research. Data is organised from the time of creation at the beginning of the research lifecycle, right through to the dissemination and archiving of the research output (Kennan et al., 2012, p. 62; Matlatse, 2016, p. 9; Whyte & Tedds, 2011, p. 1).

In addition, Bester (2018, p. 44) proffers that RDM is a section of the research process which aims at making the research process efficient in order to meet the expectations and requirements of the university, research funders and legislation. RDM as a concept consists of a number of several activities that are linked to the data lifecycle, involving the design

and creation of data, storage, security, preservation, retrieval, sharing and re-use. The different processes take into account the technical capabilities, ethical considerations, legal issues and governance frameworks (Bester, 2018, p. 44).

The RDM concept in institutions is concerned with how one creates data and plans to utilise it, organises, structures and names data, keeps it secure, provides access, stores and backs it up, finds information resources and shares it with collaborators.

For the purpose of this study, RDM implies:

the process of controlling and organising the data generated during a research project, and covers the entire data lifecycle, which includes the planning of the investigation, conducting the investigation, storage and backing up of the data as it is created, preserving the data long-term, after the research investigation has concluded, and making the data accessible for future use. (van Wyk, 2018, p. 149)

Advancements in technology have opened up the need for the management of research data across the globe. Ngéno and Mutula (2018, p. 29) assert that the availability of huge amounts of research data has transformed the methods in which research is carried out for societal benefit. Ray (2014) also emphasises that the sharing of research data has led to the return on large investments, advances in human knowledge, the promotion of economic development and reductions of costly data duplications.

Research funders are realising the need to make available to the public researches that were conducted using public funds (Cox & Pinfield, 2014; Matlatse, 2016). Additionally, Avuglah (2016) posits that even private funding organisations have become more interested in the management of research data as a valuable output. Many of the research funders have guidelines or requirements that require researchers to commit to RDM practices.

However, worth noting is that RDM on the African continent is still in its early stages, apart from South Africa which is by far the most significant example.

The National Research Foundation (NRF) has developed cyber-infrastructure initiatives to support RDM practices through data management policies and programmes (Avuglah, 2016, p. 2; Kahn et al., 2014, p. 296).

Neuroth et al. (2013) affirm that the implementation of RDM practices across the globe has been in line with scholarly communication. The scholarly community necessitates reliable and long-term access to research data to ensure that data is available when needed and that it can be verified, especially data upon which current scholarly publications are based. With regards to many academic institutions in the United Kingdom (UK), the United States of America (USA) and Australia, strategies for RDM are initiated at the grant application stage (Neuroth, et al., 2013).

Tripathi and Pandy (2018, p. 105) stress that every institution of higher learning should develop a policy that can mandate a culture of sharing research data among faculties and researchers. Such policies should clearly define different practices such as data submissions, metadata entry, data clarifications, copyright agreements and conditions under which the data withdrawal requests if any can be accommodated, the conditions of use, the protection of confidentiality of sensitive data as well as the protection of data statements and data license for use. Patel (2016, p. 11) stresses that the formulation of a data management plan is an essential process of RDM. The data management plan entails the formulation of a detailed document outlining what a researcher will do with their research data during and after the completion of their studies. Data management plans encompass detailed information on the description of data, standards for formats and metadata, short-term and long-term storage plans, legal and ethical issues and access policies (Patterson,

2016, p. 11). Furthermore, Erway (2013, p. 5) specifies that data management policies are critical to identifying various university stakeholders in the RDM process. Responsible data management policies are not only restricted to the management of research data while the research project is active and storing the data afterwards, but it is about the institutional rationale for managing research data and the ensuing implications for the university.

Likewise, the RDM legal, policy and regulation framework should be able to address diverse drivers such as storage, security, preservation, quality, compliance, sharing and jurisdiction in order to improve the management, sharing and reuse of research output (Higman & Pinfield, 2015; Ngéno & Mutula, 2016; Pinfield et al., 2014). Moreover, Erway (2013, p. 7) proffers that there are key stakeholders that can ensure that a university can succeed in the implementation of RDM practices. These stakeholders include the university management, research office, information technology department, researchers, academic units, and the library. The different stakeholders have different mandates and roles to play in the research process at a university to ensure the successful and smooth running of RDM practices.

Earlier studies at the University of Namibia (UNAM) have established that UNAM has not adopted RDM practices (Kalusopa et al., 2020; Samupwa, 2019). The studies indicate that the issue of RDM has, however, gained prominence in most African academic institutions in recent years and they are more pronounced in South African universities in the SADC region, with RDM policies emerging as a result (Patterton et al., 2018). At UNAM, RDM is slowly evolving and getting acknowledged under the UNAM Research Ethics Policy (Section I1 – I3) where there are highlights about RDM and storage (UNAM, 2013, p. 29).

1.3 Statement of the problem

A number of studies have been conducted on developing an RDM framework in Africa. Chigwada et al. (2017) undertook a study on RDM in research institutions in Zimbabwe. The study concluded that RDM is still a relatively new concept in research institutions in Zimbabwe as compared to other parts of the world. Chiware and Becker (2018) researched on RDM services in Southern Africa. The upshot of the study concluded that, to be able to respond to RDM needs, there is a need to align institutional goals and building collaborations with key institutional stakeholders to respond appropriately. Chawinga and Zinn (2020) researched on RDM at public university in Malawi. The outcome of the study concluded that, lack of an RDM policy hinders RDM initiatives at the university.

Whilst these studies and many others (Triasih et al., 2020; Verhaar et al., 2017; Rice et al., 2013) have given a perception on how institutions have implemented RDM; those that holistically assess developing a framework for RDM capabilities are thin and scarce in the literature reviewed, more particularly in Africa.

There is a paucity of studies on RDM in Namibian and specifically universities. Thus, recent studies on RDM at UNAM present several gaps that warrant further research. For example, whereas the study by Samupwa (2019) is a quantitative baseline focusing on the adoption of RDM at UNAM, it does not go into detail to examine the whole research process in all faculties and research centres. The study only proposes an outline of some RDM policy options without delving into depth on the clarity of implementation. In another study, Kalusopa et al. (2020) proposed such as an RDM policy and framework, only in passing. In fact, it is instructive to state that none of the two studies explained the detailed nature and complexion of the policy or framework at UNAM. Yet, clear and grounded institutional RDM frameworks are critical to the implementation of RDM as they can

stipulate the role of each of the stakeholders in the whole research process, underscoring prudent scholarly communication (Mushi et al., 2020, p. 3). The implication presented by lack of RDM policy at UNAM is the limiting factor in the success and implementation of RDM. It is in this context that the present study responded to the gap in literature identified in the previous studies by seeking to develop a framework that considers the whole research cycle. It is also envisaged that such an RDM framework would guide UNAM on the implementation of RDM considering all the different stakeholders.

1.4 Objectives of the study

The main objective of the study was to assess the RDM capabilities at UNAM with a view to developing an appropriate RDM framework.

The study was guided by the following objectives, namely to:

1.3.1 Ascertain the awareness and determine research data output created at UNAM;

1.3.2 Assess research data output processing and analysis within UNAM;

1.3.3 Analyse the preservation of research data at UNAM;

1.3.4 Establish how research data is accessible and the extent of data re-use at UNAM; and

1.3.5 Establish the challenges in research data management and develop an appropriate RDM framework at UNAM.

1.5 Significance of the study

This study sought to enlighten UNAM management about RDM and the importance of an RDM framework. It is expected that key stakeholders in the research process at UNAM such as researchers, be Pro-Vice Chancellor: Research, Innovation and Development, the Centre for Research and Publications and the UNAM Library, be able to assess and

recognise the existing RDM gaps and thereby ensure the effective management of research data management.

The study developed an RDM framework that takes into account the whole research process at UNAM. The RDM framework provides a detailed plan on how to implement RDM at UNAM and it specifies the role of each stakeholder.

1.6 Scope and limitations of the study

RDM has a number of critical stakeholders that ensure that it is successfully implemented within an organisation. In terms of the scope, the present study focused on UNAM's top five (5) researchers per each of the eight (8) faculties that are indexed in Scopus (SciVal) and Web of Science based on their research productivity, visibility and impact. The following were the limitations encountered during the study:

- The study was limited to UNAM academic researchers across all faculties who are indexed in Scopus (SciVal) and Web of Science based on their research productivity, visibility and impact;
- The study was limited to bibliometric analysis only;
- The study analysed UNAM's research productivity, visibility and impact in Scopus and Web of Science. However, UNAM does not subscribe to the Web of Science which made it difficult for the researcher to find ways to get access to Web of Science;
- Faculty responses varied as some faculties had little or no representation at all;
- During the time of data collection, COVID-19 was at its peak in Namibia and at UNAM. As such, some of the researchers were working from home and others were recovering from the deadly pandemic and

Despite the mentioned limitations, the researcher deployed strategies to aid the study. For example, UNAM does not subscribe to Web of Science, which prompted the researcher to carry out the analysis using credentials from a South African university.

1.7 Delimitations of the study

The study was only confined to the UNAM top researchers from all eight faculties who appeared in Scopus and Web of Science based on their research productivity, visibility and impact during the time of conducting the study. The study did not consider altmetric analysis of the UNAM researcher as captured in the UNAM Institutional Repository because repositories do not measure research productivity, visibility and impact.

1.8 Assumptions of the study

The study deployed a bibliometric analysis of UNAM's research productivity, visibility and impact using Scopus (SciVal) and the Web of Science. The bibliometric analysis helped to establish UNAM's top authors who are indexed in Scopus (SciVal) and the Web of Science based on their research outputs. The UNAM Institutional Repository (IR) was excluded because UNAM has not embraced RDM as part of the research process. The assumption was that since UNAM researchers that are indexed in Scopus and the Web of Science have published with high impact factor journals and internationally reputable publishers, they might have been exposed to RDM questions as part of their requirements. The assumptions were that the UNAM top researchers purposefully interviewed would respond to the interview questions on data management requirements in the qualitative second phase of the study.

1.9 Definition of keywords

This section outlines the key terms that were used in this study.

1.9.1 Open access

This is defined as the provision of free access to peer-reviewed, scholarly and research information to all. It envisages that the rights holder grants worldwide irrevocable right of access to copy, use, distribute, transmit, and make derivative works in any format for any lawful activities with proper attribution to the original author (Swan, 2012).

1.9.2 Research data

Research data can be defined as facts to be collected in order to fulfil the objectives of a specific research project and are collected through different means of scientific techniques and tools. These techniques may be quantitative, qualitative or mixed method, based on the nature of research at hand (Baral, 2017).

1.9.3 Research Data Management (RDM)

RDM is the organisation of data through its life cycle with an aim of ensuring reliable and accurate data. Additionally, Ray (2014) defines the concept as the collection, validation and preservation of data for re-distribution and reuse by other researchers.

1.9.4 RDM framework

Bester (2018) defines a framework as a skeleton outlining or supporting a particular approach to RDM in an organisation. An RDM framework provides a detailed recommendations plan on how an organisation should implement RDM and the needed infrastructure, the roles of stakeholders, formats and metadata description protocols (Bester, 2018).

1.10 Organisation of the thesis

The study is divided into six chapters as follows:

Chapter One: Introduction and background of the study - This is the introductory chapter for the study. The main purpose of this chapter is to give an overview of the background, as well as the conceptual and contextual information of the study. This chapter covers the background, research problem, research objectives, significance of the study, as well as the limitations and delimitation of the study.

Chapter Two: Theoretical framework and literature review - This chapter presents a detailed description of the theoretical framework that underpins the study. This chapter provides a detailed review of both theoretical and empirical literature from books, journal articles, online databases and other sources. The gaps in the literature are identified and the contribution of this study in addressing them adduced.

Chapter Three: Research methodology - This chapter presents in detail the research paradigm, research approach, research design, study population, sampling technique, sample size, data collection methods, data analysis, validity and reliability of the data collection instruments, and ethical considerations.

Chapter Four: Data analysis and presentation of findings - This chapter presents the findings of the study, guided by theory and research questions. Data collected and analysed are presented in this chapter.

Chapter Five: Discussion and interpretation of results - This chapter interprets and discusses the findings of the study using the theoretical lessons that underpinned the study and also extant literature. The originality and contribution of the study is provided.

Chapter Six: Summary, conclusion and recommendations - This chapter presents the summary, conclusions and recommendations of the study. Further areas for research are presented.

1.11 Chapter summary

This chapter presented the background to the study, the orientation of the study, statement of the problem, objectives of the study, significance of the study, limitations and delimitations of the study and the definitions of key concepts. The chapter explained how the study was limited and restricted to bibliometric analysis of UNAM's research productivity, visibility and impact in Scopus (SciVal) and the Web of Science. The next chapter discusses the theoretical framework that guided the thesis and the conceptual and empirical literature review.

CHAPTER TWO

THEORETICAL FRAMEWORK AND LITERATURE REVIEW

2.1 Introduction

This chapter discusses the theoretical framework and literature review. First, it presents the theoretical framework adopted and its relevance to the study. Secondly, the conceptual and empirical literature as obtained from available books, journals, the internet and content analysis are presented as per the objectives of the research. These are namely to:

- Ascertain the awareness and determine research data output created at UNAM;
- Assess research data output processing and analysis within UNAM;
- Analyse the preservation of research data at UNAM;
- Establish how research data is accessible and the extent of data re- use at UNAM; and
- Establish challenges in research data management and develop an appropriate RDM framework at UNAM.

2.2 Theoretical framework

Ocholla and Le Roux (2011) define a theoretical framework as an agenda, outline or theoretical construct of a research approach. The theoretical framework of an empirical study refers to the system of concepts, assumptions, expectations, beliefs and theories that inform the research (Schurink, 2009). The theoretical framework acts as a blueprint or guide for research based on an existing theory in a field of inquiry that reflects the hypothesis of a particular study (Grant & Osanloo, 2014).

In addition, the theoretical framework functions as a foundation from which a research is assembled and it helps the researcher not to deviate from the boundaries of a range of

accepted theories when doing their research (Fulton & Krainovich-Miller, 2010). This study used the UK Data Archive Lifecycle model to get an understanding on how researchers at UNAM treat their research data through the different stages of the data life cycle.

Research data flows through a data life cycle, and generally there are a number of several models in literature which are used in RDM studies such as DataOne's lifecycle, Data Documentation Initiative's Combined Lifecycle Model, DCC Curation Lifecycle Model and the UK Data Archive Lifecycle (van Wyk, 2018, p. 234).

2.2.1 DataOne Data Lifecycle

The DataOne lifecycle model proposes data lifecycle based on the following steps: plan, collect, assure, describe, preserve, discover, integrate and analyse (Wiggins et al., 2013). This model centres more on the processes with a lifecycle than the different stages and hence it is not appropriate for the current study.

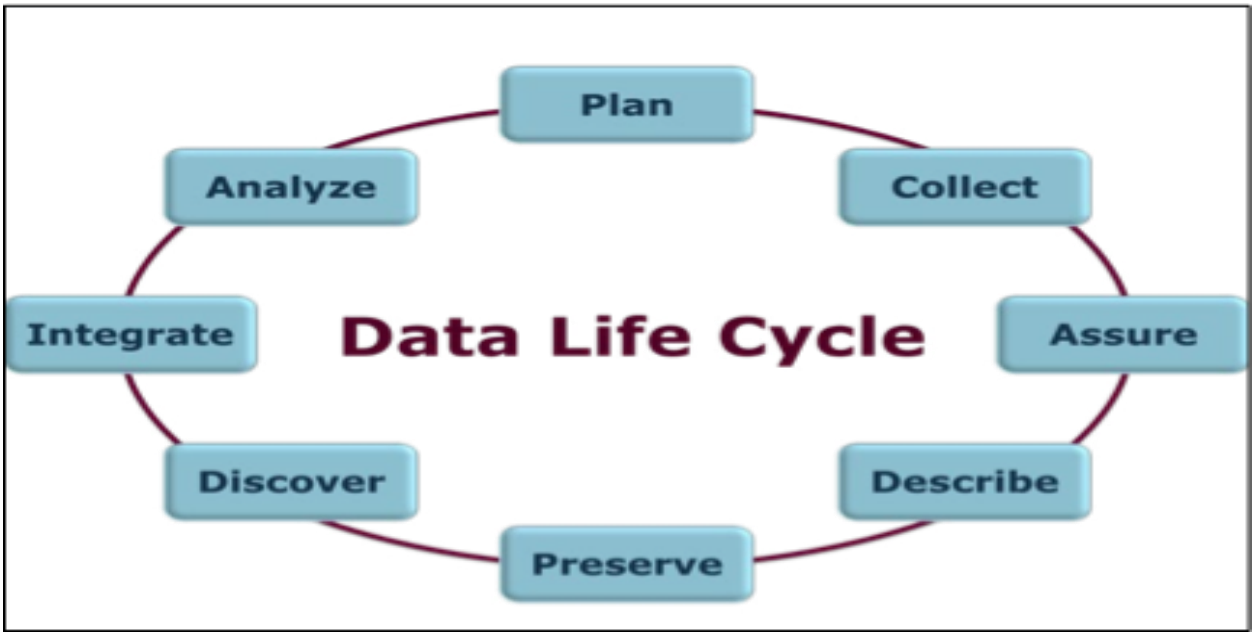


Figure 2.1: DataOne Data Lifecycle (Van Wyk, 2018)

2.2.2 Data Documentation Initiative's Combined Lifecycle Model

The Data Documentation Initiative's Combined Lifecycle Model comprises of the following sequential elements: study concept, data collection, data processing, data distribution, data discovery and data analysis, which feeds into data processing through repurposing, and an alternative path from data processing to data archiving, and then sequentially to data distribution (Ball, 2012, p. 7). This model is considered linear and hence it was regarded as not appropriate for the present study.

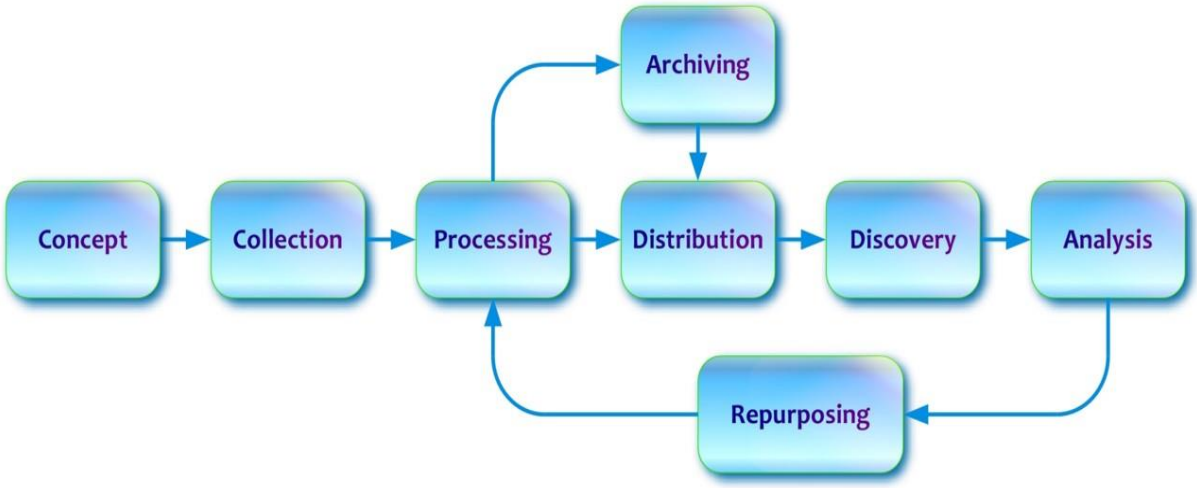


Figure 2.2 Data Documentation Initiative's Combined Lifecycle Model (Van Wyk, 2018)

2.2.3 DCC curation lifecycle model

This model includes sequential stages in its cycle, which include the following: conceptualise, create or receive, appraise and select, ingest, preservation action, store, access use and re-use, and transform. Intermittent activities include dispose, reappraise, and migrate. Although comprehensive in nature, the model does not cover all the stages of a research data cycle.

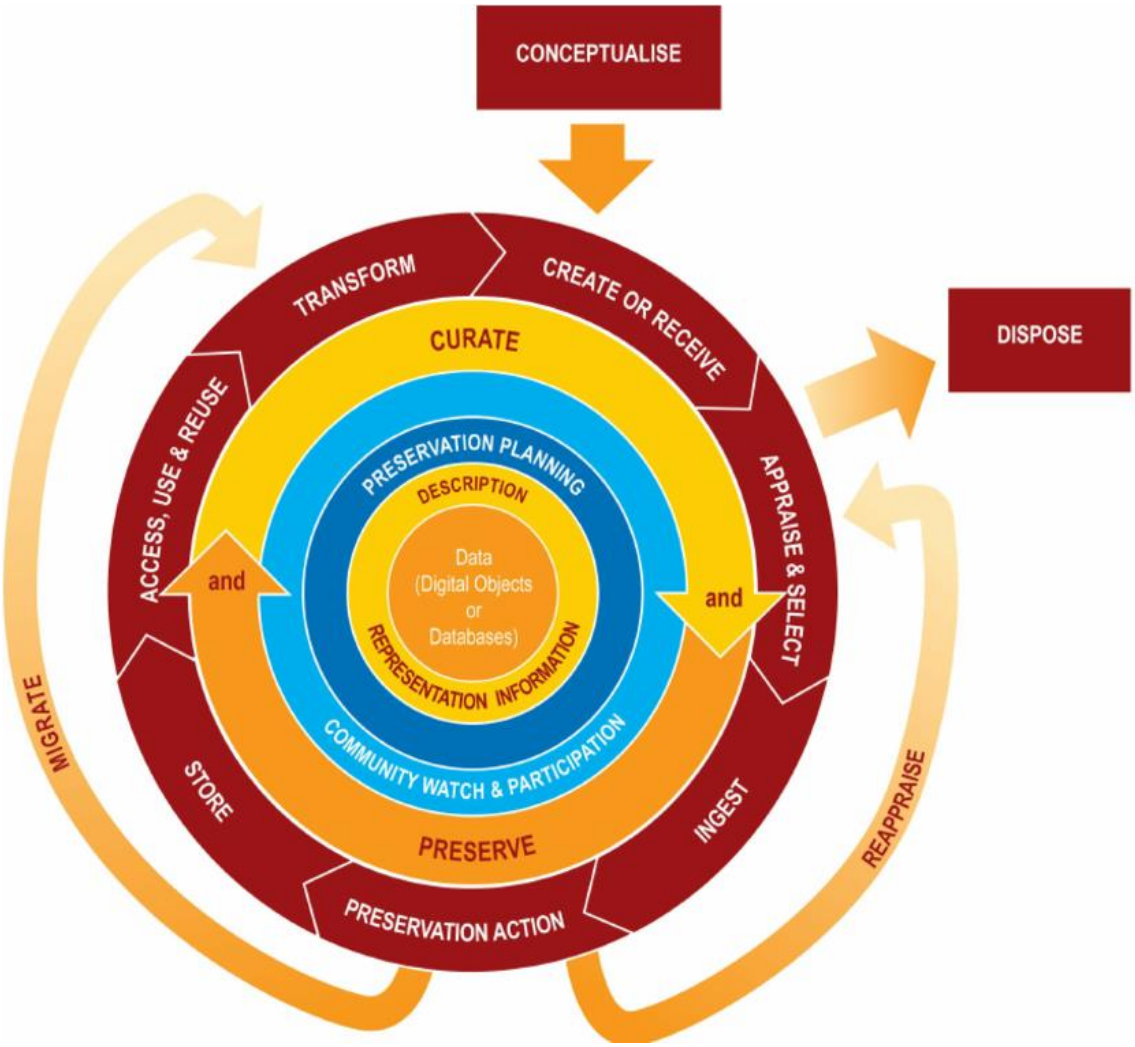


Figure 2.3: DDC Curation Lifecycle Model (Van Wyk, 2018)

2.3 UK data archive lifecycle model and its relevance to the study

The present study sought to examine the capabilities of RDM at UNAM looking at the full sequential stages from the point of creation. The UK Data Archive Lifecycle was found to be appropriate because many of the components in the other data lifecycles are integrated in this lifecycle. The UK Data Archive Lifecycle relates well with the research lifecycle, and also evidently displays discrete data management activities that can be engaged in each of the stages (van Wyk, 2018, p. 234).

The UK Data Archive Lifecycle selected for this study recommends a research data management lifecycle that comprises of sequential stages that begin at the data creation, processing data, analysing data, processing data, giving access to data and ending with re-using data (UK Data Archive, 2014).

Avuglah (2016, p. 14) proffers that the UK Data Archive Lifecycle model illustrates a sequential flow and connection between the different stages throughout the entire lifecycle.

The sequential phases in the UK Data Archive Lifecycle model provide a detailed model that can help UNAM researchers to consider how data management relates to the lifecycle of different research activities they carry out.

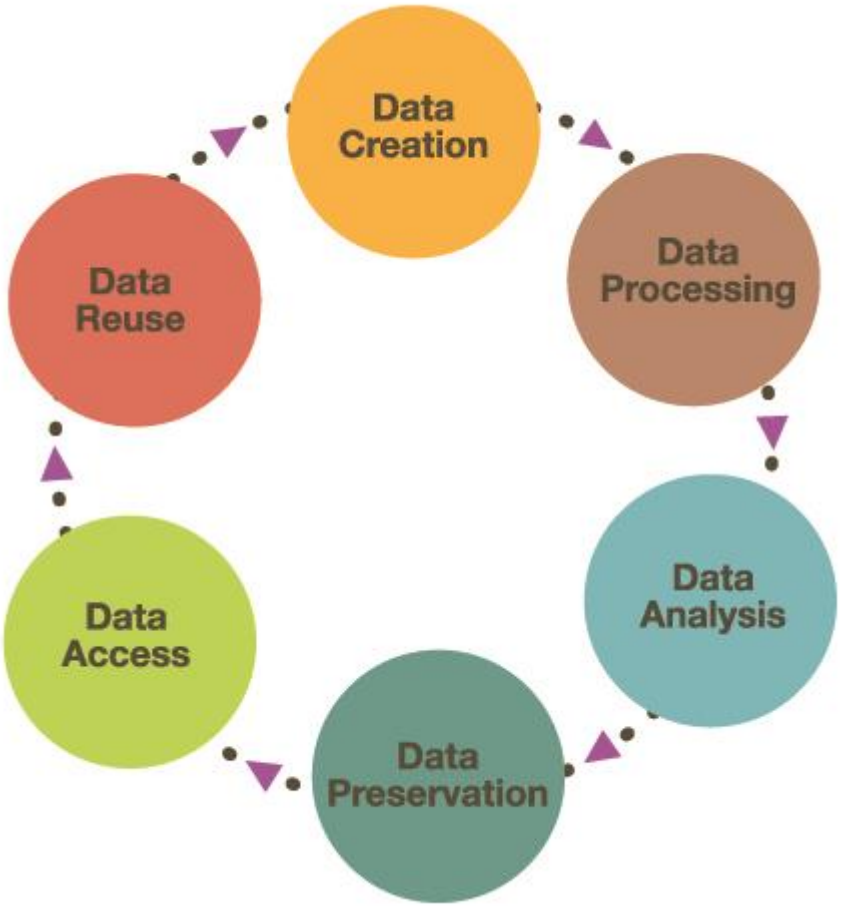


Figure 2.4: UK Data Archive Lifecycle model (Avughlah, 2016, p. 13)

2.4 Literature review

Literature review entails the process of searching for available literature that connects directly or indirectly to the research topic at hand. It provides a direction based on evidence written about the topic and guides how the topic will be investigated to fill certain gaps (du Plooy-Cilliers et al., 2014). The section below presents both a conceptual and empirical review of literature.

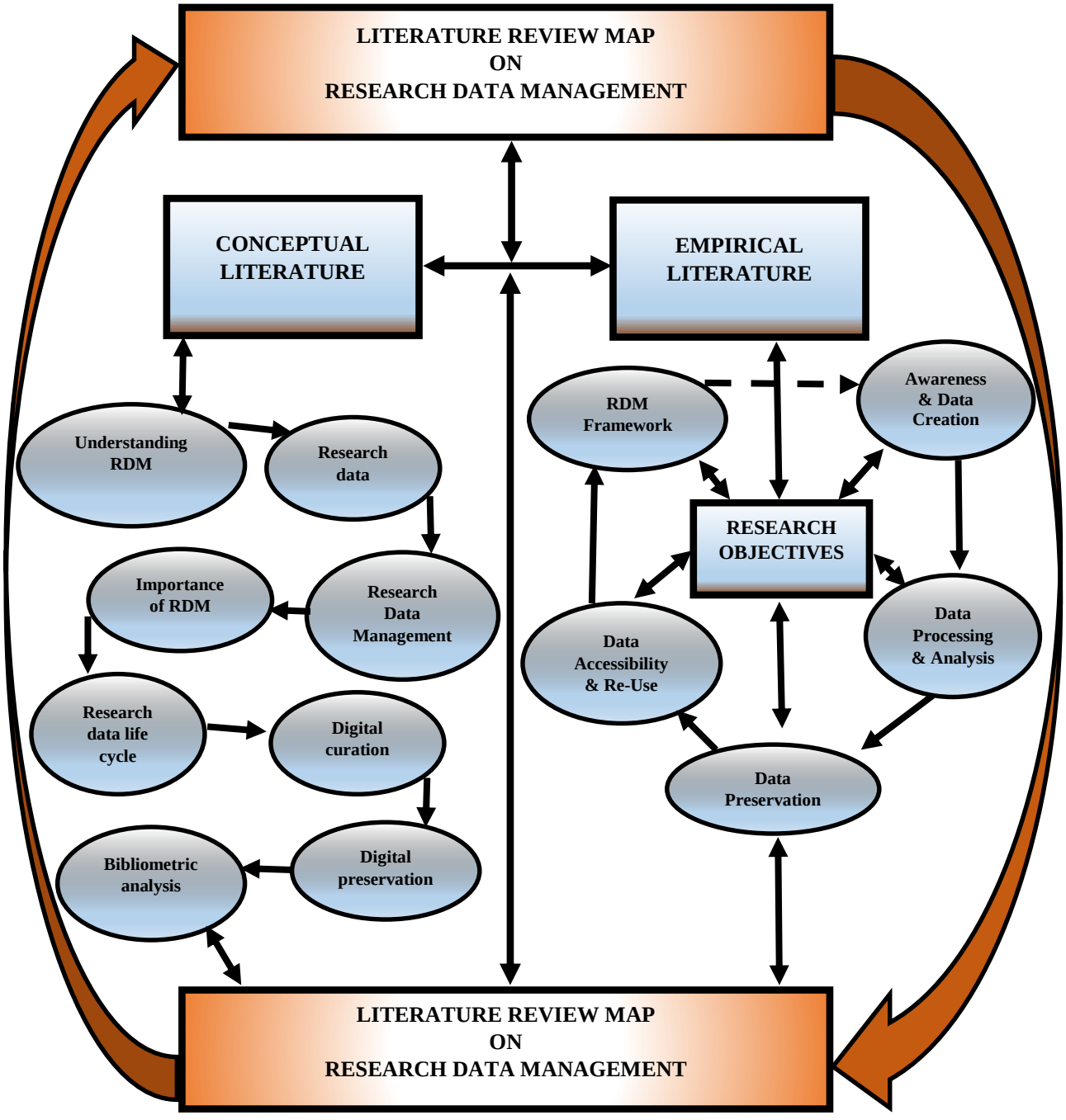


Figure 2.5: Literature review layout (Researcher, 2022)

2.4.1 Conceptual literature review

As the term implies, the section below presents and gives an elaborate understanding of the concept of RDM, which is the focal concept of the study.

2.4.1.1 Understanding RDM

On-going Open Access (OA) initiatives have contributed to the emergence of RDM across the globe. Research and academic institutions have been at the for-front of implementing RDM practices and policies (Winn, 2013. p. 3). Globally, many RDM and open access initiatives came forth discussing the importance of making public access to research data from public funded research (Johnsson & Ahlfeldt, 2015; Samupwa, 2019). Wessels et al. (2014) state that the Berlin Declaration on OA perceives OA to research data as having the potential to create a comprehensive source of human knowledge and cultural heritage approved by the scientific community.

The Berlin Declaration indicates that OA contributions include the following: original research results, raw data and metadata, source of materials, and scholarly multimedia material. Research data is essential to the foundation of scholarship due to the mere fact that it is irreproducible. Once lost, they are entirely gone and therefore will not be verifiable.

When the concept of research data is integrated in the new disciplinary context, it presents new opportunities that can answer old research questions and provide new ways from which new research questions can be generated (Neuroth et al., 2013). RDM has grown in importance as a result of funders' requirements of data management plans. There has been drastic policy shift from requiring open access to scientific publications to requiring open access to research data (Johnsson & Ahlfeldt, 2015, p. 14; Koopman & De Jager, 2016).

The OA movement encourages free access to research data produced through public funds in order to ensure that researchers will be able to re-use such data and prevent unnecessary costly duplications (Wessels et al., 2014, p. 52). In addition, OA to research data ensures that research results can be validated to guarantee quality control and reproducibility. Preserved and accessible research data is critical to helping policy makers to take key informed decisions (Wessels et al., 2014, p. 52). OA is a critical driver in National Open Science and Open Data Initiatives that inspire research data to be open. The growth in this concept is evident through emerging research data policies and RDM support systems in academic and research institutions (Chiware, 2018; Borglund & Engvail, 2014; McKiernan et al., 2016).

2.4.1.2 Research data

According to Baral (2017), research data can be regarded as facts to be collected in order to fulfil the objectives of a specific research project and are collected through different means of scientific techniques and tools. These techniques may be quantitative, qualitative or mixed method, based on the nature of research at hand. Research data means data that is presented in the form of facts such as observations, images, computer programmes results, recordings, and measurements or experiences on which an argument, theory, test or hypothesis is based (Baral, 2017). In addition, Van Wyk (2018, p. 142) defines research data as the recorded factual material that is commonly accepted in the scientific community as necessary to validate research findings.

Research data can be numerical, descriptive or visual, raw or analysed, experimental or observational. Bester (2018, p. 20) alludes that defining research data is somehow challenging since data by its nature is heterogeneous as research fields are diverse and use different data types.

Based on literature, the concept of research data has been defined as final data sets that verify and support research findings. Based on the definition above, the context in which the data was collected and processed is ignored (Bester, 2018, p. 20).

Avughlah (2016, p. 7) depicts that research data are factual materials that are acceptable in research to validate research findings. Although mostly created through digital platforms, all research data is included irrespective of format of creation. Moreover, Onyancha (2016) opines that research data should always be observed in reference to a particular subject discipline and it could refer to data from instruments like telescopes or raw data from a mass spectrometer, and to digital maps or full-text documents such as those used in the creation of critical editions. Research data is defined as the actual records, numerical scores, textual records, and images that have been used as primary sources for scientific research and have been accepted by the scientific community as necessary to validate research findings (Onyancha, 2014). According Krier and Strasser (2014), this is based on the following:

- Observational data – implies data that has been gathered through observing a particular population or phenomenon
- Experimental data – implies data that is derived from controlled and randomised experiments
- Computational data – implies data derived from the output of a computer that has taken a large set of varied data that is achieved through simulation.

2.4.1.3 Research Data Management (RDM)

Avuglah (2016, p. 11) indicated that the concept of RDM is still evolving within the research paradigm of e-research or e-science as it has attracted a lot of consideration by governments, funders, researchers and academic and research institutions.

Based on literature, there is no clear definition on RDM as a concept and it has been linked with other concepts such as digital curation, and digital preservation (Avuglah, 2016; Van Deventer & Pienaar, 2015; Wang, 2013; Whyte & Tedds, 2011).

The term RDM has been defined as a comprehensive concept that includes the data life cycle, appraisal, curation and the archiving of research data for long-term preservation, discovery and re-use, taking into consideration the technical capabilities, ethical considerations, legal issues and governance framework (Pienaar, 2011; Whyte & Tedds 2011).

Additionally, Matlatse (2016) defines RDM as a practice of storing and curating data, which is spawned from research. The process entails the storing of research data from the time of its creation at the beginning of the research lifecycle until the archiving of the research output. RDM comprises of a list of activities and processes that are actively engaged to ensure that research data is organised through the research cycle. The life span of research data overtakes the initial life span of the overall research project (Matlatse, 2016: Whyte & Tedds, 2011. p. 1).

Furthermore, the University of Leicester (2017) conceptualises that RDM is part of the research process and that it aims to ensure that the research process is efficient and is able to meet the expectations and requirements of the university, research funders, and legislation. The concept focuses on how one creates data and plans for its use; organises, structures and names data, keeps it secure, provides access, stores and backs it up, finds information resources and shares it with collaborators, and how it is published and gets cited. Sanjeeva (2018, p. 3) posits that RDM entails the organisation of data from the time it enters the research cycle through to the dissemination and preservation of valuable results and it consists of several activities and processes.

Recent developments on the management of research data emphasises on the preservation and access provision:

a) To ensure compliance with funding and regulatory requirements: Research Funding Institutions like NIH and NSF have made it mandatory to include Data Management Plans in proposals which are sent to them for Funding;
b) To ensure compliance with publishers' requirements – journal publishers increasingly require data that form the basis for publications to be shared or deposited in an accessible data centre or repository. This requirement applies to both commercially and publicly-funded research;
c) To ensure research integrity and validation of results. Accurate and complete research data are an essential part of the evidence that is necessary for evaluating and validating research results and for reconstructing the events and processes leading to them;
d) To increase research efficiency – good research data management will be helpful to organise files and data for access and analysis without difficulty;
e) Enhanced data security and minimised risk of data loss;
f) Wider dissemination and increased impact – research data, if correctly formatted, described and attributed, will have significant ongoing value and can continue to have impact long after the completion of a research project. Perhaps the most common reason to retain and manage research data is to facilitate online sharing. Initiatives such as DataCite, a registry assigning unique digital object identifiers (DOIs) to research data, helps to make data citable, traceable and findable, so that research data, as well as publications based on those data, form an important part of a researcher's output;
g) To ensure accountability – by managing research data and making it publicly available, the researcher will be able to demonstrate the responsible use of public resources to fund research;

h) To enable research continuity through secondary data use - good research data management will permit new and innovative research to be built on existing information. Sharing well-managed research data and enabling others to use it will also help to prevent the duplication of efforts;

i) Awareness about the advantages to use the data for replicability or further explorations is on the rise; and

j) The culture of "Open Data" is gaining importance and credibility in the Scientific Community (Sanjeeva, 2018).

2.4.1.4 Importance of RDM

Avuglah (2016, p. 16) specifies that the management of research data has importance and benefits attached for the researcher, scientific community, academic and research institutions and the broader public. The management of research data enables researchers to plan how they would go about planning data collection, documentation, processing, organising, managing and preserving data before commencing with a specific research project. Additionally, Strasser et al. (2012, p. 1) proffer that good RDM practices enable other collaborators to easily get access, understand and use data. This also helps with making data discoverable, accessible and intelligible for re-use in the long term.

Samupwa (2019, p. 29) asserts that the adoption and management of research data is alleged to be an obligation and the responsibility of each researcher as it is viewed as an important aspect of responsible conduct. It is the responsibility of each researcher to adopt RDM practices and ensure that they effectively manage research data that they create. Apart from the benefits that researchers are accomplishing from well managed RDM practices in the organisation, good RDM services serve as the foundation to quality in research and as such, integrity and becomes a part of the on-going research (Samupwa, 2019, p. 29).

Singh et al. (2018, p. 112) highlight that there are a number of benefits that are associated with the management of RDM in an organisation. Institutions that have adopted RDM provide their scholars with some of these benefits:

a) The preservation of research data for long-term use is different from storing data, and RDM practices ensure the authentication of data in trusted repositories for long-term curation;
b) Ensures the authenticity of data and is necessary to validate research findings without which research cannot happen and may not be completed;
c) Data can be applied in numerous other contexts by other researchers and may also be re-used by the same researcher to extend the use of such data;
d) It ensures economical means since it saves a lot of time for researchers and avails a chance to invest their time in other avenues instead of collecting data from scratch;
e) Enhances data sharing and brings transparency which improves the quality of research;
f) Provides security by keeping data safe and ensures that such data is not accessible to unauthorised parties and guarantees privacy;
g) Provides security from data hackers who may want to target and breach security; and
h) It enhances data citation and it is useful for training new researchers and also increases the efficient use of funding by promoting new research outputs.

What is coming out of the literature is that RDM is important in many ways especially for institutions of higher learning and research institutions. Research integrity and quality of research is reliant on efficient RDM practices.

2.4.1.5 Research data life cycle

Briney (2015) conceptualises that a research data lifecycle is mutual within the concept of data management as it helps with identifying the role that data plays at different phases of a research project. The data lifecycle clearly stipulates all the processes from the beginning to the end of the cycle. In addition, Wissik and Durco (2016, p. 95) highlight that research data lifecycles have gained importance in the researcher's scientific work by putting emphasis on data sharing. Data lifecycles are at the forefront of shaping the way we study digital processes and they present the whole research process in a detailed manner. Research data life cycles aid researchers and organisations by defining different phases that make it easier to identify each distinct phase of the research data. Research data is presented through different models such as the USGS Science Data Lifecycle Model, e-research and data and information model, University of Oxford Research Data Management Char, DCC Curation Lifecycle Model, and UK Data Archive Lifecycle among a series of data lifecycle models (Wissik & Durco, 2016, p. 96).

Pouchard (2015, p. 180) defines data lifecycle as the cycles that presents a structure for organising the different tasks and activities that are related to the management of data within a project or organisation. The data lifecycle model also presents the means of communicating different tasks to the intended audiences such as data managers, curators, repository specialists, and librarians, which results in a wide variety of data cycles with some focusing on the organisation, others on individual researchers and others on the community at large. Singh et al. (2018, p. 112) proffer that the data life cycle in RDM are used to give an overview to research so that they can better understand the meaning of RDM.

Data is generally an observable fact that is ***created*** on the basis of a research project and is presented as raw material and information. The data is obtained through the ***processing*** of

data that entails ***analysing*** to provide indicators and projections to arrive at valid inferences. The last three stages of the data life cycle involve the ***preservation*** of research data to be used as a basis for arriving at definitive conclusions and ***providing access*** to other researchers for ***re-use*** purposes. Emanating from the literature is that research data flows through a research cycle. The research cycle clearly defines the different stages in which data flows from the time of creation until the final stage. The type of research data cycle adopted by an organisation depends on the relevance of such data cycle in addressing the RDM needs of such an organisation or researchers.

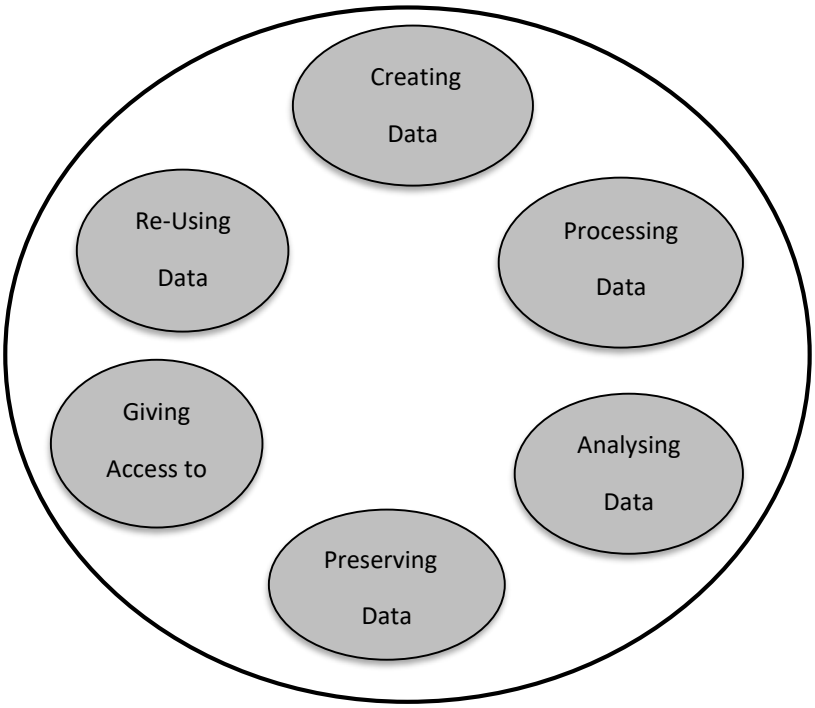


Figure 2.6: Research data life cycle (Avughlah, 2016)

2.4.1.6 RDM in Africa

Chiware and Mathe (2015, p. 1) state that the notion of RDM is being realised based on funding agencies and national governments that are gradually demanding that research grant recipients incorporate RDM practices as part of their research, with provisions for re-use. This is evident in countries such as the United States of America (USA) erecting

organisations such as The National Science Foundation, the Australian National Data service, and the e-Science Core Programme in the United Kingdom (UK). These national policies are used in order to guide, develop and enforce RDM practices at national and institutional levels.

Africa has experienced a shortage of RDM programmes, especially at institutions level as compared to other continents except for institutions of higher learning in South Africa.

RDM is perceived to be in the development stage in Africa as the majority of African countries are yet to adopt the concept at national level (Avuglah, 2016; Samupwa, 2019).

Chiware and Becker (2018, p. 3) indicate that there is progress that is being made in Southern Africa in relation to RDM as a result of the emergence of RDM institutional policies, institutional RDM working groups, institutional data centres and the development of open data platforms. In some Southern African academic institutions and research libraries, the RDM concept is driven by new institutional requirements on the better management of research data across disciplines and national mandates advocating for free access to research data generated through public funds.

In South Africa, the RDM concept has been recognised and adopted by the National Research Foundation (NRF) through its policy statement as well as the development of national infrastructure like the Data Intensive Research Infrastructure for South Africa (DIRISA) and the South Africa Data Archive (SADA) innovations (Avuglah, 2016, p. 22). The South African NRF has issued a firm statement on OA, indicating that research data generated from NRF funded research be deposited in an OA repository accompanied by a Digital Object Identifier (DOI) (NRF, 2015, online).

According to Chiware and Mathe (2015, p. 3), the NRF is responsible for maintaining several research support and knowledge networking databases that are contributing to

knowledge creation and the promotion of research development in South Africa. The databases are facilitating access to research results and technology and innovation. RDM innovations include the Current and Completed Research Projects Database; the National Electronic Theses and Dissertations Portal; the NRF Funded Projects Database; the Forthcoming Conferences Database; the South African Professional Associations Database; the Research Information Management Systems (RIMS); Data Resources in Africa; and the Digital Initiatives Register and Heritage Repository (Chiware & Mathe, 2015, p. 3). South African universities developed RDM related policies in order to align their research practices to the RDM concept. These include universities such as University of Pretoria (UP), Stellenbosch University, University of Cape Town (UCT), University of South Africa (UNISA), and University of the Witwatersrand among others in the country (Kahn et al., 2014, p. 297).

The reviewed literature has revealed that the appreciation of RDM as part of the research processes is yet to be fully integrated in many universities on the African continent except for South African case studies.

2.4.1.7 RDM at UNAM

According to Kalusopa et al. (2020, p. 497), UNAM has the responsibility, as an institution of higher learning, to conduct research, which emanates from its legal establishment by Act 18 of 1992. In order to accomplish this obligation, UNAM established two centres of excellence, the Multidisciplinary Research Centre (MRC) and the Sam Nujoma Marine and Coastal Resources Research Centre (SANUMARC), with full time employed researchers. The rationale behind the establishment of the two centres were primarily to conduct research and strengthen research collaboration nationally and internationally by engaging both the private and public sector. In addition to strengthening research at UNAM, academics are required to dedicate 30 percent of their work to research activities (UNAM,

2019). Moreover, in order to advance research, UNAM established a Centre for Research and Publications under the leadership of the office of the Pro-Vice-Chancellor: Research Innovation and Development. The mandate of the Centre for Research and Publications is to administer policies associated with research and also to support research within the institution through the procurement of relevant infrastructures and software required by researchers to effectively conduct research (UNAM, 2019). Furthermore, UNAM developed policies that guide and spearhead research at the institution. These policies comprise of the research and ethics policy, and the scholarly communication policy.

The UNAM research and ethics policy, regulations and guidelines have a highlight on RDM management and storage (UNAM, 2013, p. 29). However, besides the plans and formulation of some policies around research at UNAM, there are no clear directives regarding RDM at UNAM, as there is at present no policy that addresses and accentuates RDM practices. Studies addressing RDM have been carried out at UNAM. Kalusopa et al. (2020) conducted a study on research productivity, visibility and impact at UNAM.

The study recommended that UNAM needs to work on an RDM framework that can include designing a research data management policy, guidelines and procedures of the RDM process (Kalusopa et al., 2020, p. 511). Samupwa (2019) conducted a study on adopting research data management practices at the University of Namibia and the study recommended that UNAM develops policies to drive RDM.

The study further established that RDM at UNAM is not integrated into the research process and only researchers practice at an individual level and also refer to it as a concept that is important and yet it is not implemented at UNAM. Furthermore, the study underscored the fact that there is lack of direction at UNAM, guiding the implementation and integration of RDM as part of the UNAM research processes.

2.4.1.7 Digital curation

Mills (2013, p. 47) indicated that digital curation was the sifting and aggression of the internet and other digital resources into manageable collection of what is relevant in institutions of higher learning. In addition, Ungerer (2016, p. 6) posits that digital curation is an active process whereby content or artefacts are purposefully selected to be preserved for future access. In addition, Poole (2016) indicates that digital curation is known to be the umbrella term that includes data curation and digital preservation. The term comprises of vital counterweight activities to problems of data loss, and it adds value to trusted data assets for current and future usage (Poole, 2016).

Neuroth et al. (2013) specify that digital curation is about making research data digitally available for long-term and sometimes even as independent publications in their own right. The intention is to ensure that they are verifiable, interpretable and re-useable, and to cross-link research data using research infrastructure in order to increase the potential of inter-disciplinary re-use.

2.4.1.8 Digital preservation

According to Najar and Wani (2019), digital preservation incorporates activities that are undertaken by a digital curator to guarantee that digital content for which the digital curator has responsibility is maintained in useable formats that are usable over a long period of time and is able to be made available in meaningful ways for currently existing and future users beyond the limits of media failure or technological obsolescence (Najar & Wani, 2019). Digital preservation involves a number of activities that need to be considered to ensure that there is continued access to digital resources over time (Velmurugan, 2013). Additionally, Avuglah (2016) defines digital preservation as the activities needed to ensure accurate and consistent access to digital resources over a long period of time.

Moreover, Sadiku et al. (2017) indicate that digital preservation necessitates a process whereby there are clear policies that can guarantee the long-term preservation of digital recourses through preservation methods and technologies that can warrant digital longevity. The overall intention of digital preservation is to guarantee long-term availability and access to digital resources (Sidiku et a., 2017). Shimray and Ramaiah (2018, p. 2) refer to digital preservation as simply the preservation of rare and delicate materials and objects through digitisation by the use of computers or electronic equipment. Moreover, Qasim et al. (2018, p. 3) state that digital preservation is the active management of digital content over time to ensure ongoing access. It involves a series of activities that involve selecting content for preservation, preparing and maintaining it in an environment that enables access, and having strategies in place to ensure that this content can be made available over time.

2.4.1.9 Bibliometric analysis

Ellegaard and Wallin (2015) advances that research impact of scientific publications and journals have gained importance over the last decade. Consequently, bibliometric analysis methods have been adopted by institutions in order to provide an analysis of written publications. Methods of bibliometric analysis are playing a significant role in the ranking of research and academic institutions (Ellegaard & Wallin, 2015). In response to emerging trends, libraries across the globe responded by launching efforts to prove their worth and their impact through supporting research activities. In support of research, libraries have shifted their focus from traditional forms of support for researchers towards data management and curation and bibliometric citations (Corrall et al., p. 5).

Bibliometric analysis requires data sources such as Web of Science and Scopus that can be carried out based on an institution's external data. Web of Science which is produced by Clarivate analytics and Scopus which is produced by Elsevier are the two most commonly

used bibliometric data sources in research analysis (Waltman & Noyons, 2018, p. 6). The reviewed literature has revealed that studies on bibliometric analysis of universities' research productivity, visibility and impact are of critical importance as they give a clear outline on universities' research productivity on an annual basis, collaboration with other universities across the globe, and the impact of research produced and frequency of such research.

2.4.2 Empirical literature on RDM

The section below considers previous empirical literature and documents on RDM. The literature review which follows is discussed according to the five respective objectives of the study and they are thematically presented as follows:

a) Awareness and data creation
b) Data processing and analysis
c) Data preservation
d) Data accessibility and re-use
e) RDM framework

2.4.2.1 Awareness and data creation

One of the objectives of the study was to analyse the awareness and creation of research data at UNAM. Manu (2021) indicates that good RDM practices deal with planning, preservation, organisation, access, sharerability, description and publishing of research data. It allows researchers to have a better understanding of the data lifecycle.

Therefore, it is important that universities understand their researchers' perceptions and awareness of RDM. Researchers' awareness of RDM is critical to the management of research data as they are the role players in the collection, organising and sharing of such data.

The sharing of research data is largely dependent on their willingness and level of awareness (Rafig & Ameen, 2021, p. 2). Chigwada (2021) posits that researchers always present mixed feelings about having to share their research data, which is clearly linked to their level of awareness. Data protection from misuse by other researchers is one of the key areas of concern, which require good RDM practices as well as assisting researchers to draw good data management plans that are inclusive of use, reuse and preservation of such data (Chigwada, 2021, p. 424).

The creation of research data entails actions whereby researchers are involved in designing data management plans, collection, storage and metadata creation. The different processes help researchers to determine how they can deal with their data after completing their research (van Wyk, 2018). In addition, Kabir (2016) depicts that data creation is common in all fields of study and it is created through various approaches emanating from either quantitative or qualitative methods. However, data storage and back-ups are prepared in the creation stage (van Wyk, 2018, p. 240). Metadata is critical to serving in the description of data and ensuring access to data sets. The described data offers information on the aim, origins, time references, geographic location, creating author, access conditions and terms of use of data sets. Researchers are expected to be aware of institutional policies that guide data collection (Corti et al., 2014). This however, depicts a gap when it comes to RDM awareness in Namibia, especially at UNAM, where the concept is not fully integrated into the university's research processes.

2.4.2.2 Data processing and analysis

Data processing ensures that research data is cleansed and free of possible errors that would affect the quality of data. Research data should be able to make sense and sensitive data should be identified while at the same time preserving its format and data type (Rughunathan, 2013).

Shukla (2018) posits that data processing is done in order to check illegible, incomplete, illogical or inconsistent responses emanating from the collected data. The process helps researchers to deal with errors that may affect the reliability of the study. Data analysis on the other hand consigns significance to the gathered data and affirms the conclusions, significance, and implications of the findings (van Wyk, 2018, p. 242).

Additionally, data analysis publishes data sets, making research data to underpin research findings published in peer-reviewed articles available in the public domain as supplementary materials (Corti et al., 2014, p. 197). Sharma (2018) indicates that the analysis of data converts data into information and knowledge and explores the relationship between variables. It helps the researchers to appreciate the meaning of scientific methods which include testing of hypothesis and statistical significance in relation to the research questions. Several issues are critical to the researchers in respect to data analysis:

a) Having the necessary skills to analyse;
b) Distinguishing data types;
c) Distinguishing different types of statistical tests;
d) Identify the selection of a right test;
e) Determining statistical significance;
f) Distinguishing between parametric and non-parametric tests with their applying criteria;
g) Distinguishing between correlation and regression;
h) Drawing unbiased inferences;
i) Inappropriate subgroup analysis;
j) Lack of clearly defined and objective outcome measurements;
k) Partitioning ‘text’ when analysing qualitative data;
l) Reliability and validity; and
m) Extent of analysis.

2.4.2.3 Data preservation

One key objective of the study was to understand how researchers at UNAM deal with issues of data preservation. The term data preservation has always been confused with digital preservation. The term data preservation in RDM can be defined as the process dealing with the maintenance of research data and ensuring that such data can be retrieved, understood, accessed and used (Patterton, 2016). Data preservation entails ensuring that research data continues to be accessible over time even when the current technology becomes obsolete (Rahmanto & Riasetiawan, 2018). This however, gives an indication of the struggle faced by UNAM researchers when it comes to preserving their researchers' data using platforms that ensure long-term access.

Emerging technologies need to be monitored during this stage in order to ensure that there are migration plans in place and that no data shall be lost (O'Donoghue & van Hemert, 2009). Data is preserved and made available so that anyone rather than the original creator is able to access, use and interpret the data. This process ensures that research data is retrievable and accessible in the long-term future (van Wyk, 2018, p. 244). The preservation of research data connects data to specific research outputs through using a Digital Object Identifier (DOI) (Callaghan et al., 2013). In contrast, digital preservation refers to the preservation of materials that are born digital and never existed in print related formats. The process includes a set of processes and activities that ensure continuous access to digital content over a period of time (Pal, 2012). Velmuragun (2013) posits that digital preservation takes into account a number of series and actions taken to ensure the long-term management of digital materials for as long as necessary. It combines policies, strategies and actions (Velmuragum, 2013, p. 2).

Furthermore, Nilesh and Verma (2012) narrate that digital preservation is driven by maintaining the ability to display, retrieve and use digital collections in the face of evolving

technologies, organisational infrastructures and elements. These processes clearly specify the importance of data preservation by researchers in order to continue having access to their research data through available technologies and platforms.

2.4.2.4 Giving access to data

The production of born digital data means data-led science and open access to research data. When data and research results are shared, this enables other researchers to check the accuracy of the data, verify analyses and conclusions and build on previous work.

Developments in e-research through the use of digital technologies to support new and existing research has fostered a re-reconsideration of strategies and ways in which scientific and scholarly knowledge is produced and shared (Wessels et al., 2014, p. 49). Smith et al. (2021) posit that ensuring that research data are more accessible fuels scientific analysis and discovery by making data more open to scrutiny, re-analysis and extension. Open scholarship and open access to research data accelerates the pace of discovery and its application to societal problems, as well heighten the visibility and reputation of an institution and its scholars (Smith, et al., 2021, p. 4).

This process entails ensuring that research data is accessible to other researches through different formats based on outlined access policies and guidelines. The availing of research data presents an opportunity for researchers to validate research outputs (Corti et al., 2014, p. 2). Additionally, data sharing is significant to the improvement and augmentation of scientific inquiry through a variety of other activities such as testing theories and hypotheses and methodologies used for particular studies (van Wyk, 2018, p. 245).

Furthermore, data confidentiality is assigned through restrictions, which may be linked to legal, health and or commercial reasons. Jones (2011) posits that giving access to research data helps with reassuring funders by outlining when, how and where data will be made

available. Publishers and research funders may require that you share your data, so it is worth investing time to plan for effective data management. Several funders ask for data plans as part of grant proposals. Based on the reviewed literature, the available research clearly indicates that giving access to research data.

2.4.2.5 Research data re-use

Data re-use has gained significance in academia thus providing new impetus for academic research. The notion of data re-use has re-shaped a new scholarly debate in the context of data-driven research paradigm that has attracted attention in academia worldwide (Wang et al., 2021, p. 1161). In addition, van de Sandt et al. (2019) posit that the ability to re-use research data has become critical to the benefit of the wider research community. Many researchers across disciplines are faced with demands to share their research data so that it can be re-used. Such demands on data use and re-use present implications on how we document, publish and share research. Furthermore, Pasquetto et al. (2017) narrate that arguments for data sharing are to reproduce research to make public assets available to the public as a way to leverage investments in research and to advance innovation.

This process involves the technique whereby captured data can be re-analysed, re-worked or re-used for new analyses that can be associated with contemporary data. This process is essential to situations where researchers find it difficult or expensive to collect research data (Corti et al., 2014, p. 169). In addition, data re-use allows for data to be cited in a similar manner with traditional sources of information (Corti et al., 2014, p. 197).

Pryor (2012) also indicates that published research may be accessed and used by other scholars. The scholars can use the published data to plan a new research project or simply to compare such research data with their own findings.

Furthermore, Matlatse (2016) indicates that data re-use is one of the critical aspects of managing research data.

The motivation for managing research data is the fact that research is enhanced and validated by the availability of the data behind it. Re-used and shared data should be accompanied by explanatory documentation of its production, processing, analysis and workflows.

Data re-use entails the use of data by one or more community by using data collected for one purpose to study a different new problem. It allows the reanalysis of evidence, the reproduction and verification results, minimising duplication of effort, and building on the work of others (Thanos, 2017, p 1). Based on the reviewed literature, research data re-use has many derived benefits if integrated in the research process because researchers would not always have to collect data, but analyse existing data to address their information needs. Although sensitive in nature based on concepts of academic freedom, this practice is yet to be integrated into research practices at UNAM.

2.5 RDM framework

The final objective of the study was to develop an RDM framework. RDM responsibility lies with the researcher and the institution. The first requirement is to ensure that RDM is covered under the policy framework of the institution. It should provide a detailed structure that is geared towards improving the researcher's management practices and implanting new services (Singh et al., 2018, p. 113).

Erway (2013) indicates that RDM policies are important in identifying key stakeholders in the research process and outlining different key elements pertaining to data management from the time it is created until the end of the cycle.

RDM policies are critical to ensuring that research data created will be managed appropriately throughout its entire lifecycle and will be able to be discoverable and made available to allow re-use (Erway, 2013).

RDM frameworks are fundamental to mandating data sharing in an institution by spelling out the purpose, scope, applicability and guidelines to the data contributors relating to data submission, licensing, metadata entry, data classification, copyright agreements and conditions under which the data withdrawal requests, if any, will be considered. This also includes the terms and conditions of the use of data, the protection of confidentiality of sensitive data, protection of data against security breaches, and Intellectual Property (IP) concerns (Patel, 2016, p. 228). Literature indicates that RDM in an organisational context can only be guided by a policy or framework. It showcases a gap existing at UNAM whereby there is no RDM policy, hence the current challenges about RDM at UNAM.

2.6 Chapter summary

This chapter aimed at presenting the theoretical framework and reviewed pertinent literature that guided the current study. The theoretical framework revealed that there are a handful of research data models. The reviewed literature in this chapter affirmed that RDM studies have been carried out in the SADC region, particularly in South Africa. However, very few studies exist in Namibia and specifically about UNAM. Through the review of the literature, it was established that there was a dearth of empirical literature, which investigated and outlined the different phases of the UK data archive lifecycle model. The next chapter discusses the methodology of the study.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction

The previous chapter reviewed literature on the study of RDM. The review of literature was guided by the study objectives with a focus on RDM awareness and data creation: data processing and analysis; data preservation; data accessibility and re-use, and the RDM framework. The purpose of the present chapter is to present the research methodology used in the study. Luthuli (2021, p. 63) posits that it is important to elaborate on the research design and methodology. This implies that during the planning and execution stages of a research project, the research methodology discourses the researcher's ideas about what to investigate, which methodologies to use and the appropriate data gathering and analysis tools to rely on.

The research approach helps the researcher to select the most appropriate approach based on the research problem, the study questions, experience, resources and the target audience (Creswell & Creswell, 2018). In addition, Ngulube (2015, p. 125) posits that knowledge produced in any given scientific field depends primarily on the methodology adopted for that particular study. This chapter therefore, describes the current study's research methodology.

Approaches, methods and instruments used in conducting this study are all discussed in detail in Figure 3.1. Additionally, the chapter focuses on the research methods used, the study population and justification, ethical matters and the data collection instruments. The illustration below depicts a detailed research methodology road map that guided the current study.

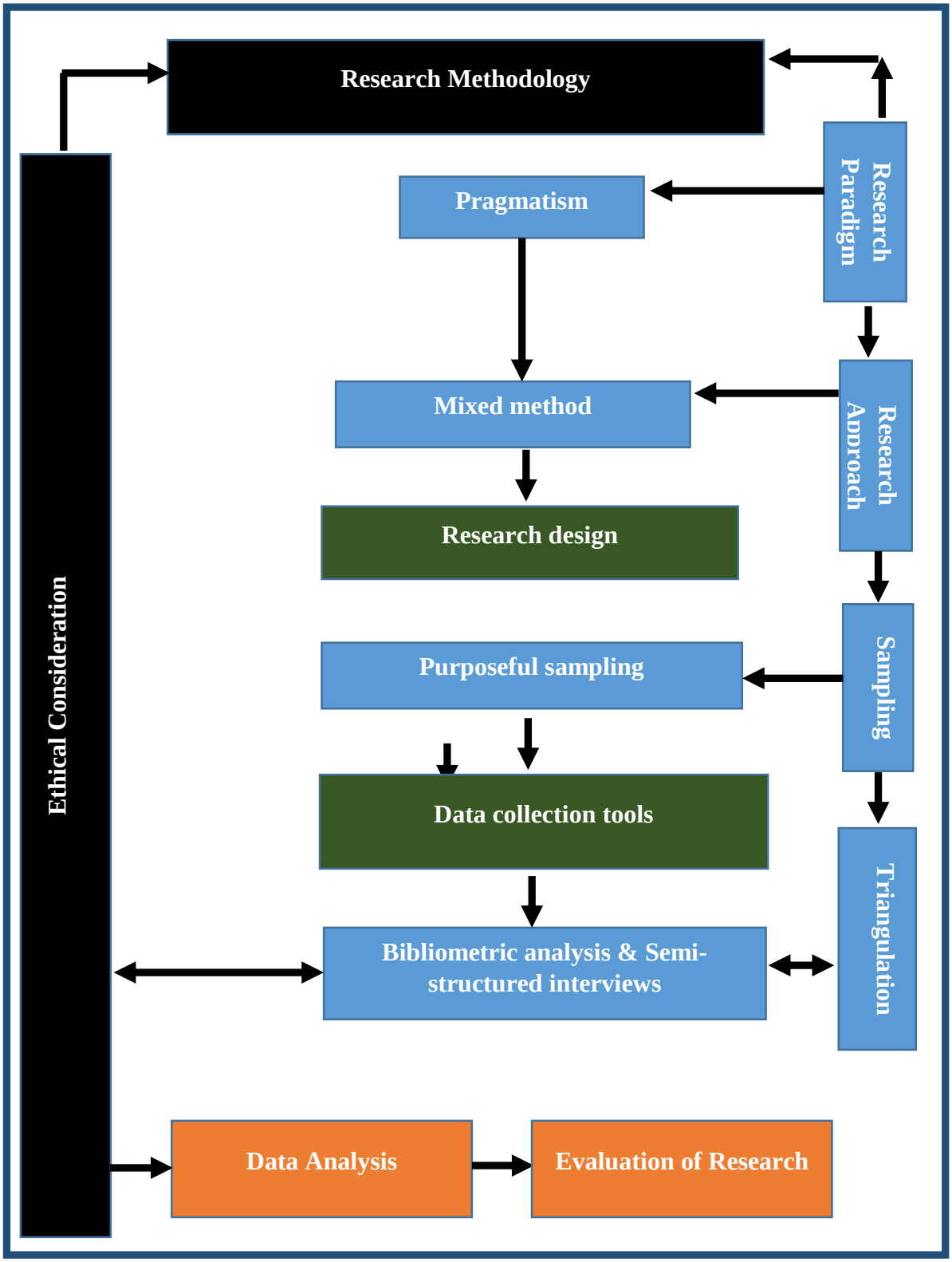


Figure 3.1 Research methodology road map for the current study (Researcher, 2022)

3.2 Research paradigm

Rehman and Alharthi (2016, p. 51) posit that a research paradigm is a basic belief system and theoretical framework coupled with assumptions about ontology, epistemology, methodology and methods. Kamal (2019, p. 1388) also advances that a paradigm can be defined as a set of beliefs around the way in which particular problems come into existence and a number of agreements on how such problems can be explored.

Research is divided into five paradigms: positivist, post-positivist / postmodernism, interpretivist, critical and the pragmatic paradigm (Kivinja & Kuyuni, 2017). This study adopted the pragmatic research paradigm which is based on an explanatory sequential mixed method design because it involves a two-phase data collection process. The pragmatic paradigm was adopted for the study based on the problem of the study. The pragmatic paradigm was found to be appropriate for the reason that collecting different types of data would best provide a more complete understanding of a research problem than either the quantitative or qualitative alone. The study began with a quantitative bibliometric analysis of Scopus (SciVal) and the Web of Science with a view to establish UNAM's top researchers based on their research productivity, visibility and impact. The second phase was qualitative and it used interviews to collect detailed views from purposefully sampled key informants to help explain quantitative results. Creswell (2018, p. 13) acknowledges the works of Murphy (1990), Patton (1990), and Rorty (1990) and his own ideas in arriving at the conclusion that pragmatism provides a philosophical basis for research: Pragmatism is not committed to any one system of philosophy and reality; Individual researchers have freedom of choice; Pragmatists do not see the world as an obsolete unity; and, Truth is what works at the time. It is not based in a duality between reality independent of the mind or within the mind. Thus, in mixed methods research, investigators use both quantitative and qualitative data because the idea is to provide the best understanding of a research problem:

Pragmatist research looks to the what and how to research based on the indented consequences; Pragmatists agree that research always occurs in social, historical, political or other contexts; Pragmatists believe in the external world independent of the minds as well as that lodged in the mind. Thus for mixed methods researchers, pragmatism opens the door to multiple methods, different worldviews and different assumptions, as well as different forms of data collection and analysis (Creswell, 2018, p. 11).

3.3 Research approach

Creswell (2018, p. 3) defines research approaches as the plans and procedures for any research that span the steps from broad assumptions to detailed methods of data collection, analysis and interpretation. The research approaches vary based on the theoretical assumptions and worldviews, research strategies and methods used.

The three (3) research approaches are namely distinguished as qualitative, quantitative and mixed methods.

3.3.1 Qualitative research approach

Creswell (2014, p. 4) posits that qualitative research is an approach that explores and understands the meaning that individuals or groups assign to human problems. In addition, Bezuidenhout et al. (2014) also indicate that qualitative research is mostly regarded as a whole-world experience due to the interest that is tied to human experience, which comprises of personal and subjective originalities based on human experiences associated with a phenomenon.

3.3.2 Quantitative research approach

Creswell (2018, p. 4) advances that a quantitative research approach is commonly known for testing objective theories by examining the relationship among variables.

The variables are measured, typically on instruments, in order to have numbered data analyses by using statistical procedures. Quantitative research involves numeric and statistical approaches to research and it employs inquiry strategies such as experiments and surveys, and collects data with predetermined instruments to produce statistical data. The approach tests numbers and mostly uses open-ended questions in data collection. Quantitative approaches put emphasis on greater sample size hence making it possible for the generalisation of findings (Creswell, 2014, p. 4).

Moreover, Apuke (2017, p. 41) highlights that quantitative research methods is the explanation of a subject or phenomenon through the gathering of data in numerical form and analysing the data with the help of mathematical techniques, in particular statistics.

3.3.3 Mixed method research approach

This study was informed by a mixed method approach. Creswell (2018) conceptualises mixed methods research as an approach to inquiry that entails collecting both quantitative and qualitative data, integrating the two forms of data and using distinct designs that may involve philosophical assumptions and theoretical frameworks.

Schooneboom and Johnson (2017) specifies that mixed method research is the type of research whereby a researcher combines elements of qualitative and quantitative research approaches such as viewpoints, data collection, analysis and inference techniques for the broad purposes of breadth and depth of understanding corroboration.

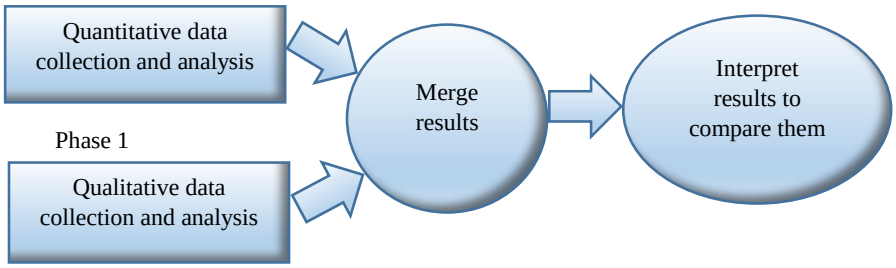
Pragmatism is known as the paradigm that supports mixed method research. Creswell (2018, p. 10) conceptualises that pragmatism is a philosophical worldview that arises from actions, situations and consequences rather than antecedent conditions. Authors such as Maarouf (2019); Mitchell (2018) have stated that pragmatism is considered to be the

philosophical partner of the mixed method research approach as its underlying assumptions provide the essence for mixed research methods.

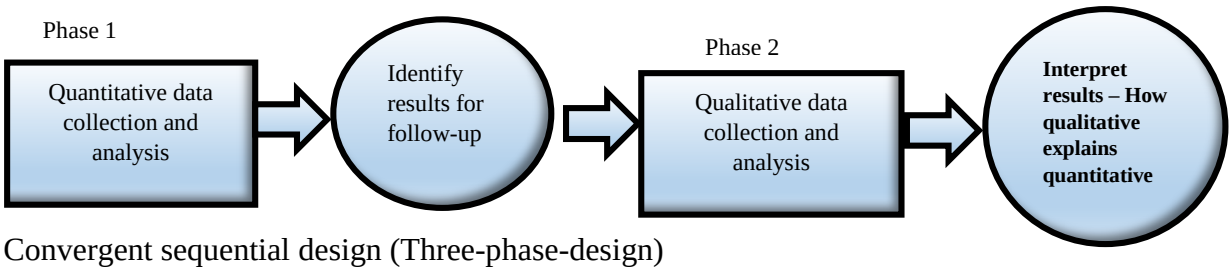
Moreover, Biddle and Schafft (2015) indicate that pragmatism does not only justify the mixed method research approach, but also opens all methodological choices for a researcher where the mixed research approach becomes one way of applying the pragmatic philosophy. Pragmatic researchers are presented with a choice to undertake quantitative, qualitative or mixed research grounded on what serves their research purpose or objective.

This study adopted a mixed-method research approach in which both qualitative and quantitative methods were used in order to avoid the disadvantages of using just one research approach.

2.3.4 Three core mixed methods designs



Explanatory sequential design (Two-phase design)



Convergent sequential design (Three-phase-design)

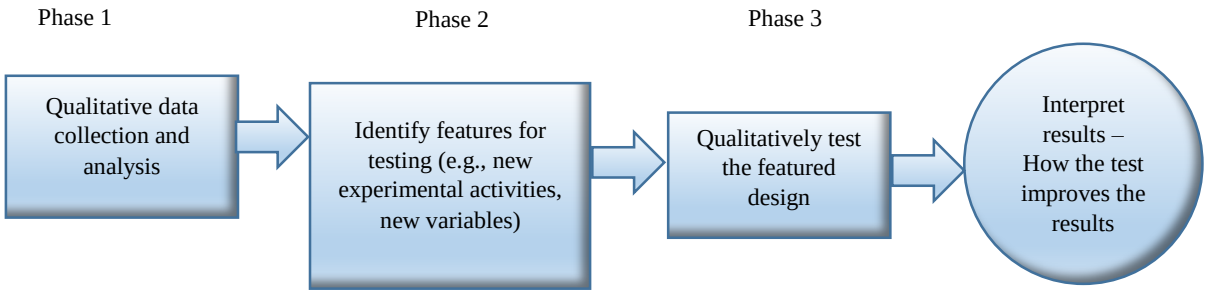


Figure: 3.2: Content adapted from Creswell (2018, p. 218)

3.4 Justification for choosing Mixed Method Research (MMR)

The type of approach and design in a research study is determined by the researcher's intensions and the objectives of the study. The current study has adopted an explanatory sequential mixed methods research design. This was motivated by the fact that the researcher perceived that the sequential data collection procedures would maximise the strengths of both quantitative and qualitative phases by complimenting one another. Mabhiza (2016, p. 113) asserts that combining quantitative and qualitative methods means bringing together the strengths of both methods to compensate for the weaknesses found in the other method.

Creswell (2018, p. 15) conceptualises that with the explanatory sequential mixed methods, the researcher first conducts quantitative research, analyses the results and then builds on the results to explain them more in detail using qualitative research. It is considered explanatory due to the fact that the initial quantitative data are explained further with qualitative data. It is also considered sequential because the initial quantitative phase is followed by the qualitative phase (Creswell, 2018).

The present study employed the mixed method in order to carry out a bibliometric analysis of UNAM's research productivity, visibility and impact using two internationally recognised databases, namely, Scopus (SciVal) and the Web of Science in phase one (1). The first objective of the study was to ascertain the level of awareness and data creation by UNAM researchers. The bibliometric analysis was important in order to establish UNAM's top researchers indexed in Scopus (Scival) and the Web of Science. Secondly, RDM does not officially exist at UNAM, thus the second qualitative phase was to obtain a detailed understanding from UNAM's top researchers on how they deal with issues of RDM. The study's mixed method designed served its purpose, as it was able to answer the objectives of the study.

3.5 Design of the study

The study involved a two-phased design in which the researcher collected quantitative data in the first phase, analysed the results, and the types of questions guided the second qualitative phase. The explanatory research design was selected in order to compliment the quantitative design with more detailed explanations in the qualitative phase. Phase one involved a bibliometric analysis of UNAM's research impact and visibility through Scopus (SciVal) and Web of Science; while in phase two, qualitative data was collected using face to face semi-structured interviews.

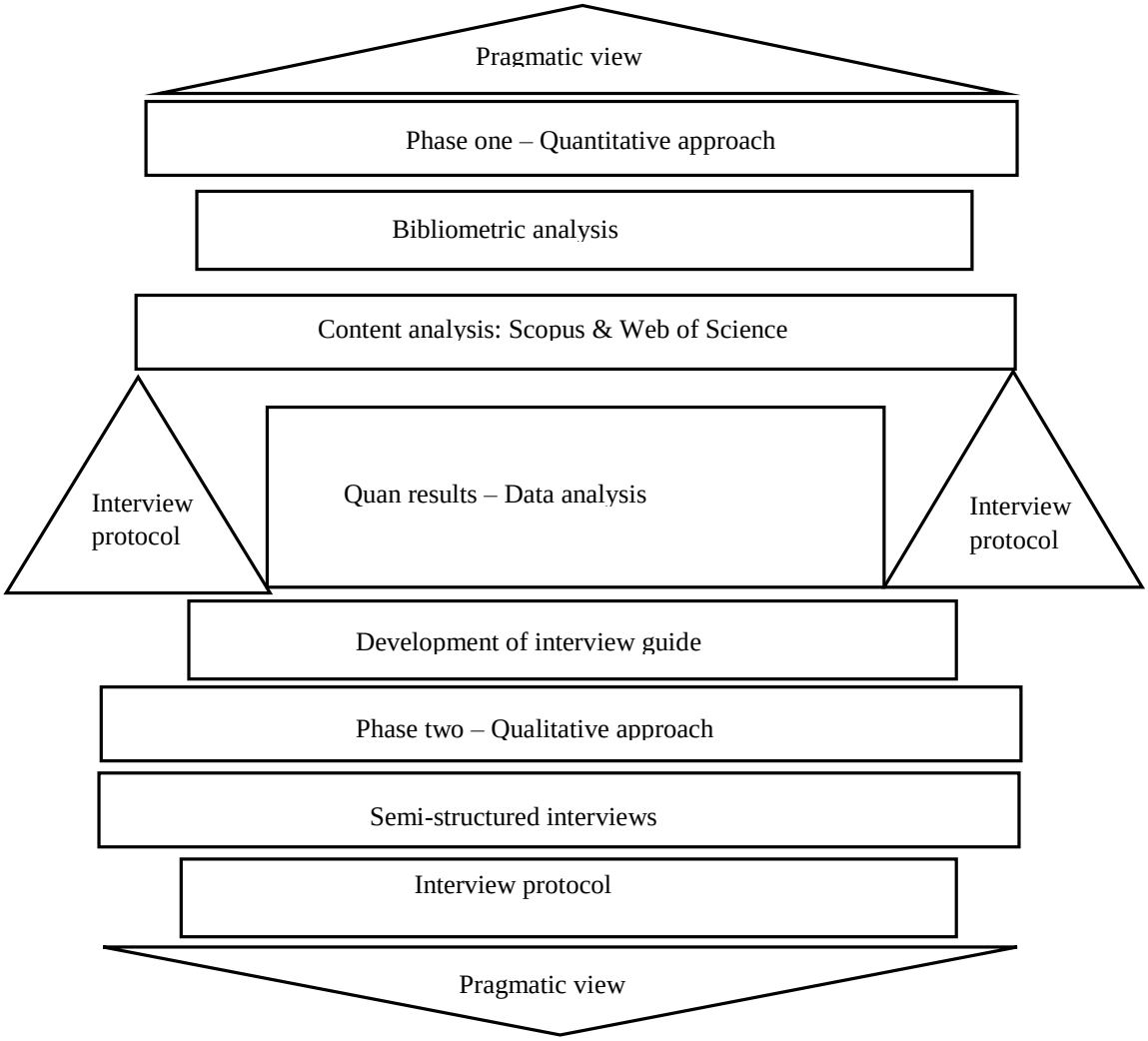


Figure 3.3: Visual model of the pragmatic philosophical worldview of the explanatory sequential mixed methods research design (reproduced from Mabhiza, 2016, p. 117)

3.6 Population of the study

The population of the study is defined by Hassan and Madugu (2015) as any collection of a specified group of human beings or non-human entities like objects, educational institutions, time units, and geographical areas that are drawn by individuals. When a population is not well defined, the researcher will not know what units to consider when selecting the sample (Hassan & Madugu, 2015). Kalusopa (2011, p. 150) also refers to the population of a study as the body of people or collection of items that are under consideration for research. Collis and Hussey (2013) supported the same views. In the current study, a census was used and all UNAM researchers from eight (8) UNAM faculties were targeted. The study population for this study consisted of researchers from all eight (8) UNAM faculties (as of 2020). These are:

a) Faculty of Agriculture and Natural Resources,
b) Faculty of Economic and Management Science,
c) Faculty of Education,
d) Faculty of Engineering and Information Technology,
e) Faculty of Health Sciences,
f) Faculty of Humanities and Social Sciences,
g) Faculty of Law, and
h) Faculty of Science.

This totals to nine hundred thirty-four (934) researchers. [1]

[1] To date, UNAM went through a restructuring process and the number of faculties were reduced to four (4) faculties namely: Faculty of Agriculture, Engineering & Natural Sciences, Faculty of Commerce, Management & Law, Faculty of Education & Human Sciences and Faculty of Health Sciences & Veterinary Medicine. However, restructuring did not dis-align UNAM researchers as restructuring was more for administrative purposes.

3.7 Sample

A study sample is the practice of selecting a statistically representative sample from the population of interest. This practice is important in research because the population normally consists of too many individuals for any research to include as participants. A good sample is a statistical representation of the entire population of interest that is big enough to address the objectives of the study (Bhardwaj, 2019, p. 158; Majid, 2018, p. 3). This study took a census of all the nine hundred thirty-four (934) researchers from eight (8) faculties at UNAM. Sampling took a two-staged approach. Stage one (1) involved a bibliometric analysis of all the nine hundred thirty-four (934) researchers in Scopus (SciVal) and Web of Science. Results from the bibliometric analysis produced the top five (5) UNAM researchers in each faculty based on their research productivity, visibility and impact. In total, forty (40) researchers from all eight (8) faculties at UNAM made up the sample for the current study.

3.7.1 Purposive or judgmental sampling

Etikan et al. (2016, p. 2) narrate that the concept of gathering data in research is of critical importance as gathered data is meant to contribute significantly to a better understanding of a theoretical framework. The purposive or judgemental sampling technique is known as the deliberate choice of a participant based on the qualities that they possess. The purposive or judgemental technique requires no underlying theories or a particular number of participants as the researcher decides what needs to be known and outlines how to find participants that are able to provide the required information by virtue of knowledge or experience (Etikan et al., 2016, p. 2; David & Sutton, 2011).

In addition, Gentles et al. (2015, p. 1778) highlight that the logic and power of purposeful or judgemental sampling lies in selecting information-rich cases for in-depth study.

Information-rich cases are those from which a researcher can learn and gain a great deal on matters of critical importance to the purpose of the inquiry. The concept of studying information-rich cases is known to enable the author to obtain insights and an in-depth understanding of the subject matter. UNAM researchers that were interviewed in the second qualitative phase were purposefully selected based on results emanating from the first bibliometric phase. The top five (5) UNAM researchers were purposefully selected from each of the eight (8) faculties at UNAM based on their research production, visibility and impact as indexed in Scopus (Scival) and Web of Science. Therefore, the purposeful or judgemental technique was the most appropriate for the current study, and it enabled the researcher to identify key participants with significant knowledge and the experience which was preferred to yield knowledge on RDM at UNAM.

3.8 Data collection methods and instruments

Kalusopa (2011, p. 151) emphasises that data collection refers to the way in which research data is collected in the field, to a specified sample or study population. Data collection is influenced by the research methodology as it determines the types of techniques and instruments to be used and that such techniques ensure their validity and reliability. In this current study, data was collected in two (2) sequential phases. Phase one (1) quantitative data was collected through bibliometric analysis tools based on international bibliographic databases namely Scopus (Scival) and Web of Science. In Phase two (2), qualitative data on the understanding of the depth of RDM was collected from UNAM's top five (5) researchers from each faculty as established through bibliometric data using semi-structured interview guides. The researcher made certain that the data was categorised according to its objective in order to facilitate the transition into meaningful information.

The in-depth interviews enriched rigour and trustworthiness to enhance the reliability and validity of the qualitative data. The second qualitative phase engaged UNAM's top researchers to get an in-depth of their understanding about research data management and how they deal with their research data throughout the data lifecycle. In this study, two (2) techniques were employed for data collection, namely, content analysis and interviews. Each of these tools is described briefly in detail below.

3.8.1 Bibliometric analysis

Ellegaard and Wallin (2015) depicts that bibliometric methods are now firmly established as scientific specialties which forms an integral part of research. In addition, Carpenter et al. (2014, p. 162) underscore that bibliometric analysis refers to the quantitative analysis of publication data using a document, author or source to uncover characteristics, patterns, and relationships so as to demonstrate individual investigators' or research teams' productivity, quality or impact.

According to Kalusopa et al. (2020, p. 493), bibliometric analysis measures the research productivity of an individual and groups by applying different metrics such as the number of publications, author status and publication sources in which researchers publish, the impact factor score, citation, and h-indices using popular international bibliographic databases such as Scopus (SciVal) and Web of Science. The same sentiments are shared by Van n (2005), who indicates that bibliometrics are common research instruments for systematic analysis. The current study explored the research productivity and impact of UNAM using bibliometric citation and views. Scopus (SciVal) and Web of Science databases were used to analyse the impact of research into two general classes namely the journal level metrics and article-level metrics.

Journal-level metrics rank the journal in a field of research as it is assumed that articles published in high impact journals are of high quality and therefore, researchers with

publications in high impact journals are doing high quality research. Article-level metrics are about the individual published articles into a single measure to determine the quality of the work.

3.8.2 Interviews

Bolderston (2012, p. 68) defines interviews as a way of listening to and gaining an understanding of people's stories that are attached to their experiences and knowledge. Interviews are critical in research as those interviewed can express their viewpoint without a framework imposed by the researcher.

According to Alsaawi (2014), there are different types of interviews that can be adopted in research. The different types of interviews have their own objectives and focus. The objectives and focus are determined by research questions and the information that is needed to provide holistic answers to the questions. The different types of interviews range from structured, unstructured and semi-structured interviews.

Semi-structured interviews are known to be common in social science research. Semi-structures interviews are a combination of structured and unstructured interviews whereby questions are pre-planned prior to the interview but the interviewer gives the interviewee a chance to elaborate and explain particular circumstances through the use of open-ended questions (Alsaawi, 2014).

The current study used semi-structured, face to face interviews. These are the type of interviews whereby the researcher is physically required to be present when asking questions and to clarify where the interviewer does not understand clearly. Bolderston (2012, p. 68) highlights that face-to face interviews entail one to one sessions between the interviewer and interviewee and they allow for probing and clarifications of matters that may be raised. In addition, Alamri (2019) proffers that using interviews helps the

researcher to obtain the perspective of the interviewee by interpreting the meaning of the described phenomenon and it is flexible as it allows the interviewer and interviewee to elaborate on the topic to pursue an idea or response in greater detail.

Therefore, the interview method was adopted and used to gather quality information about RDM practices at UNAM. The interviews conducted were ideal for understanding the experiences and practices of RDM as required by different publishers across the globe.

3.9 Procedure

Creswell (2014, p. 177) defines a research procedure as the different steps that a researcher should undertake in order to get access to the participants of their study. In the current study, the researcher applied for a research permission letter from the University of Namibia before collecting any research data. An informed consent letter was distributed to the participants to read, agree and sign that they were going to voluntarily participate in the study.

The participants were contacted telephonically and via email to book for a time slot that they were available for face to face interviews. Interviews were arranged to take place at the participants' offices. The researcher introduced himself and the research topic before the commencement of the interview and explained why the participants were purposefully chosen for the current study. The researcher gave each participant a short brief about the RDM context at UNAM and what the researcher aimed to achieve at the end of the study.

An informed consent letter was then given to the participants to sign before the start of the interviews. Interviews were recorded using a digital voice recorder with the participants' consent. Each interview persisted for 10-35 minutes. During the interviews, the researcher took short notes to help during transcribing. The researcher thanked the participants for their time and expressed how their participation was a contribution towards the study.

3.10 Validity and reliability

Surucu and Maslakci (2020) enunciates that validity and reliability are critical in research. Although closely linked, they express different properties of measuring instrument as an instrument maybe reliable without being valid, but is measuring an instrument is valid, it is also likely to be reliable. Reliability alone is not sufficient to ensure validity even if the test is reliable, it may not accurately reflect the desired quality (Surucu & Maslakci, 2020). The concept of validity and reliability are vital in determining the quality of research as research is considered valid and reliable when conclusions are accurate and the findings are repeatable (Powel, 2010, p. 60). Validity and reliability are key concepts in modern research as they are used for enhancing the accuracy of the assessment and evaluation of research (Mohajan, 2017, p. 2).

It is difficult to describe the effects of measurement errors on theoretical relationships that are being measured without assessing the validity and reliability of research. Validity and reliability can be enhanced by using numerous types of methods to collect data in order to get factual information. Creswell (2014) emphasises that questions of validity and reliability are worth discussion in qualitative research just as in quantitative research although they are treated differently. In qualitative research, validity can be attained through collecting and analysing valid, strong and convincing arguments from the participants.

Validity and reliability as concepts are central to a study such as this one. The current study used a mixed method approach. The researcher had to discuss the validity of qualitative findings in order to determine how accurate the findings are. Qualitative data was collected through semi-structured interviews with the top ten researcher from each faculty at UNAM as well as faculty research representatives. Umana (2020, p. 57) confirms that one way to guarantee validity and reliability is by piloting data collection instruments. Pre-testing interviews helps to determine whether interview questions are capable of attaining the

information needed to answer the research objectives. To ensure that the interview guide was valid and reliable, it was given to five librarians for scrutinising and proofreading.

Interview guides were simplified to ensure that participants understood the questions with ease, which would enable them to give clear, consistent and accurate answers, thus minimising errors.

3.10.1 Methodological triangulation

Honorene (2017) acclaims, that triangulation refers to the application and combination of several research methods in the study of the same phenomenon. The practice involves mixing data or methods so that diverse viewpoints or standpoints cast light upon a given topic. It helps with validating claims that may arise from an initial pilot study. In addition, Kalusopa (2016, p. 146) refers to the concept of methodological triangulation as the use of numerous methods, techniques and tools for data collection that is combined in the same study, and this is meant to overcome the potential bias and desolation of a single method approach. Triangulation decreases the inadequacies that are found in individual methods and the threats to internal validity. Triangulation takes into account the advantages from both the quantitative and qualitative methods by ensuring that they complement each other, thereby making a stronger research design which results in more valid and reliable findings (Kalusopa, 2016, p. 146).

Kalusopa (2016, p. 146) indicates that there are four types of triangulation which are commonly used in mixed method studies:

1. Data triangulation: which strengthens research findings by applying multiple ways to collect and analyse data involving time, space and persons;
2. Investigator triangulation: entails the usage of multiple researchers rather than a single observer in an investigation;

3. Multiple triangulation: this type of triangulation entails the researcher combining multiple observes, theoretical perspectives, sources of data and methodologies in one investigation; and
4. Theory triangulation: involves using more than one theoretical scheme in the interpretation of the phenomenon.

This study adopted the methodological triangulation of both the quantitative and qualitative data collection methods. The triangulation of bibliometric analysis and interviews were used. Triangulation was used to strengthen the limitations and weaknesses of the quantitative and qualitative data collection approaches used in the study, hence ensuring enhancement in validity and reliability of the research results.

3.11 Data analysis

Data analysis discusses the use of reasoning in an attempt to understand the data that have been assembled and it involves for stages namely: description, interpretation, conclusion and theorisation (Masenya (2018, p. 70; Zikmund & Babin, 2013, p. 68).

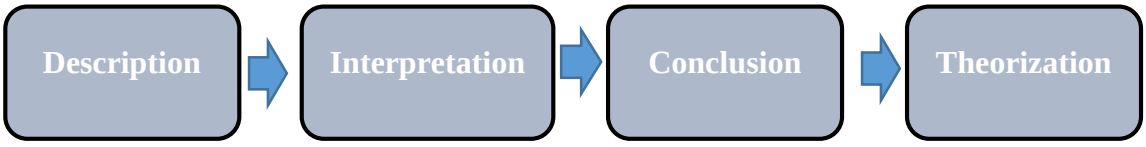


Figure 3.4 Source: Masenya (2018, p. 70)

At the description phase, the researcher describes the data by defining what they see in the study's collected data. The researcher then indicates what they think the data means during the interpretation phase of data analysis. After the data has been interpreted, the researcher is then responsible for making conclusions by drawing minor assumptions from the data which contribute to the study's conclusion. The last theorisation phase involves the researcher referring back to their literature review and analyses where their findings fit with or contradict the findings of the theorists that were mentioned in the study (Masenya, 2018,

p. 71; Quinlan et al., 2015). Onwuegbuzie and Combs (2011, p. 2) proffer that the term mixed analysis is used for analysing data in mixed method research. Mixed analysis entails the utilisation of both quantitative and qualitative analytical techniques within the same framework. Mixed analysis might be based on one of the existing mixed methods research paradigms (e.g. pragmatism, transformative-emancipatory) such that it meets one of more of the following purposes: triangulation, complementarity, development, initiation and expansion (Onwuegbuzie & Combs, 2011, p. 2).

3.11.1 Quantitative data analysis

Kalusopa (2016) asserts that quantitative data analysis applies statistical techniques for collecting, organising and interpreting data. Statistical data analysis is vital to the field of research in library and information discipline. Analysing quantitative data varies from simple descriptive analysis to more elaborate reduction and multivariate associating techniques.

The quantitative data gathered for this study was analysed using bibliometric analysis. Bibliometric analysis is known as the quantitative analysis of publication data by using data sources, publication counts (full and fractional counting), citations and co-citations and factors that affect citation impact. The productivity, visibility and impact of UNAM researchers was be analysed through Scopus (SciVal) and Web of Science.

3.11.2 Qualitative data analysis

Flick (2013) avows that qualitative data analysis is the classification and interpretation of linguistic materials to make statements about implicit and explicit dimensions and structures of meaning-making in the material and what is represented in it. Additionally, Kalusopa (2016, p. 148) affirms that qualitative data analysis refers to data collection techniques by means of interviews, observation, documentary and open-ended

questionnaires. Mabhiza (2016, p. 131) emphasises that qualitative data analysis is a wide analysis of the data, splitting it into various components and defining the characters and dimensions in order to make an inference about the whole object. For this study, qualitative data analysis adopted the use of coding and thematic analysis. The qualitative data was analysed by utilising content analysis methods by transcribing, grouping, coding and interpretation based on the research objectives. Umana (2020, p. 20) indicates that content analysis entails the analysis of written, verbal or visual communication messages. Content analysis in qualitative research involves researchers to adopt the use of coding. Data was analysis based on the following themes: research data creation, data processing and analysis, data preservation and data accessibility and the extent of data re-use at UNAM.

3.12. Research ethics

Research entails collecting data from people and about people, and this requires researchers to anticipate ethical issues. Researchers need to protect their research participants, develop trust, and promote the integrity of research, as well as guard against misconduct and impropriety that might reflect on their organisation or institution (Creswell, 2018, p. 88). Ethical issues need to be considered prior to conducting the study, beginning a study, during data collection and analysis and in reporting, sharing and storing the research data.

According to Showkat and Parveen (2017), researchers have to own the sole responsibility for ethical conduct for their own research. Researchers have the responsibility of ensuring the safety, dignity, rights and well-being of participants. Good ethics entails that the research is conducted with honesty, objectivity and integrity.

Consent from participants should be agreed upon before they are involved in the study. Culture, values, religions, and economic status should be considered and valued (Showkat & Parveen, 2017).

Kalusopa (2016, p, 149) stresses the importance of scholars and institutions of higher learning to comply with research ethics. This entails researchers being familiar with risks to research before, during and after completion of their studies.

Risks comprise of privacy, breaches of confidentiality, deception and informed consent. Ethical clearance was considered during the research as the anonymity and confidentiality of participants were observed. In addition, Louw (2014) emphasises that researchers are tasked to ensure that ethical issues are considered, especially when human participants are involved.

The current study involved human participants, thus the researcher was required to apply for ethical clearance from UNAM's Research and Ethics Committee (UREC) which was granted before the research embarked on data collection. The UNAM Policy on Research Ethics clearly specifies how researchers have to adhere to issues of ethics in order to avoid acts of misconduct in research, which include data fabrication, falsification and plagiarism.

All the sources cited in the current study were referenced and original authors were acknowledged to avoid plagiarism. The current study adheres to ethical guidelines as the researcher adopted the voluntary participation and anonymity since the participants in the study were not asked to state their details. The recorded interviews and transcripts data are preserved on Google Drive. Google offers cloud computing services and the researcher embraced such available technologies to preserve the data in electronic format. No raw data will be deleted and it will be migrated to new available technologies to ensure that it continues to be accessible to the researcher.

Moralities of confidentiality were firmly upheld as the researcher used coding instead of participants' names in order to conceal and protect their identities. The current study is for academic purposes and there is no conflict of interest surrounding the research.

3.13 Evaluation of research methodology

Kalusopa (2016, p. 151) observes that research methods have to be evaluated in order for them to give a detailed brief on what information was required, the collection methods and analysis. It is a practical norm in research to evaluate the research methodology in terms of research design, limitations and ethical issues that would have surfaced during the study. In addition, Shaambeni (2019) indicates that the evaluation of research methodology notifies other researchers on possible challenges encountered and what worked best with the research in order to avoid similar mistakes if the study is to be replicated. In this current study, emphasis on the evaluation of the research methodology was placed on the research approach and the data collection instrument employed. The current study was based on a mixed method research design and it was undertaken in two sequential phases. The mixed method design was applied to establish RDM practices among UNAM researchers who are visible on Scopus (SciVal) and Web of Science. Scopus (SciVal) and Web of Science are the only recognised international databases measuring impact factors across the globe. The two databases were considered because UNAM does not have an RDM framework in place.

Despite the popularity of Web of Science, UNAM does not subscribe to the Web of Science. To solve the problem, a Librarian from South Africa volunteered to rescue the study in harvesting UNAM data on research production, visibility and impact which is available on the Web of Science. The study utilised semi-structured interviews for the second qualitative phase. One of the challenges was that it was difficult to arrange face to face interviews with the participants due to Covid-19 protocols that were in place nationally.

3.14 Chapter summary

This chapter discussed the methods and techniques that were used to assess RDM at UNAM. The explanatory research design was selected in order to compliment the quantitative design with more detailed explanations in the qualitative phase. Phase one involved a bibliometric analysis of UNAM research impact and visibility through Scopus (SciVal) and Web of Science; while in phase two, qualitative data was collected using face to face semi-structured interviews. Reasons for the selection of each instrument for data collection were explained, and principles such as validity, reliability and the ethical standards that informed the research process were presented. The analysis of data and evaluation of the research methodology were also discussed in the chapter. The results of the assessment are presented and interpreted in the chapter that follow.

CHAPTER FOUR

PRESENTATION OF RESEARCH FINDINGS

4.1 Introduction

The purpose of the presentation of findings in research is to present empirical findings in an attempt to answer the research objectives (Garaba, 2010, p. 186). This chapter presents the findings of Phase one (quantitative) and Phase two (2) (qualitative) of the study. The findings of phase one (1) were obtained from data gathered through a bibliometric analysis of UNAM's research productivity, visibility and impact using Scopus (SciVal) and Web of Science databases. The findings of phase two were obtained through interviews that were held with 40 purposefully selected UNAM top researchers from the eight (8) faculties at UNAM.

4.2 Presentation of quantitative data: Bibliometric analysis - Phase one

Bibliometric analysis is key in measuring the degree of research or academic output. It measures the number of academic publications and citations of a single researcher, group or an institution (Byl et al., 2016). The bibliometric analysis of universities research has gained importance in higher education as a practice of strategic input to decision making. Bibliometric analysis is measured through two major units, which are scientific publication as an indicator of research output, and citations established based on impact on the scholarly community (Haustein & Lariviere, 2015).

The current study undertook a measure of UNAM's research productivity, visibility and impact from 2011 to 2020.

4.3 UNAM's research outputs

UNAM is an emerging university in Southern Africa, with a relatively comparative research and development headcount based on research outputs. Established in 1992, UNAM features relatively very well among the Southern African Development Community (SADC) member states.

Research productivity between UNAM and South African universities can be attributed to the different research environment, rewarding platforms, financial expenditures and development. In line with the above, the University of Namibia has about 1450 academics who are obliged to commit about 30% of their work to research activities (Kalusopa et al., 2020).

4.3.1 Research productivity

Research productivity is carried out to establish evidence of accountability or to consult scholarly behaviour and performance, therefore, measuring research productivity involves the co-existence of numerous factors at input and output levels (Pal, 2020).

In academic institutions, lecturers are the main research actors, as they are required to conduct research, which is one of the three main responsibilities apart from teaching and community service. The increase in research productivity, both in terms of quantity and quality, is measured to determine the extent of research in an institution (Sanmorino et al., 2021, p. 184). Oruc (2021) specifies that researcher's research output serves as a critical measure of research productivity based on quality and quantity. Four criteria used for this purpose include impact factor (CiteScore), citation number, journal collection, and the number of the publication. Quality is represented in the elements of impact Factor (CiteScore), number of citations, and quality of the journal collection. Meanwhile, the number of publications reflects quantity.

MacLeod and Urquiola (2021) depict that evaluating the results of productive researchers can be challenging. One researcher might be a co-author on all the papers that a research group has produced, while another researcher may prefer to write papers alone. One approach is not necessarily more productive than the other. In measuring performance, another critical factor is that researchers' performance should be compared to that of other researchers within the same research field. In terms of publication type, there are more articles produced by UNAM's researchers indexed in Scopus (1085) and Web of Science (1079). This is justified by the fact that articles are frequently published through different volumes and issues. Data depicts that there are more publications indexed per type in Scopus as compared to Web of Science (e.g. article, conference paper, book chapters and others) as illustrated in Figure 4.1 and 4.2

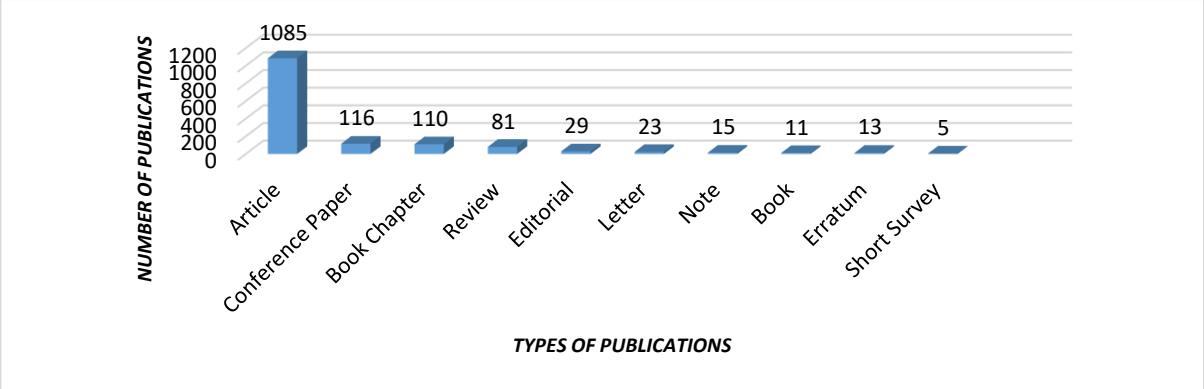


Figure 4.1: Research outputs in Scopus

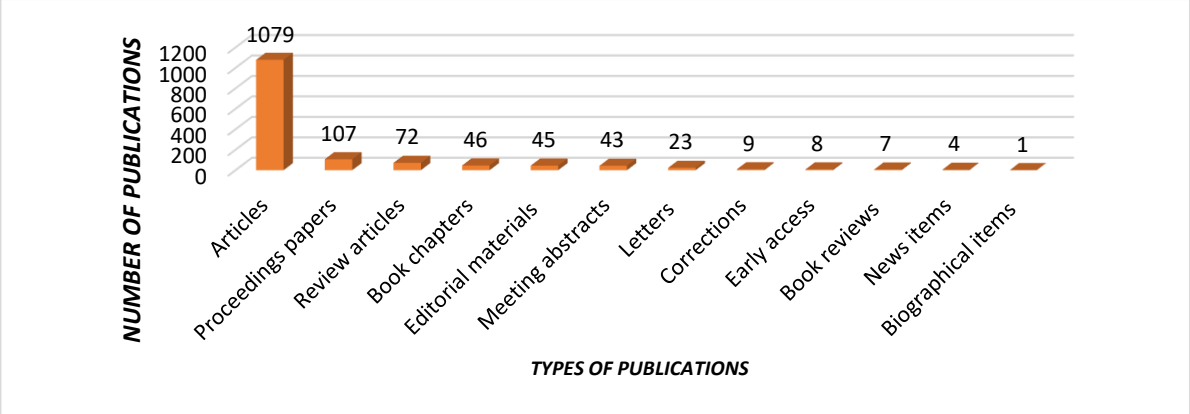


Figure 4.2: Research output in Web of Science

Data illustrates that there has been a gradual increase in the number of UNAM's research publications presented in Scopus and Web of Science between 2011 and 2020. Although there has been a gradual increment, data from Scopus shows that there was a considerable decrease in publications recorded between 2012 and 2013, followed by a rapid increase between 2013 and 2014. Overall, in total, as indicated in figure 4.3 and figure 4.4, there are more publications documented in Scopus (1488) as compared to Web of Science (1247).

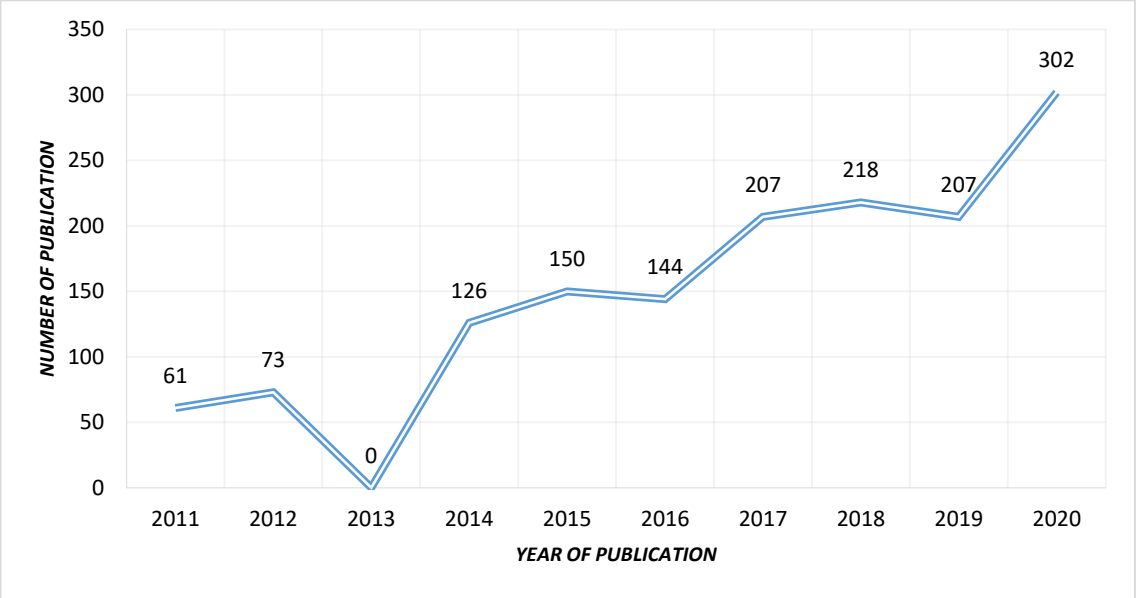


Figure: 4.3 Scopus annual publication indexed

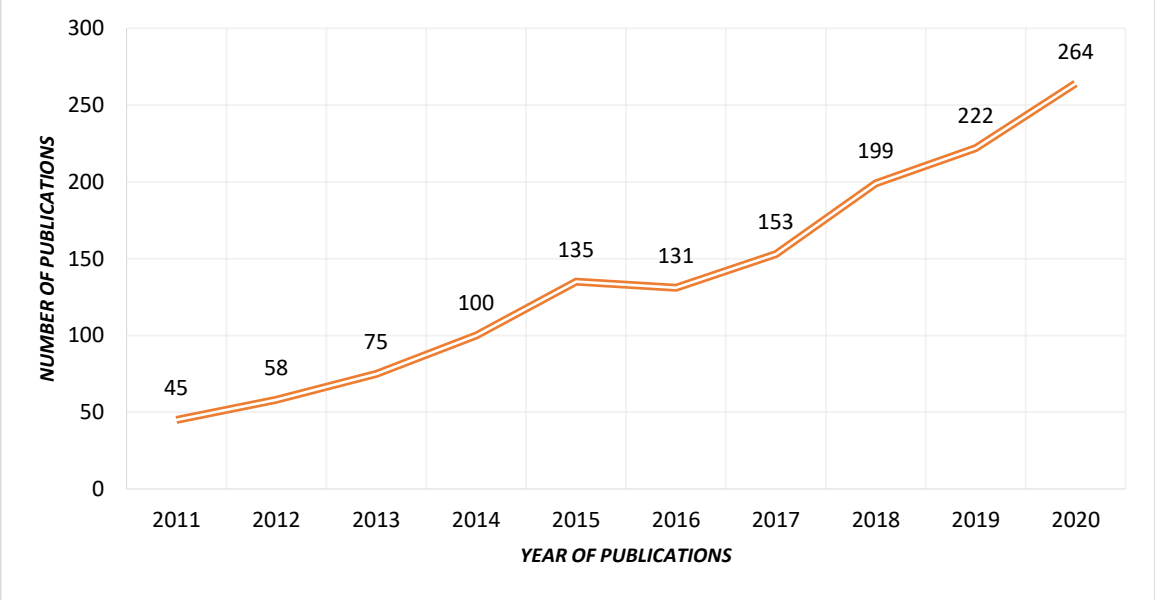


Figure 4.4 Web of Science annual publication indexed

Subject areas covered by Scopus and Web of Science are important to the analysis of bibliometric data to show the areas in which UNAM researchers are actively involved. This gives a comparison of disciplines from which more research is indexed as compared to others. As demonstrated by Figure 4.5 and 4.6 the data indicates the number of publications per subject area. Data from Scopus and Web of Science depicts that there is more presence from Medicine and Social Sciences as compared to other disciplines such as Engineering and Computer Science, which may be attributed to UNAM publications indexed from those particular disciplines.

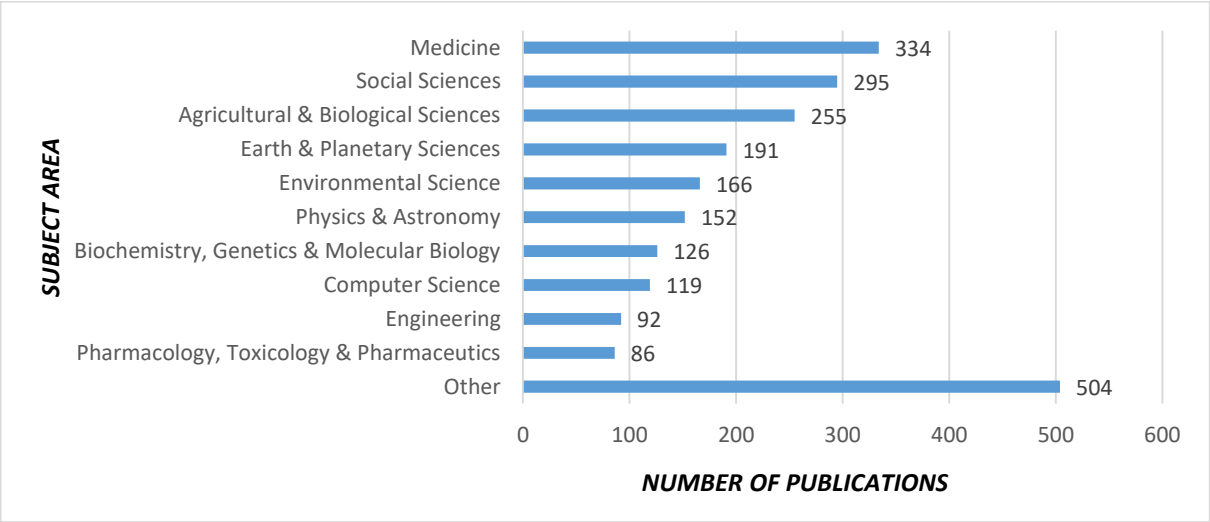


Figure 4.5 Number of publications per subject area in Scopus

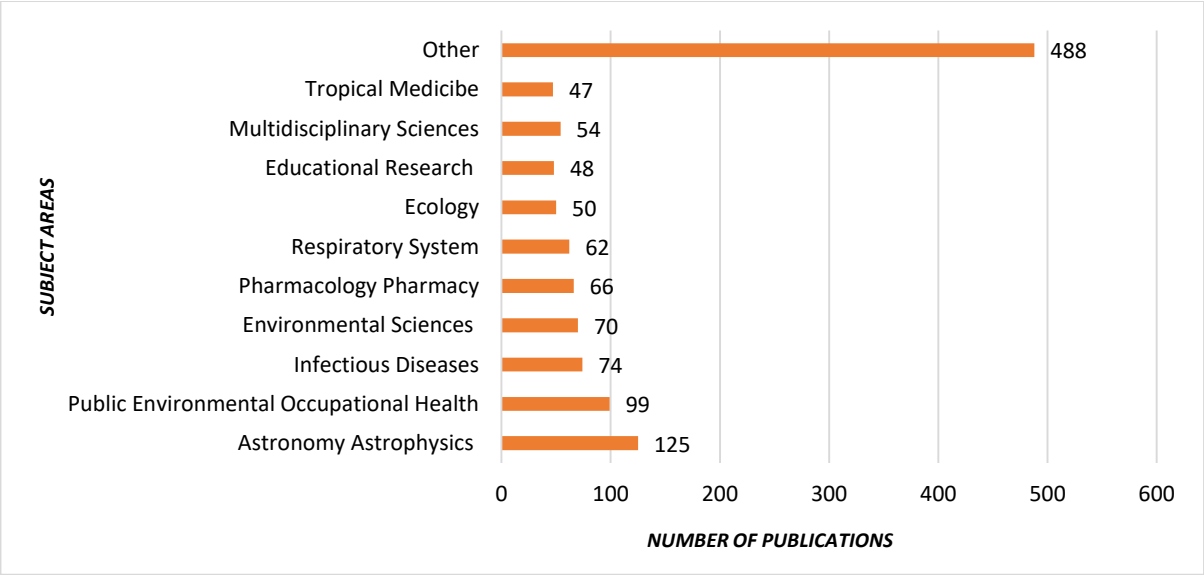


Figure 4.6 Number of publications in subject area in Web of Science

4.3.2 Research visibility and impact

The main measurement that determines author productivity is attributed to the number of citations which is related to the visibility of research. Citations in research describes the number of times that particular research has been used by other scholars. The basic citation analysis is linked to how often the paper has been cited which determines the influence in the field (Ebrahim et al., 2014, p. 120).

Figure 4.7 shows the citation of UNAM publications in Scopus and Web of Science over the study period (2011 to 2020). Data designates that UNAM publications are getting a fair number of citations, which has an impact in the global system. Additionally, in Scopus, there were 1584 publications cited between 2011 and 2020, excluding self-citations with an h-index of 58.

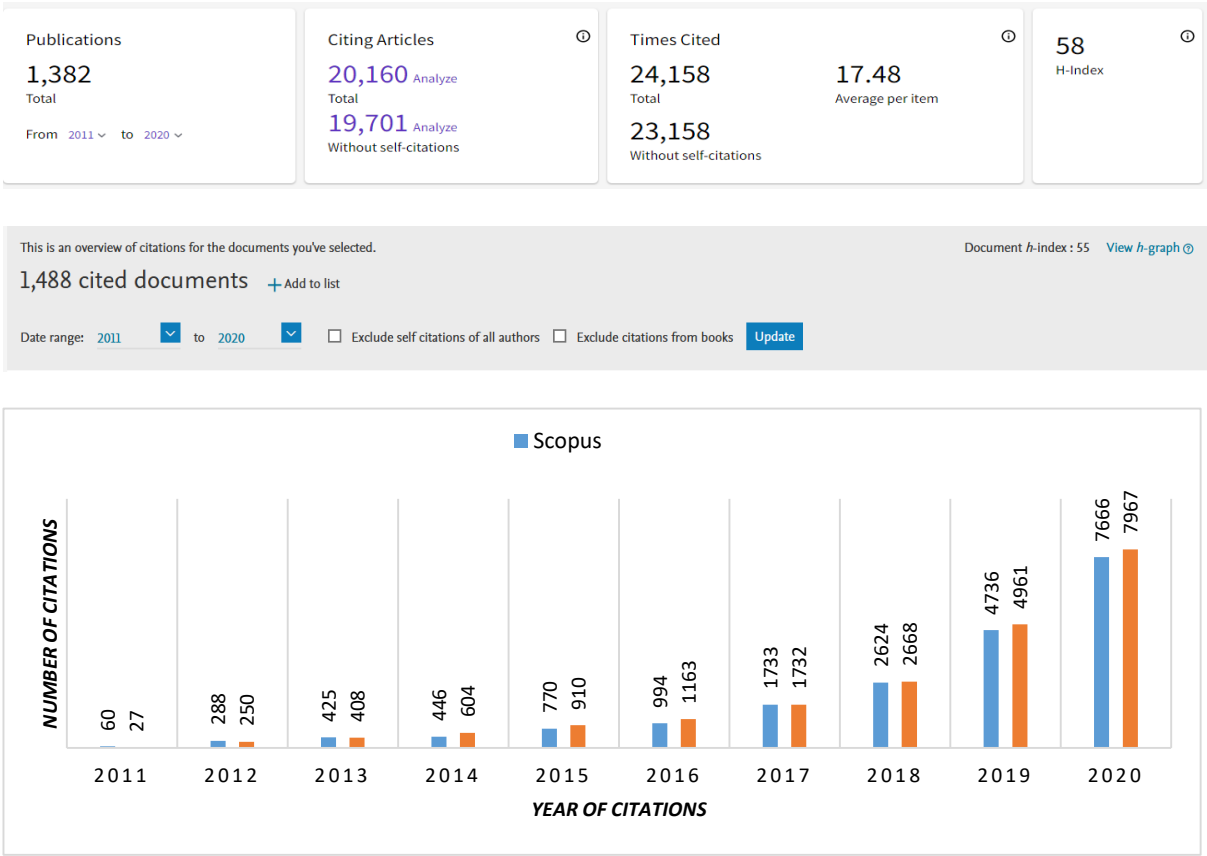


Figure 4.7 Citation of UNAM publications in Scopus and Web of Science

4.3.3 Field-weighted citation impact

Purkayastha et al. (2019, p. 636) define field-weighted citation impact as an indicator of mean citation impact that compares the actual number of citations received by a document with the expected number of citations for documents of the same document type. The metric is always defined with a reference to a global baseline of 1.0 and intrinsically accounts for differences in citation accrual over time, differences in citation rates for different document types (e.g., reviews typically attract more citations than articles), as well as subject-specific differences in citation frequencies overall and over time and document types (Purkayastha et al., 2019, p. 636). Data reveals that Medicine continues to have a high number of citations which contribute to great **field-weighted citation** impact as compared to other disciplines such as Mathematics and Psychology.

Table 4.1 shows the number of scholarly outputs per subject distribution produced by UNAM between 2011 and 2020 in SciVal (Scopus). The table depicts the number of citations per distribution area and the field-weighted citation impact based on subject areas.

Table 4.1 Number of scholarly outputs per subject distribution

Subject area	**Scholarly output**	**Citations**	**Field-weighted citation impact**
Medicine	362	14,880	5.56
Social Sciences	317	1,521	0.90
Agricultural & Biological Sciences	278	2,395	0.75
Earth & Planetary Sciences	207	5,679	1.75
Environmental Science	177	1,354	0.77
Physics & Astronomy	164	6,631	2.34
Biochemistry, Genetics and Molecular Biology	129	1,314	0.75
Computer Science	123	470	0.49
Engineering	97	650	0.90
Pharmacology, Toxicology and Pharmaceutics	88	669	1.07

Immunology and Microbiology	85	902	0.98
Arts and Humanities	68	215	1.04
Multidisciplinary	66	2,038	1.57
Chemistry	65	406	0.61
Business, Management and Accounting	47	90	0.39
Materials Science	43	793	2.73
Mathematics	40	79	0.36
Psychology	40	982	1.91
Economics, Econometrics and Finance	39	79	0.57
Energy	39	244	0.76
Veterinary	38	256	0.91
Chemical Engineering	36	281	0.69
Health Professions	27	113	0.64
Nursing	19	224	2.82
Decision Sciences	17	3	0.08
Neuroscience	4	175	3.29
Dentistry	3	8	0.51

4.3.3 Research collaboration

Collaboration among academic and research institutions is recognised as a measure of facilitating knowledge exchange (Hottenrott & Lawson, 2017). Additionally, Salmi (2016) highlights that institutional affiliation affects the value that is credited to individual researchers through institutional status.

As shown in Figure 4.8 and 4.9 below, Scopus and Web of Science, there is more optimistic collaboration internally among UNAM scholars as compared to external collaboration. Apart from internal collaboration, the data clearly shows that there is an average level of collaboration between UNAM, South African universities and European universities. The data explains that UNAM collaborates more with North-West University and University of

the Witwatersrand in South Africa and collaboration with European universities are greater with Sorbonne University (France) and Heidelberg University (Germany).

Data does not show any collaboration between UNAM and other universities in the SADC region or other African countries on the continent. This calls for more collaboration between UNAM and other universities on the continent.

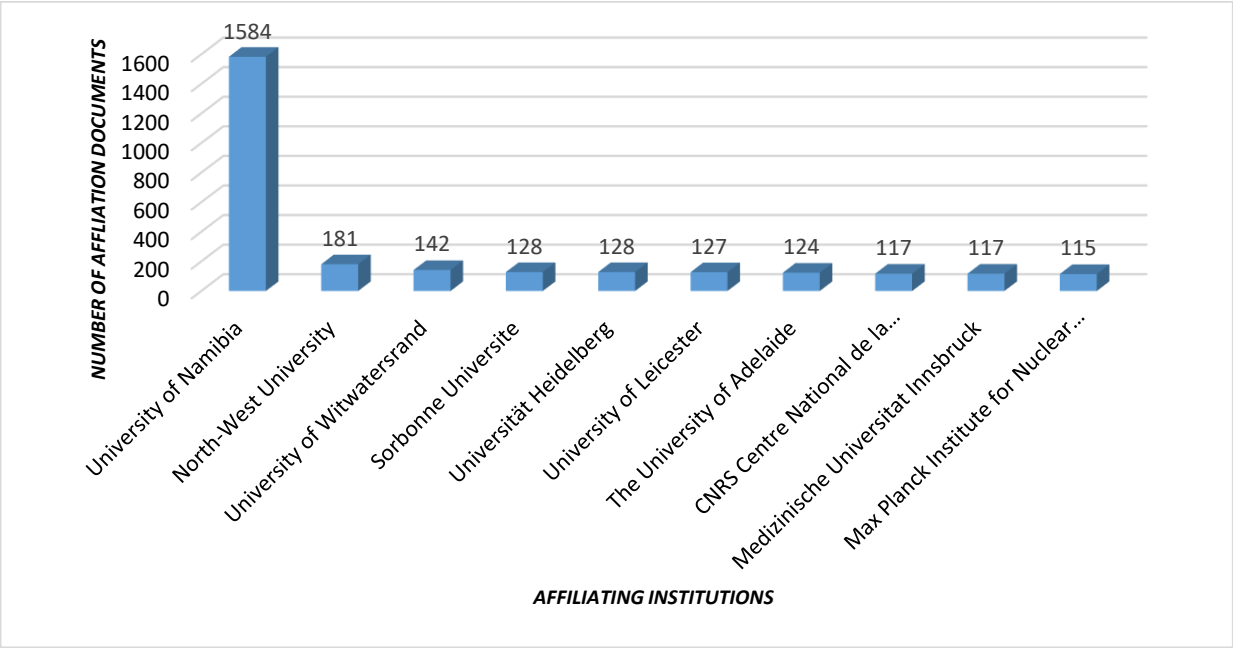


Figure 4.8 Number of documents because of collaboration in Scopus

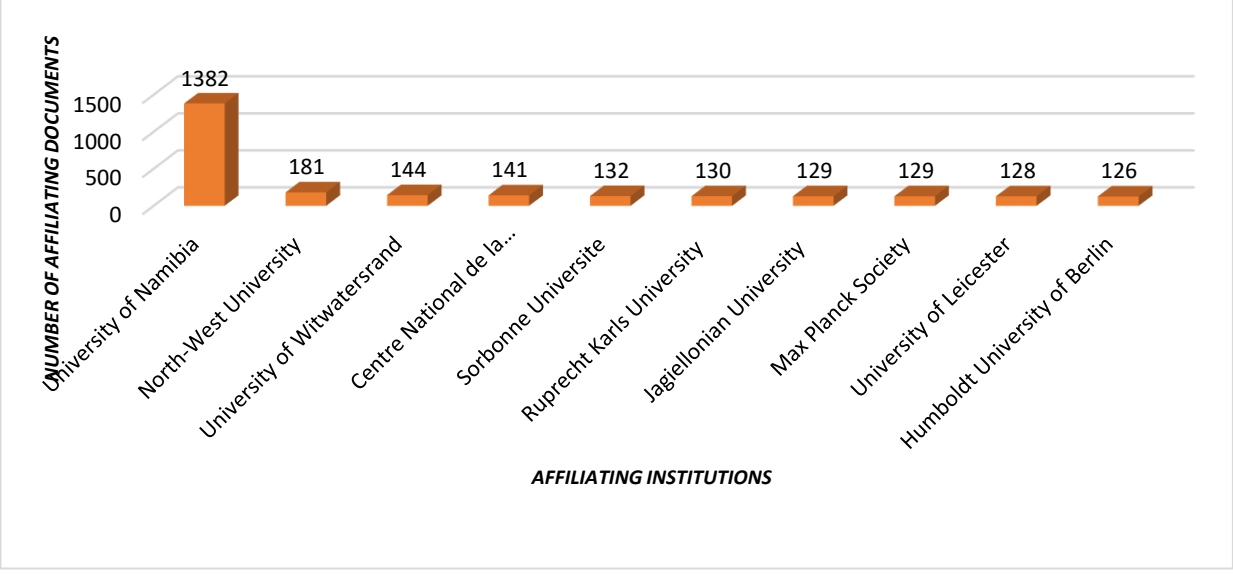


Figure 4.9 Number of documents because of collaboration in Web of Science

Data depicts collaboration of co-authored publications between UNAM, South African universities and the Namibia University of Science and Technology (NUST). The figure shows that UNAM has mostly collaborated and co-authored with the University of the Witwatersrand (100 publications) and North West University (99 publications) in South Africa. UNAM has collaborated more with seven South African universities than with one of the leading local Namibian university, which is the Namibia University of Science and Technology (NUST) with 35 publications. Although UNAM has strong collaboration with seven South African universities, the data shows that collaboration is low as indicated in Figure 4.10.

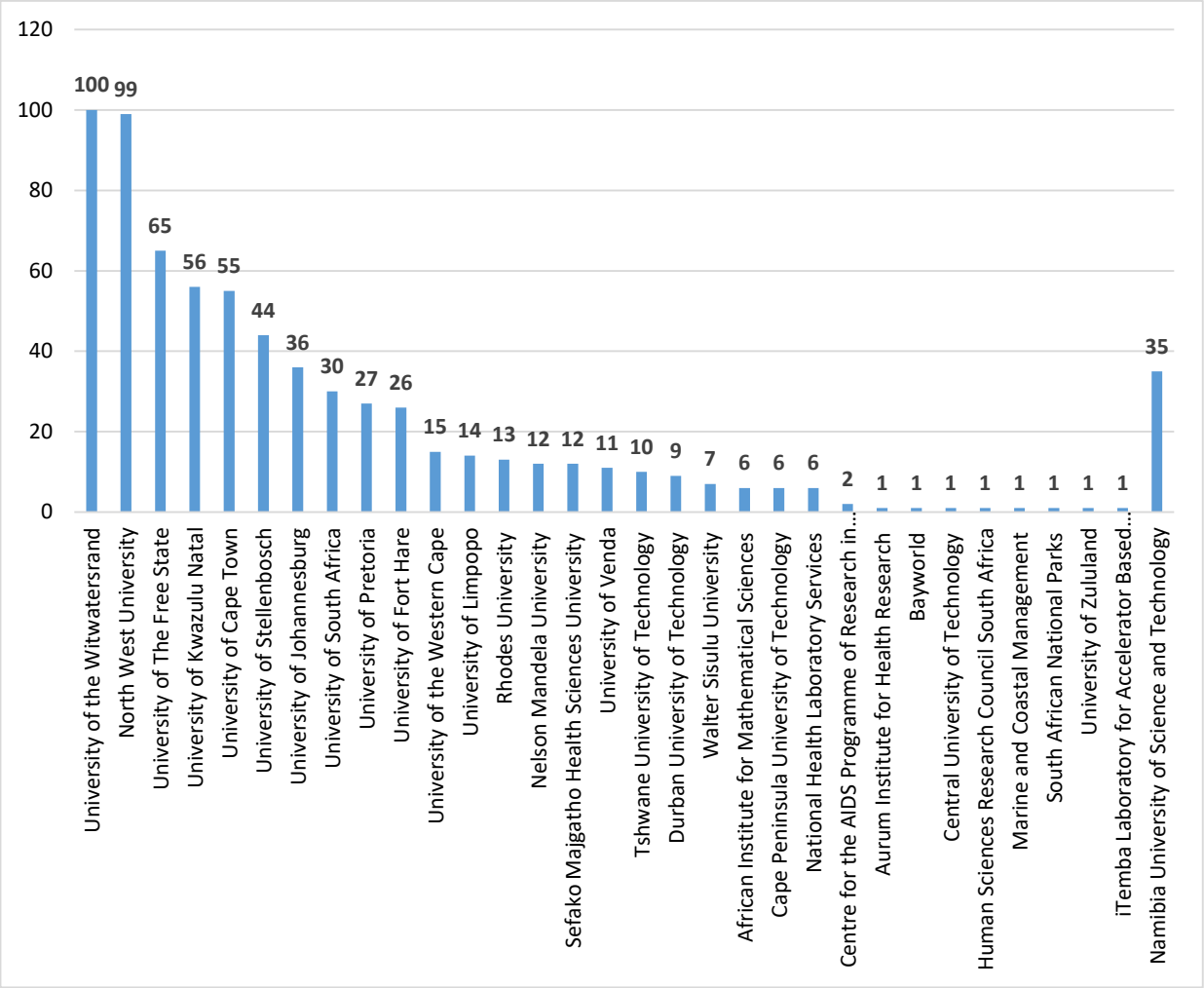


Figure 4.10 Collaboration of co-authored publications Between UNAM and South African universities

4.3.4 UNAM's top five (5) researchers per faculty

Universities are established to enable staff members to generate new knowledge through teaching and research. The number of publications output determines research productivity in a university. Publications count is one of the strategies used by experts to assess and evaluate the research performance of individual researchers, groups of researchers, disciplines, universities and countries at large (Ani & Onyancha, 2011). The current study used Scopus and Web of Science databases to establish UNAM's research productivity, visibility and impact from 2011 to 2020. Most importantly, the two databases were used to establish UNAM's top five (5) authors per faculty based on publication productivity.

Data advances that the top five (5) researchers from the Faculty of Agriculture and Natural Resources are mostly indexed in Scopus as compared to Web of Science. However, within the faculty, the top indexed researcher, Mushonga with thirty-seven (37) publications from veterinary science has also 6 publications captured in the Web of Science. As presented through Figure 4.11, other researchers' publications range between 16 and 21 respectively. In addition to the top five (5) shown, other researchers including Samkange, Von Bach, Fabiano, Madzingira and Iitembu made the top ten (10) with a fair number of publications ranging between 11 and 14.

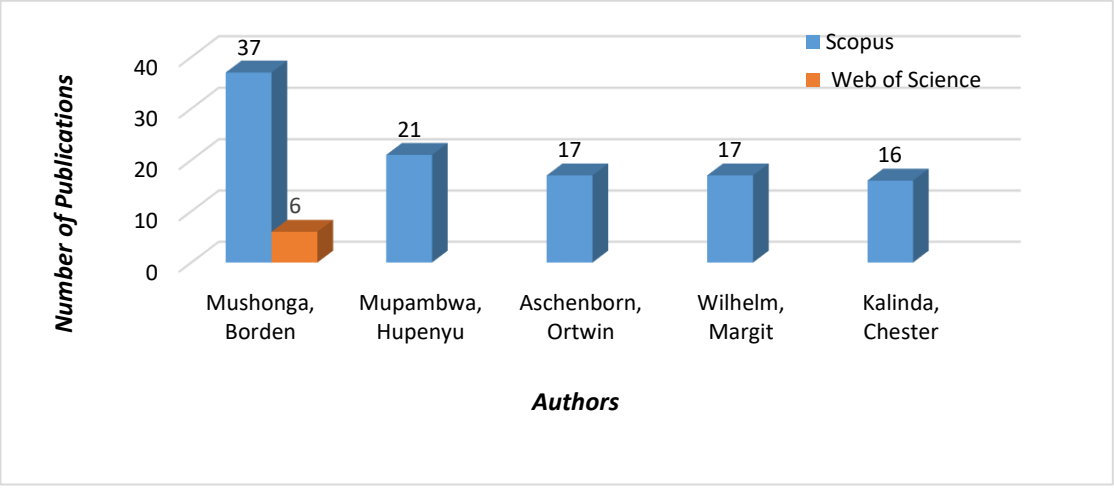


Figure 4.11: Top five authors from the Faculty of Agriculture and Natural Resources

Figure 4.12 presents the top five (5) researchers from the Faculty of Economics and Management Sciences. Data posits that the top five researchers within this faculty are only indexed in Scopus and there is no visibility in Web of Science. Ziramba from Economics is the leading researcher followed by Makanyeza, Mukong, Amadhila and Amupanda with four publications. Researchers from the Faculty of Economics and Management Sciences who made the top ten (10) include Blaauw, Mukong, Marenga, Kaapama and Chata.

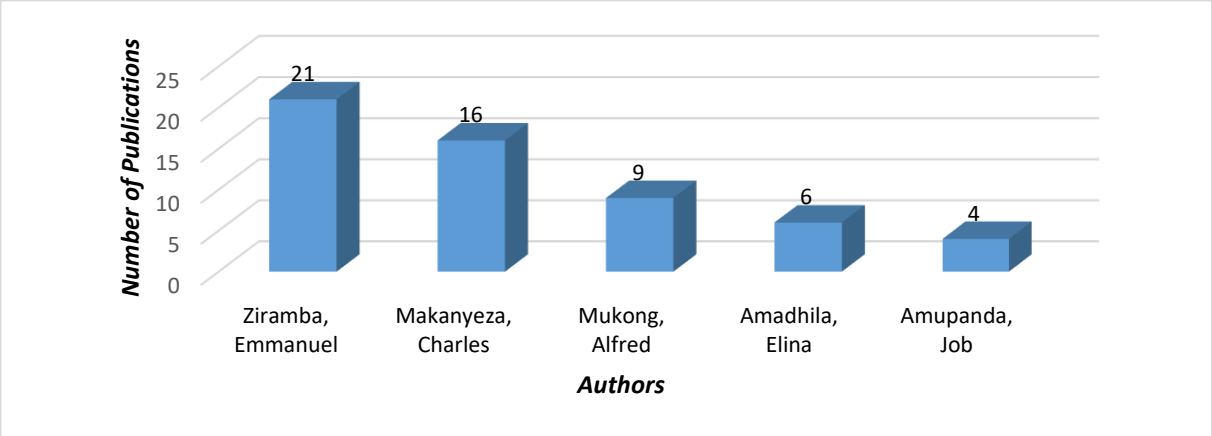


Figure 4.12 Top five authors from the Faculty of Economics and Management Sciences

Data show that the top five (5) researchers from the Faculty of Education are all indexed in Scopus. No visibility is recorded in Web of Science as illustrated by figure 4.13 Other researchers who made the top (10) include Haihambo, Zulu, Abah, Albin and Anyolo.

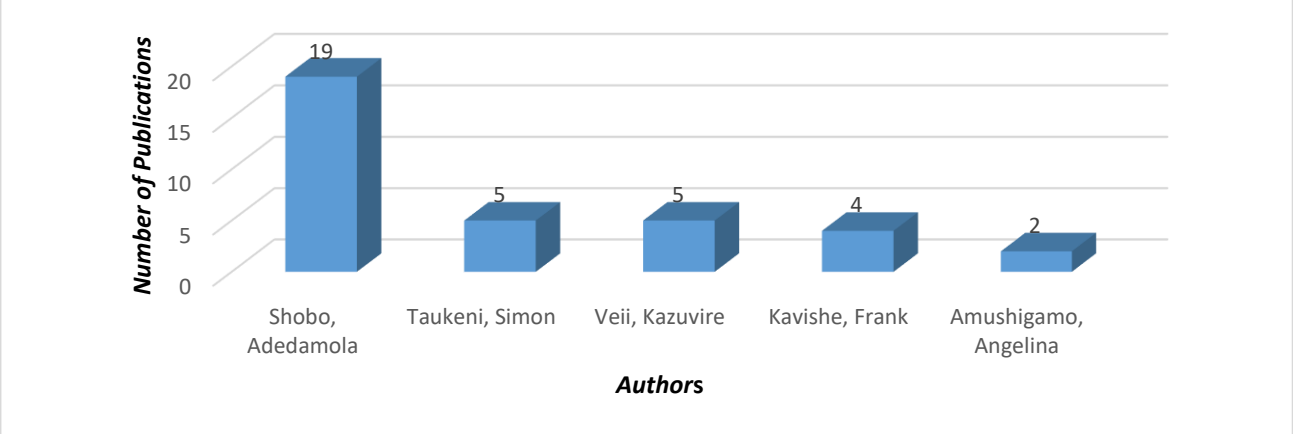


Figure: 4.13 top five authors from the Faculty of Education

Figure 4.14 shows the publications productivity of the top five (5) researchers from the Faculty of Engineering and Information Technology. Data depicts that the top five (5) researchers from this faculty are only indexed in Scopus and there is no visibility in Web of Science. Oluwagbenga is the leading researcher with 43 publications. Respectively, Imoru, Wanjekeche, Sverdlik, Lengricht and Suresh are the other researchers who made up the top ten whose research productivity is also only captured in Scopus.

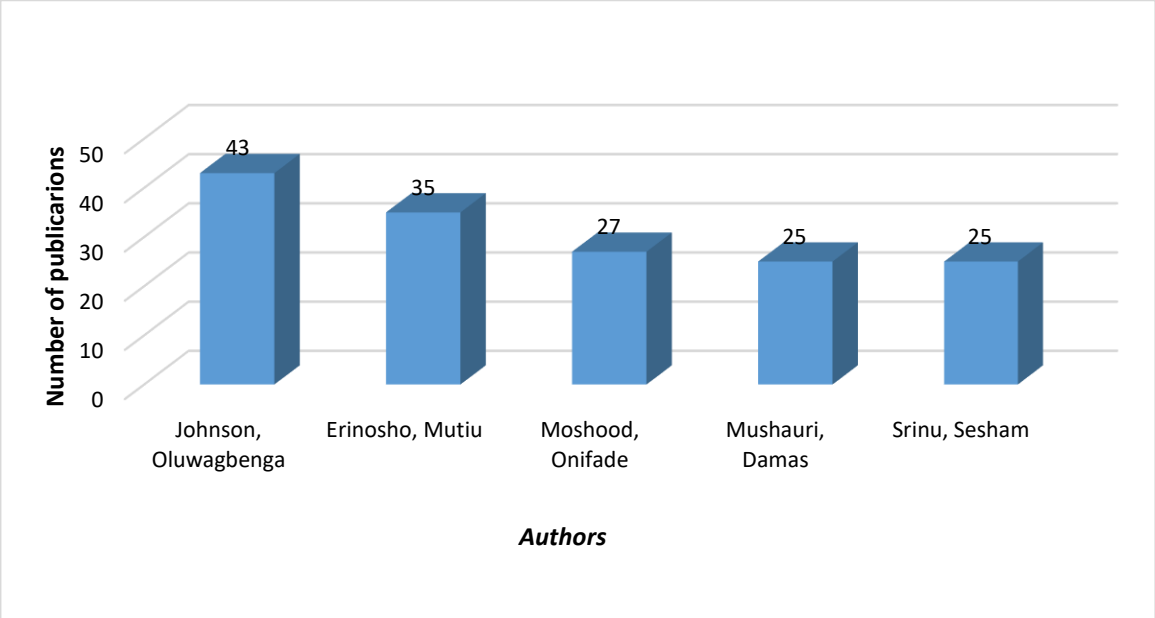


Figure 4.14: Top five authors from the Faculty of Engineering and Information Technology

Data from the Faculty of Humanities and Social Sciences advances research productivity for the top (5) researchers. Nhemachena from Sociology is the leading researcher as shown in figure 4.15. Although research productivity for this faculty is mostly indexed in Scopus, Nengomasha who appears second (2nd) also has one (1) appearing in Web of Science. The other five researchers who make the top ten are: Gentz, Uutoni, Abankwah, Mlambo and Hlatywayo.[2]

[2] From the Humanities, Kalusopa is also indexed as one of the top researchers. But because he is already involved in the current study, he was removed to avoid conflict of interest.

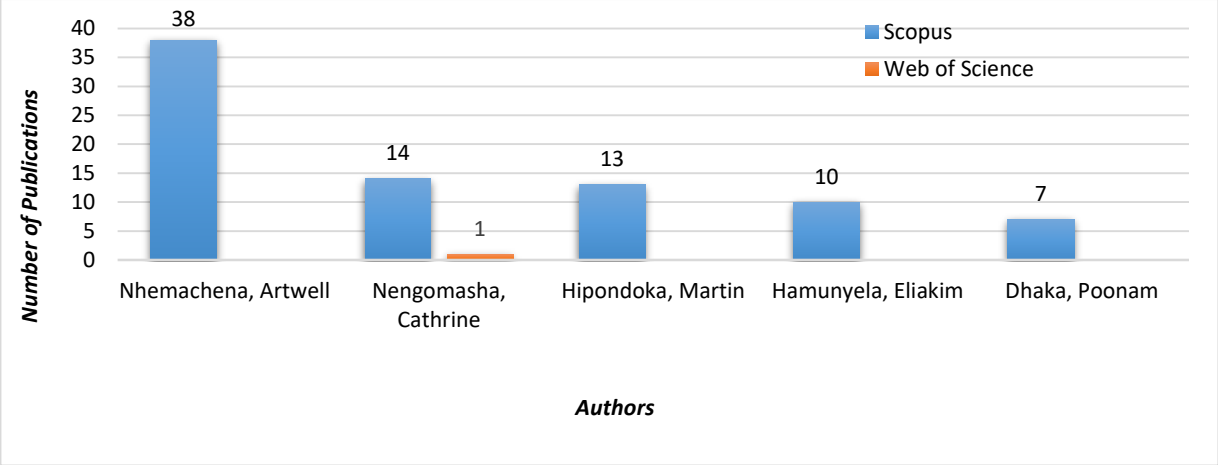


Figure 4.15: Top authors from the Faculty of Humanities and Social Sciences

Data for the Faculty of Health Science shows that the top (5) researchers are all indexed in Scopus as illustrated in figure 4.17. Research on Health, Pharmacology and Therapeutics has more visibility as compared to Human, Biological and Translational Medical Sciences. In addition, Misihairabgwi, Mwangana and Mitonga are some of the researchers with a great number of publications as indexed by Scopus.

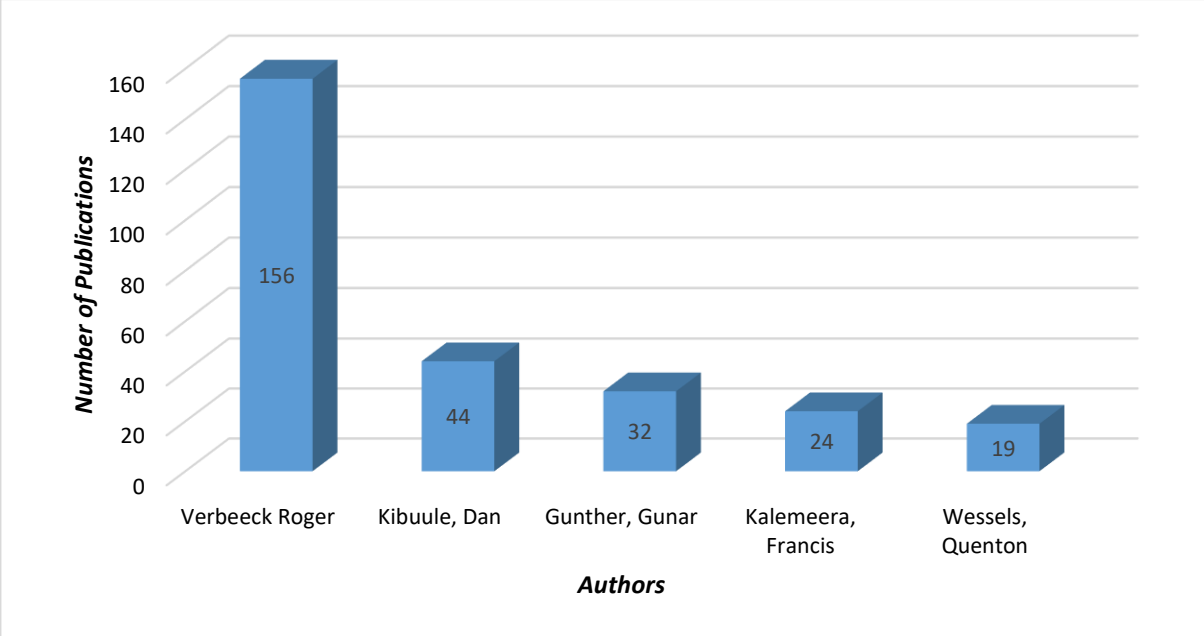


Figure 4 16: Top five authors from the Faculty of Health Sciences

Figure 4.17 shows the top five (5) authors from the Faculty of Law. Data extracted depicts the top five (5) authors indexed in both Scopus and Web of Science. As illustrated, the top two authors also have publications indexed in Web of Science respectively.

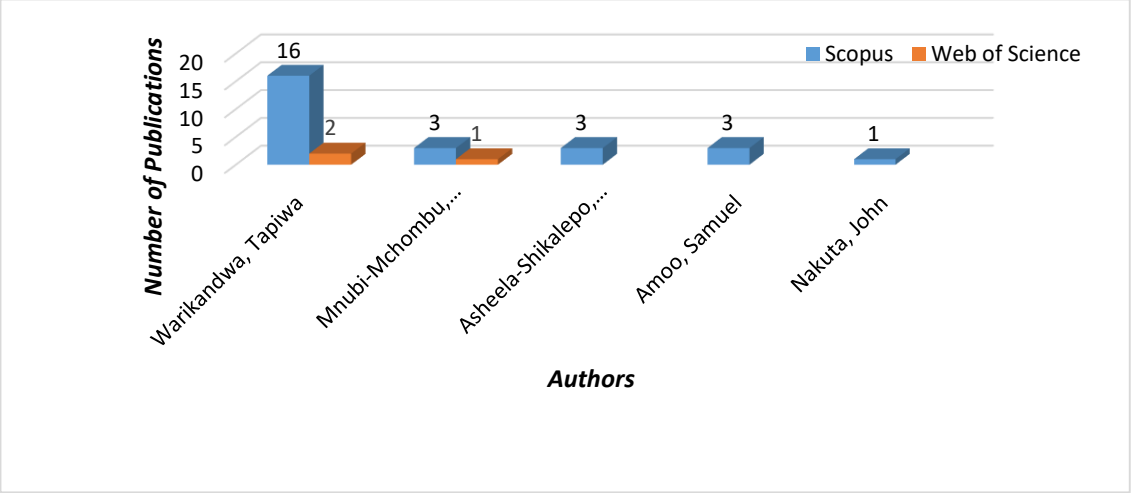


Figure 4.17 Top authors from the Faculty of Law

For the Faculty of Science, data depicts that all top five (5) researcher's research productivity and visibility is only indexed in Scopus. Steenkamp from Physics, Chemistry and Material Science is the leading researcher within the faculty as illustrated by figure 4.18 Other researchers who had more publications included: Onyemaechi, Sibanda, Mapaure, Shiningayamwe and Rahman.

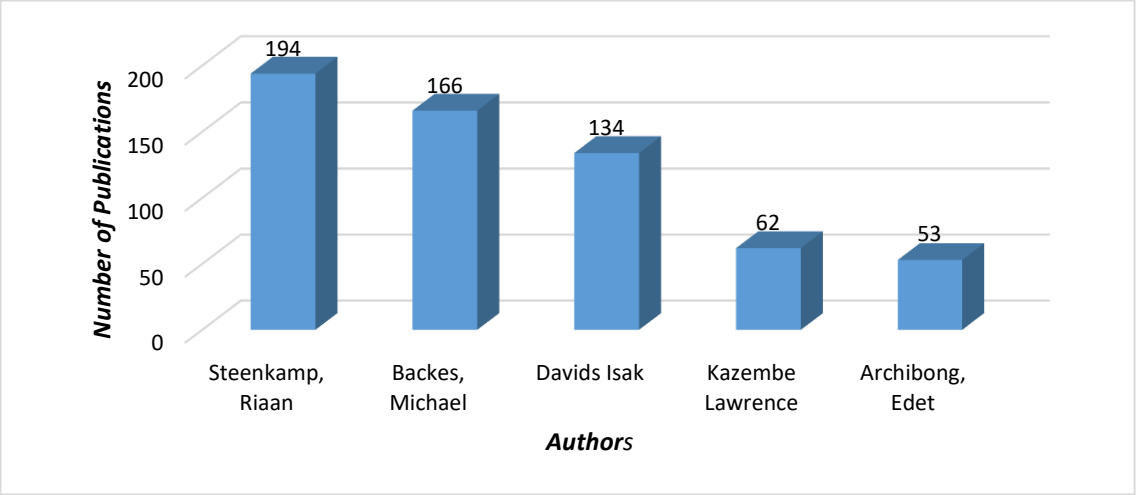


Figure 4.18 Top five authors from the Faculty of Science

4.5 Presentation of qualitative data (Phase two)

4.5.1 Introduction

A semi-structured interview guide was used to collect data originating from the themes of the study. A sample of UNAM researchers indexed in Scopus (SciVal) and Web of Science was used as the research population.

A total of 22 participants from the target sample of 40 researchers participated in the study. The study achieved only about 52% of representativeness of the whole target sample. This was due to the fact that the qualitative data was collected in June, July, August and September 2021. During this period, COVID-19 was extreme in Windhoek and specifically at the University of Namibia. Most researchers were working from home and a number of them were recovering from the deadly pandemic. Data collection was discontinued because the study reached a point whereby it became counter-productive as more interviews could not yield new information that could add value to the scope of the study. Data saturation in qualitative research is commonly used to determine and assess the adequacy of data for a purposive sample. Saturation in research is reached when issues in the collection of data starts to repeat themselves making data collection redundant (Hennink et al., 2019).

Table 4.2 Number of interviewees

Faculties	**Agriculture & Natural Resources**	**Economics & Management Science**	**Education**	**Engineering & Information Technology**	**Health Sciences**	**Humanities & Social Sciences**	**Law**	**Science**
Number of participants	4	0	1	3	4	5	1	4

Total number of participants N=22

4.5.2 Participants' profile

The study sought to collect data from UNAM's top researchers indexed in Scopus (SciVal) and Web of Science databases. Research permission was granted to collect data from all UNAM top researchers. The study interviewed the top five (5) researchers from each faculty with more research productivity, visibility and impact. For ethical considerations, no personal details were mentioned in the qualitative phase.

4.5.3 List of participants

All these participants were a representation of RDM in their respective faculties at UNAM. The total number of participants for the current study were 22 and they were coded R1 to R22

4.6 Qualitative data presentation

Interviews conducted during the second qualitative phase were grouped into themes that emanated from the research objectives. The following research questions originated from the objectives of the study were used:

- Are you aware of RDM and data creation?
- Do you undertake data processing and analysis?
- Do you consider data archiving and have you ever preserved your research data?
- Do you support data sharing and re-use by other scholars?
- What appropriate RDM framework can be developed at UNAM to ensure effective management of research data?

4.6.1 Awareness and data creation

The first objective of the study was to ascertain the awareness of and determine research data output created at UNAM.

The level of awareness and data creation are presented in the following sub topics:

- Level of awareness,
- Data created during research processes,
- Data storage, and
- Metadata

4.6.1.1 Level of awareness and data creation

Globally, RDM is regarded as a practice for good research in an organisation as it allows researchers to fully understand the lifecycle of data management and documentation. Different researchers have their own practices, perceptions, attitudes, and concerns on the management and sharing of research data (Manu, 2019).

To understand UNAM researchers' awareness of RDM, the researcher asked the following question: *Are you aware of the RDM concept?*

The study revealed that 20 participants indicated diverse levels of awareness about RDM and what it entails in general. This was evidently captured from the narrations that emanated from the researchers.

For example, one participant narrated the following*:*

"Yes, I am aware of RDM even way before the whole concept became known at UNAM and I do use it as part of my research processes. The only limiting factor is that I have not seen this done at UNAM and if it exists I am yet to come across it". (R3)

This clearly shows that although RDM is not integrated into the UNAM research processes, researchers have adopted the concept at an individual level and they have been practicing in it for some time.

Another participant reported that:

"I have been practicing RDM for over two decades now, but only at an individual level since I have not heard about this concept at UNAM. We used to do it at the university where I came from before joining UNAM and I thought it was a good idea to continue doing it". (R4)

Depending on previous experiences and how different researchers were introduced to RDM, some researchers saw it fit to continue adopting RDM as part of their research processes.

Moreover, another participant recounted that:

"I heard about it but decided to ignore it because I do not support keeping my data in those so called cloud services. I am just sceptical about the cloud. Are you sure we are not running the risk of losing everything in future? I would rather remain old fashioned". (R22)

Contingent to RDM, each researcher developed his/her own modus operandi on how they deal with the concept, which among other things, includes concepts such as data security and the safety of cloud computing.

Research data can be generated by humans, machines or through a combination of both humans, machines, and can be generated and stored using structured and unstructured formats (Ghotkar & Rokde, 2016).

In order to gather information on data creation by UNAM researchers, the researcher asked the following question: *"What type of research data do you create and formats"?*

The study revealed that all 22 participants stated that they create research data using different formats based on the type of research methodologies that they adopt for their research.

For example, one participant had this to say:

"*I create both quantitative and qualitative research data. If I employ the quantitative research method, the data format consists of statistical analysis using surveys, polls or questionnaires sent across to a specific section of a population. This usually applies to corporate governance matters, trade law statistics and any related commercial and human rights issues. If it pertains to quantitative research method, I collect data through methods of observations, one-to-one interviews, conducting focus groups, and similar methods". (R15)*

Based on whichever research methodology adopted by researchers at UNAM, different data is created using different research tools. It implies the types of research methodology that the particular researcher is comfortable with in order to answer the objectives of that particular study.

Contrary, another participant indicated that:

"I focus mostly on quantitative data through surveys from households based on different settings such as data pertaining to food security, environment and it is mostly statistical data". (R5)

Depending on the type of research that the researcher prefers, data created in terms of formats is defined by the methodology.

*4.6.*1.2 Data storage

The storage of data is one of the first most considered aspect as it allows the researcher to quickly access data. The medium of storage plays an important role in ensuring access. Storing data on a local machine can be advantageous but presents obstacles when attempting to share such data. It requires that only the owner of the machine can be fully

responsible for its preservation (Blumzon & Panescu, 2019). In order to garner information on data storage techniques from UNAM researchers, the researcher asked the following question: "*How do you store your research data?"*

All 22 participants indicated that they store the research data using different mediums or platforms at their disposal. Responses to this question had diverse answers based on the researcher's level of awareness and exposure to different storage devices and platforms. When asked about how researchers store their data, one participant had this to say:

"I store my data using external hard drives, personal emails, Google drive and office space especially for all the print files". (R12)

This clearly indicates that researchers are well aware of different storage platforms and technologies that are available and are at their disposal to store research data.

Apart from some researchers having been exposed to different platforms of storing data, there are other researchers who are not exposed to such platforms and their only way of storing data is through emails and laptops. This has risk implications as indicated by one participant:

"I store all the data I create in personal folders on my laptop. When this computer breaks down, that will be the end of most of my data. I also store some of my data on my emails and I always revert back to my previous emails whenever I need such data. I know it is risky storing data like this due to the availability of other platforms, but I do not have the time to explore them". (R19)

The level of exposure, know-how, past experiences and collaborations have proven to be critical to the researcher's level of connection on how they store their research data. Based on the importance of RDM to certain researchers, they continue utilising data storage

platforms from their previous employment as a means to continue practicing RDM. This narrative is evident from the participant who narrated that:

"I belong to a large research group in South Africa and I store my research data with their cloud services. I also store some data using external disks and on the computer for back-up purposes. The data I store on the computer is actually the one to work on during that specific time. When I am done analysing that data, I transfer it to the cloud for long-term storage and in case, I might need to refer to it to address another phenomenon using the same data. This has worked for me for many years and I hope to continue using this practice for as long as I am able to continue conducting research". Unfortunately, UNAM does not have RDM services and fortunately for me I can continue using South African platforms from my previous employers" (R15)

4.6.1.3 Metadata

The term metadata is commonly known as data that is associated with either an information resource or system for purposes of description, use, retrieval and preservation (Dashrath, 2014). The author sought to get information on whether UNAM researchers attach metadata to their research data by asking the following question: *Do you attach metadata to your research data?*

Fifteen participants confirmed the importance of metadata to information resources or research data to enhance access and retrieval.

In response to the question on metadata, one participant narrated that:

"I do attach metadata to my research data because it makes it easier for me to get it with ease sometime down the line. The way I define my metadata includes the year, type of data and there is an ID specifying the date of collection and progress". (R8)

Although metadata is crucial to the retrieval of research data, in the absence of clear metadata descriptions at UNAM, researchers have come up with their own ways of describing their research data based on methods and techniques that they know and are comfortable with. This was narrated by one participant indicating the following:

"Oh yeah, I do that with all my work. For example, storage location of the data (including identifier information) Origin of the data (creation/acquisition of the data) Time references for the data (key dates associated with the data: start, end, release)". (R19)

The above narrative clearly demonstrations that researchers understand the need for metadata and why it is important. Although, defined based on how the researcher sees it fit, it works and solves the researcher's problem when searching for needed data.

Although some researchers demonstrated the importance of metadata and the techniques used, some researchers have overlooked the importance of metadata when filling their research data. This has presented them with consequences of data retrieval and easy access. This narrative is demonstrated from the participant below:

''I store my files in boxes in my office. Unfortunately, the boxes are not labelled or defined in any way. I wonder how I will get those files easily because I do not even know where to start. I have my PHD data in one of those boxes which I need to use but the thought of having to go through those boxes is so demoralising". (R.22)

4.6.2 Data processing and analysis

The second objective of the study sought to assess if researchers at UNAM undertake data processing and analysis during their research processes. Consequently, this objective discussed whether UNAM researchers undertake data cleansing and validation; data anonymisation, data interpretation and analysis and whether they publish their research

data. Data analysis and processing plays an important role in data generated through various sources. Data cleaning is to prepare data for processing in an organised manner. It deals with fixing inconsistencies within data and handling massive values and includes various steps such as data qualities, identification and analysis (Rahul et al., 2020, p. 386).

In order to gather significant information on data processing and analysis, the researcher asked the following question: *Do you undertake data cleansing and validation?*

All 22 participants specified that they undertake data cleansing and validation as part of their research process. They revealed that for a researcher, data cleansing and validation are critical to determining the research outputs and one of the process that a researcher is not able to overlook. Participant 3 stated that:

"Yes, I do clean my data after I am done collecting it. This actually helps me end up with the data that I need to use for analysis because not all collected data is relevant". (R. 3)

During data collection, researchers collect many data, and the cleansing and validation of such data ensures that researchers have data that is able to answer the objectives of the study. In addition, the credibility and quality of data is only visible after cleansing and validation. The processes are termed to be time consuming and they require detailed attention and it is not possible to overlook any detail as indicated by the below participant:

"Most definitely, when you want to deal with quality and accurate data then one has to make sure that you clean and validate what you have collected". (R.12)

4.6.2.1 Data anonymisation

Anonymity in research is a form of confidentiality, which includes keeping the participants' identities a secret. Data anonymisation ensures maximising the protection of participants' identities and maintaining the value and integrity of the data. Anonymisation includes

hiding aspects such as people's names, places, religious and cultural background, occupation, family relations and other potentially identifying information (Saunders et al., 2014, p. 620). The present researcher sought to get information on whether UNAM researchers undertake data anonymisation by querying the following: *Have you ever done data anonymization of your research data?*

All 22 participants emphasised the importance of data anonymity in research. Research participants are key to the research outputs, and as such, their identity should be protected when analysing research in order to protect them and to ensure that research that involves human objects is not jeopardised. This was clearly evident from the participant who stated that:

"I hope you have thought about this concept and you will not reveal our names in your research findings. Anyway, yes I do pay close attention to data anonymization because you would not want to risk your participants by revealing their identities". (R.20)

When anonymity is not addressed in research, it results in many consequences especially pertaining to the participants and there are many ways to deal with anonymity. This was apparent from this participant:

"*It is now a norm in research that whenever you collect data involving human objects, you should protect their identities. There are many ways to do that and I suggest I do not need to mention them to you since you are a master's students. Data anonymity is very crucial because otherwise UNAM will terminate our contracts. Anyway, yes I do data anonymity in my research processes and it works well when you apply different coding styles*". *(R3)*

4.6.2.2 Data interpretation and analysis

The study sought to obtain information on whether UNAM researchers undertake data interpretation and analysis as part of their research processes. Twenty-two participants

specified that they fully practice data interpretation and analysis. This helps researchers to make meaning from the data that they have collected. Researchers were asked the following question: *Have you ever done data interpretation and analysis?*

For example, one participant indicated that:

"This is actually the mayor reason why we do research. After stressing with collecting data having to manage between teaching and research. This is the final part whereby you get answers to your questions". (R.8)

The research process does not have meaning without data interpretation and analysis as collected data needs to be interpreted and analysed to answer or provide tangible evidence. Although time consuming, this part of research does not require one to run up and down after one has the required data.

In the research process, data interpretation and analysis inform the findings of the study and it is important that this process is well understood by researchers.

In addition, another participant reported that:

"The research process forces you to do data interpretation and analysis. There is no way you can overlook it because otherwise it is as good as not having done research at all. Although time consuming and focus oriented, I do enjoy it a lot because after I am done with this process then the publication is almost due for publishing". (R6)

4.6.2.3 Data publishing

Data publishing is a process of making data publicly available for use by other researchers in accordance with associated policies, while guaranteeing that data provenance and attribution are tracked. Data publishing is multifaceted and occurs at different stages of the research process and it comprises of controlled sharing and openness (Candela et al., 2017,

p. 73). The researcher sought to attain information on whether UNAM researchers are publishing their research data as part of their research processes. To obtain information on data publishing, researchers were asked the following question: *Do you publish your research data?*

Thirteen participants indicated that they publish their research. It was revealed that among those who shared their data, they did so as it was part of the publisher's requirements or collaboration with other researchers across the globe. The 13 researchers alluded that they do not share their data, some researchers indicated that they support data sharing as it jeopardises their academic freedom. Other researchers were sceptical about sharing their research data because other researchers would publish publications using their data.

For example, one participant had this to say:

"I publish my data through Data in Brief which is under Elsevier and also through Mendeley. Publishing my data serves as a way of storing my data and also making it available to other researchers." (R15)

Researchers who do not support data sharing had reasons as to why they were sceptical about sharing their data. One participant stated that:

"I am sceptical about data sharing, not that I do not support it but there are so many questions that ought to be addressed before I can share my research data. Look, we collect data using our resources and sometimes it is a very costly process. So far I have not considered sharing it and I will only stress about this when I happen to come across on day". (R5)

Emanating from the narrative above, it is clear that based on the researcher's level of RDM awareness and exposure, researchers pronounce their position on how they deal or view aspects of data sharing.

4.6.3 Data preservation

The third objective of the study sought to obtain information on data preservation strategies adopted and used by UNAM researchers. Consequently, this objective discussed whether UNAM researchers consider data archiving; if they have ever considered preserving their research data; and if they consider long-term data preservation and whether they link their data to research outputs.

Data preservation ensures that digital data preserved today remains accessible over time. Massive data emanate from research and such data needs to remain readable and interpreted through changing mediums and platforms (Rahmanto & Riasentiawan, 2018). To gather information on data preservation, the researcher asked the following question: *Do you consider data archiving?*

Twenty-two participants stressed the importance of data archiving especially in the digital era whereby there are vast platforms and technologies to rely upon. For example, one participant narrated that:

"*Yes I do consider data archiving a lot. So far, I have created multiple Google accounts that I use to preserve my data on the Drive. It works for me because it makes my data mobile and accessible from wherever. Situations on the ground have forced us to think about our own alternatives ton how to preserve our research data". (R7)*

Another participant reported that:

"Yes I have always been considering data archiving of my research work. So far, I keep migrating those data from one laptop to the next. It has been working for me, but it is not one of the safest way at all. I would not really call this data archiving, but perhaps desperate measures to ensure that I do not lose my data" in the absence of proper platforms". (R16)

4.6.3.1 Long-term data preservation

With immense production in digitised documents as well as those born digital, more efforts are required to preserve data for next generations. Long-term preservation of digital data involves processes and tools to ensure the availability and readability of the data over a long period. The process covers looking at significant changes in technology and data formats. Over time, many factors such as deterioration, obsolescence, and storage media can threaten the integrity of data and long-term data preservation address all these factors (Kremser et al., 2012). The researcher sought to attain information on whether UNAM researchers consider long-term data preservation during their research processes by probing this question: *Do you consider long-term data preservation?*

Twenty-two researchers stressed about the importance of long-term preservation to research data. Data to be preserved for long-term access should be clearly defined due to storage capacities that institutions are able to offer.

The implication at UNAM is that there is no RDM policy to guide the type of data to preserve over a short period of time as well as long-term preservation. Researchers who preserve their data on a long-term practice do it based on available platforms and some by subscribing to cloud services.

One participant indicated:

"I do consider long-term data archiving for as long as the system can be trustworthy and no one will steal such data. Data is very important and I personally do not support any ideas about destroying data. My only concerns are about who gets access to the data. Strick measures have to be put in place in order to protect such data and access should be based on consent from wherever that data came from". (R22)

In addition, another participant highlighted that:

"The idea of long term preservation of research data is very important, but it should be clearly stipulated as to which data one should preserve for long term. Some disciplines are generating massive data sets over a short period; one will run out of space. There is a need for some type of document to guide researchers when it comes to this part of long-term preservation". (R3)

In the absence of an RDM policy at UNAM, there is nothing in place to indicate which data to keep, delete, when and how. This makes it difficult for researchers who are not well exposed to technologies that are available and support data preservation.

4.6.3.2 Research data links with research outputs

Open access to research results and reproduction are visible through publisher policies that encourage authors to provide data availability statements. As a result, many data availability statements have surfaced in recent publications (Colavizza et al., 2020). Eight researchers revealed that they link their data to research outputs, especially when it is part of donor requirements or publisher requirements. This implies that researchers have only practised this practice when mandated by external forces and not necessarily willingly. Among those who indicated not linking their data indicated that either they are sceptical about sharing data or they have not come across such requirements. The researcher sought to get insights on whether UNAM researchers link their raw research data to their research outputs by asking the following question: *Do you link your data to research outputs?*

"In the case of our department, the codes used during the analysis are published as part of the appendix and anyone can use those codes to verify, validate and check authenticity. We have learned this from some examiners who needed to verify if there is a correlation between the data collected and the presented results".

Depending on expertise and disciplines, researchers demand for research data to be linked with the research outputs to ensure others to validate their research results. Other researchers have adopted linking their research data to their research outputs as narrated by one researcher:

"Some publishers require that you submit your research data as part of the final document. I have done this with several of my publications and it is a good norm because you cannot mess with your findings". (R.4)

The willingness to link research data to research outputs is more external and upon requirements. Although some researchers link their publications when requested to do so, they are equally comfortable with the practice if they are not required to do so as narrated by one researcher:

"Yes, I do this with most of my publications, but it is dependent on the publisher's requirements. Some publishers have it as part of their requirements while others do not really mind. Where it was requested I have shared the data and where it was not requested I did not bother". (R.21)

4.6.4 Data accessibility and re-use

The fourth objective of the study sought to obtain information on data accessibility and re-use. Therefore, this objective looked into whether UNAM staff support data sharing and re-use by other scholars. The open access to research promotes data sharing and open data, which entails reproducing research, and making it available in the public domain. In order to achieve the benefits of data sharing, data is supposed to be re-used by other researchers (Pasquetto et al., 2017, p. 1). The researcher sought to acquire information on data accessibility and re-use from UNAM researchers by asking the following question: *Do you support data sharing and re-use by other scholars?*

Seven researchers indicated that data sharing has been around for a while and depending on exposure, external collaboration and understanding of RDM, it positions researchers to either share or hold on to their research data. For example, one participant stated that:

"I advocate for open access to research data, especially after we have published our articles. Let us allow others to verify our findings because we have nothing to lose since the publication is already in the open. Those who are not willing to share have either not been exposed to the concept or it is just fear of the unknown that they can also publish their data or they are afraid that others my question that there is no correlation between the data collected and the findings. If one if confident of what they have published, then there is no way you can hide your data". (R.15)

Some researchers stressed that it is good for research to be validated by other experts in the discipline. If one's research is authentic and there is a correlation between the data and the findings, then it should not be a problem sharing research data.

This narrative was confirmed by the response below:

"I strongly belief in the openness of research data and truly support data sharing with other scholars. Anybody can use it. If I publish my simulation codes, anybody can use them. I had question to share data in the past and I have done that. I also benefited from other researcher's data". (R2)

Data collection is time consuming and costly, whereas advocating for data sharing and re-use also helps with ensuring that others can address new areas using data collected by other researchers. For example, one participant depicted the following:

"Yes, this allows one to avoid unnecessary duplication of efforts to generate research data. It allows me to integrate data from different studies, sites, and disciplines

and thus to open up important new avenues of research. It further eases the burden on over-researched populations". (R10)

Researchers are the primary tool in collecting, organising, and sharing data. They hold their data rights, and sharing data for others largely depends on their will and attitude (Rafig & Ameen, 2022, p. 392).

4.6.5 RDM challenges and framework

The fifth and final objective of the study sought to investigate possible RDM challenges that UNAM staff encounter during their research processes and at the same time to propose an RDM framework that can address RDM challenges at UNAM. Consequently, the researcher sought to acquire information from UNAM researchers on whether there are any RDM challenges that they encounter at UNAM and a recommendation for an appropriate RDM framework.

Five researchers revealed that new developments in research can only be a success if integrated into organisational (UNAM) research processes, especially through a policy that all researchers can abide to. To acquire information on RDM challenges, the researcher posted the following question: *What challenges do you encounter in reference to RDM?*

One participant stated that:

"*Coming from an institution where data management was part of the organizational structure, I had to find means to deal with my own RDM issues at UNAM. I was a very big challenge when I started with UNAM, but since I was exposed to the concept already and I had to survive on my own. This is one major aspect that UNAM should look into and I am glad you are carrying such a study". (R15)*

The absence of an RDM policy at UNAM is the limiting block on the development and implementation of RDM practices at UNAM. Researchers at UNAM are left to practice

RDM using their own means, which is not supposed to be the case. Respondents revealed that it is totally not yet late for UNAM to consider RDM as indicated by the participant below:

"Although RDM remains a big challenge, it seems as if the institution is moving into the right direction and the hope is that RDM will be implemented just like in other pioneer institutions in the region. Not having RDM platforms at UNAM is a big challenge and a shift is needed because electronic data cannot be preserved under lock and key. Data used should be made available somewhere. There should be a policy in place to ensure that data can be made available". (R.7)

4.6.5.1 RDM framework

The researcher wanted to know if there was a need for an appropriate RDM framework at UNAM by asking the following question: ***What appropriate RDM framework can be developed at UNAM to ensure the effective management of research data in the whole research life cycle?***

Twenty participants stressed that it is of critical importance that UNAM develops an RDM policy that would be able to guide RDM practices and ensure that UNAM is at the same level as other universities especially in the region. For example, one participant stated the following:

"*Any RDM framework that takes cognisance of the following issues:*

- *Research Data management framework that supports the evolving global data-intensive research environment.*
- *Research Data management framework that will support all aspects of the data lifecycle, through creation and collection, storage, manipulation, sharing and collaboration, publishing, archiving and reuse.*

- *Research Data management framework that promotes the effective data management through teamwork and collaboration between researchers, research offices, information specialists and technical support staff. Such an RDM framework must constitute of the following four elements:*

- *Institutional policy and procedures;*
- *Information Technology Infrastructure: the hardware, software and other facilities which underpin data-related activities;*
- *Support services: people and other means of providing advice and support, such as online toolkits and research data interviews; and*
- *Metadata management: so that data records can be used for both internal and external purposes. Examples of such plausible RDM which UNAM could learn from in developing a comprehensive RDM framework are:*

1. *The Australian Code for the Responsible Conduct of Research; and*
2. *The United Kingdom Community Capability Model for Data-Intensive Research. (R17)*

It is evidently clear from the respondents that a number of researchers have integrated RDM as part of their research processes and that they are well acquainted with what is happening and the requirements thereof.

Another participant indicated that:

"*Very good idea to have an RDM framework because some of the data may be needed beyond UNAM and making it available will increase the visibility of UNAM and the researchers. If UNAM provides RDM services, researchers across the globe will be conducting UNAM inquiring for data housed under UNAM. Statistics should be kept and it will be a good thing as it will expose the academic community beyond Namibia". (R.5)*

In addition, another participant narrated that:

"UNAM should bridge the RDM gap presently available. An RDM framework taking all aspects (technology & available software's) into consideration will help solve this problem. Currently I am preserving my data through my own personal means and if I am to leave UNAM tomorrow, I will take along my data which partially also belongs to UNAM and any new expert taking over will have to start from scratch which is not supposed to be the case in a university setup". (R1)

The visibility of academic institutions across the globe is not only limited to the research outputs, but also the data the researchers collect, analyse, preserve and link to their research outputs for other researchers to validate or consult the data to provide solutions to global phenomenon using already existing data.

4.14 Chapter summary

This chapter has presented research data from the bibliometric analysis and semi-structured interviews organised according to the thematic areas of the study. The analysis shows that UNAM research productivity, visibility and impact have been on the increase between 2011 and 2020. Although UNAM is a relatively small university in the SADC, the university is doing well as compared to other small universities in the region. The bibliometric analysis also managed to confirm UNAM's top publishers indexed in Scopus (SciVal) and Web of Science per each of the eight faculties at UNAM. The findings from the qualitative data revealed that UNAM researchers are aware of RDM and they have integrated RDM as part of their research processes. Research data is guided by a number of processes in the data lifecycle. Different processes require different levels of understanding and action. Depending on the researcher, each of the researchers had their own ways of handling research data at each phase of the data lifecycle. Furthermore, the findings revealed that there is no institutional RDM support at UNAM. The next chapter discusses the findings of the study based on the research objectives by incorporating the literature.

CHAPTER FIVE

DISCUSSION AND INTERPRETATION OF FINDINGS

5.1 Introduction

The previous chapter analysed and presented the results of data obtained from bibliometric analysis in Scopus (Scival), Web Science and interviews. The present chapter provides the interpretation and discussion of findings of the results. The interpretation and discussion are based on conceptual and empirical information as discussed in Chapter two. Luthuli (2021, p. 121) proffers that the interpretation of the results necessitates the researcher interpreting insinuations from the results or the research questions, hypotheses, and the broader meaning of the obtained results. The interpretation and discussions are deliberated based on thematic areas derived from the objectives of the study and in relation to the UK data archive lifecycle model as follows:

- Are you aware of RDM and data creation?
- Do you undertake data processing and analysis?
- Do you consider data archiving and have you ever preserved your research data?
- Do you support data sharing and re-use by other scholars?
- What appropriate RDM framework can be developed at UNAM to ensure effective management of research data?

5.2 Awareness and data creation

The level of awareness and data creation are presented using the following sub topics:

- Level of awareness,
- Data created during research processes,
- Data storage, and
- Metadata.

5.2.1 Level of awareness

The level of RDM awareness and its significance plays a critical role in how those presented with it respond to its development and implementation (Higman & Pinfield, 2015). In Namibia, RDM is a relatively new concept. By the time of this research, only two publications were written about RDM in Namibia, specifically at UNAM. The present study acknowledges that in order for RDM to succeed, there should be some level of awareness of the procedure, policies and standards governing RDM. The findings showed that many researchers at UNAM were aware of RDM either through studies with other universities, previous experiences/ affiliation and through donors' or publishers' RDM requirements.

Although RDM does not officially exist at UNAM, 20 researchers indicated that they were aware of RDM even though awareness and practice are two different scenarios. The implication of the findings on the level of RDM at UNAM clearly posits that the only limitation to the practice and integration of RDM as part of research processes at UNAM is entirely linked to the absence in UNAM of an institutional RDM policy.

The above was underscored by Chigwada (2021) that RDM is still at an infant stage in Sub-Saharan Africa, which contributes to the different levels of awareness. In the same vein, the data may not be a true representation for all researchers at UNAM because of the scale of the study and the number of participants. Overall, the level of RDM awareness among the participants is a clear indication that UNAM should look into integrating RDM as part of the research process, provide storage facilities that all researchers can use to store their data based on available guidelines, security, privacy and the re-use of such research data where possible.

5.2.2 Data creation during research processes

Data created through research takes a wide range of forms such as statistical, experimental, interview recordings and transcripts. RDM is at the helm of organising research data from its entry into the research cycle up until dissemination and archiving of valuable outputs (Sanjeeva, 2018). Data analysed discloses that data is created by researchers through their research processes at UNAM.

The study acknowledges that all 22 researchers are actively involved in data creation through different research tools and methodologies that they are comfortable with in addressing research objectives and to obtain relevant answers to the questions. Some of the research data created by some researchers involves big data which becomes difficult to keep as an individual unless parent institutions avail storage facilities that are able to cater for big data.

The implication of the findings reveals that researchers at UNAM create massive data through their research processes and many of the researchers are facing challenges on how to store such data.

5.2.3 Data storage

Many universities across the globe adopt RDM platforms to accommodate research data created by its scholars and they have few restrictions on data formats, standards and or scope. This is important in making sure that all sorts of raw data can be stored irrespective of the format.

The study revealed that all 22 researchers store their created research data using different platforms and the available technologies.

Techniques on storing data revealed were based on expertise and the digital divide, and some attributed them to safety and the security of their research data. This includes CDs, Dropbox, Google Drive, laptops and external drives.

Tripathiel et al. (2017) also add that researchers keep their data in computers, CDs, DVDs, and hard drives. The implication on the findings points to the fact that researchers are not supposed to be struggling through their own means to store their research data. This puts UNAM at a risk because should a researcher leave UNAM, they leave with their data which could benefit the university. This implies that UNAM should avails data storage facilities as a way to ease the burden on researchers using their own money subscribing to Dropbox and others to use external drives which are a risk.

5.2.4 Metadata

Metadata is critical to the discovery of information resources and this was evident from the data analysed that researchers attach metadata to their research data when storing data.

Farnel and Shiri (2014, p. 75) define metadata as structured information that provides context for information objects of all kinds, which include research data and enables the use, preservation and re-use of those objects. The study acknowledges that 15 researchers understand and appreciate the importance of metadata. The implication on UNAM is that those who attach metadata only do so according to methods that are known to them. This implies that should RDM be implemented at UNAM, researchers will be introduced to the available metadata standards and how to best describe data.

Empirical literature details that metadata is at the helm of providing context to different objects, including research data which facilitates the use, preservation and re-use of such data (Farnel & Shiri, 2014). In addition, another study by Tenopir et al. (2011, p. 75.) underlined that there is a lack of awareness about metadata and its importance among the

scientific fraternity and recommended that institutions and the scientific community should do more to enhance researchers to prepare metadata to support the discovery, preservation and re-use of their data.

5.3 Data processing and analysis

Discussions in Chapter four pointed out that data processing is critical to ensuring that research data is cleansed and free of errors before publishing data sets and making data to underpin research findings. Therefore, it is no surprise that researchers at UNAM undertake data processing and analysis. The findings are presented and discussed as follows:

- ✓ Data cleansing and validation,
- ✓ Data anonymisation,
- ✓ Data interpretation and analysis, and
- ✓ Data publishing

5.3.1 Data cleansing and validation

Ridzuan et al. (2019) indicate that data collected from various sources will affect the accuracy and prediction of the results if not processed and analysed. When data is cleaned, it offers better data quality which enhances its usability and interpretation. Although the process might seem as time consuming and complex, it assures verification and validation (Ridzuan et al., 2019).

All researchers stressed the importance of data cleansing and validation as part of their research process at UNAM. The findings revealed that when collecting data, a researcher is presented with all sorts of data which includes errors, as well as inaccurate and unnecessary files. The process enables researchers to deal with relevant data that is able to address the research needs of the researchers at the time.

Some data may be of value but not relevant at the time and the process ensures that the data is processed and analysed in such a manner that it can be used efficiently me to address other research needs that are closely linked to the current ongoing topic.

5.3.2 Data anonymisation

The study findings show that UNAM researchers undertake this process to deal with participants' identity. All the researchers emphasised that every research that involves human objects should by all means protect their identity.

Majeed and Sungchang (2016) portray that data anonymisation is regarded as a practical solution for preserving users identity in data publishing. Data anonymization is undertaken to protect the privacy of users. Although data sharing brings innovation and enables better decision making, it may jeopardize the privacy of users due to the existence of sensitive information in the data (Majeed & Sungchang, 2015).

Dube et al. (2014, p. 208) posit that exposing the identities of participants can be a problem particularly if the readership of a publication finds the content controversial. It puts a major challenge in maintaining a clear balance between confidentiality and autonomy. Loss of confidentiality and anonymity present implications on the validity or integrity of research.

5.3.2 Data interpretation and analysis

Boaduo (2011) posits that the interpretation of data for any study is vital to attaining meaningful information for analysis. Data obtained through research instruments should be analysed and interpreted in an acceptable manner that would make it relevant to practitioners and the public for articulation. The research findings revealed that UNAM researchers treat data interpretation and analysis as a crucial process within the research process because nothing in research will happen without the process.

Raw data is nothing when this process is not executed through the research process and the only way to get that meaning is to do data interpretation and analysis. Researchers indicated that meaning from the whole research carried out is realised when interpreting and analysing the collected data.

The process enables researchers to get answers to their research questions. Researchers at UNAM indicated that once one is done with this important process, the publication is almost done and due for publishing.

5.3.3 Data publishing

Austin et al. (2017, p. 11) describe research data publishing as the release of research data, associated metadata and all accompanying or associated information for re-use and analysis by other researchers. The findings of the study showed that the position of UNAM researchers and their views on data publishing is influenced by the level of RDM awareness, integration, policies and the benefits that they got exposed to in their past experiences.

This study acknowledges that UNAM researchers had diverse responses about publishing their research data. Many researchers had questions about privacy, acknowledgement, and integrity. There is a strong indication that data publishing is one of the most sensitive topics in the whole research process if not well addressed. Some researchers support data publishing based on how they understand the importance of the process, but the same cannot be said about those who oppose data publishing.

5.4 Data preservation

Access to research data whether short-term or long-term is vital to researchers for them to be able to have access to their data over time. Data collected for a specific purpose can be used to provide answers to different global phenomenon. Its preservation over time ensures its relevance and accessibility.

The findings are presented and discussed as follows:

- Data preservation,
- Long-term data preservation, and
- Linking data to research outputs.

5.4.1 Data preservation

The preservation of research data is important in the research processes and it encompasses a number of incentives such as meeting the obligations when an organisation has a mandate for preserving data; maximising resources for data production relating to costs; maintaining its status of data preservation; and avoiding losing data and provide ongoing access to such data. Maurya and Sabaveerapandiyan (2022, p. 259) advances that data preservation is vital in institutions of higher learning for students and researchers. Proper RDM plans, implementation and curation of research data becomes a challenge. Research data is required to be preserved safely during and after completion of research by using proper and reliable storage devices that could provide necessary security to confidential data.

Researchers viewed it as a cost and time saving technique because you may not be required to collect such data again. In reference to the preservation of data, researchers are using different mediums for preserving their research data. While some researchers have explored the use of online cloud devices, other researchers are using the traditional mediums such as memory sticks, external drives and typically on their laptops or work stations and some through printed files.

The implication on UNAM is that there is a need for an RDM policy that can clearly outline the data preservation models to be used, which data needs to be preserved and the duration.

5.4.2 Long term data preservation

Empirical literature in Chapter two highlighted the need for long-term preservation of research. The present study acknowledges that when there is an RDM policy in place, UNAM will ensure that research data is preserved for its researchers for a long-term period which will ensure long-term access. The long-term preservation of research data has gained importance in recent years and there has been on-going awareness raising and developments to address the long-term preservation of valuable research data. Data storage media are all presented with a specific lifespan and the management of data has to ensure that data is migrated to new and available media to avoid media from becoming obsolete. Some file formats are also becoming obsolete and as such, some word processing formats make it difficult to interpret- the data. Long term data preservation accounts for all processes that ensure that data remains accessible over a long period (Layne et al., 2011).

Researchers support the long-term preservation of research data because there are some types of data that one would wish to keep for as long as they are active in research. But, many respondents indicated that there should be a guiding policy or document stipulating which data to keep for such a long period. The implication on UNAM and researchers is that data might be lost along the way if RDM is not implemented to provide solutions to practices already in place.

5.4.3 Research data linkages with research outputs

Empirical literature in Chapter two indicated that recent developments in research require that researchers submit their raw data as part of their research work submissions, thus making raw data available. This helps other researchers to validate their research outputs and ensuring that there is a link between the raw data collected and the findings of the study. Eight researchers were in support of linking data to research findings. The present study

acknowledges that linking data with research outputs is sensitive in nature and that it requires proper understanding from the researcher's point of view. The findings revealed that UNAM researchers had mixed feelings about sharing their research data. The implication of the findings is that the lack of an RDM policy is a serious contributing factor to much of what is happening at UNAM. The lack of RDM integration and implementation was key to answering the mixed feelings and perceptions by UNAM researchers.

5.5 Data access and re-use

The availability of new digital environments has placed data into the centre of scientific discovery. Data has become a centrepiece of modern scientific discovery and the sharing is now a crucial element of scientific progress which makes it a foundation for Open Science initiatives which advocate for scientific research data to be accessible to all (Tenopir et al., 2020, p. 2). In addition, Curty et al. (2016, p. 1) depicted that open science initiatives have paved the way for increased transparency and reproductively in research. The expedition on open science to research data has witnessed a recent propagation of research data repositories which enables research data to be shared. Despite advances in data repositories, there are still existing barriers on data sharing and re-use which include among others, perceived risk of sharing and re-using others data, quality concerns and lack of incentives (Curty et al., 2016).

Empirical literature in Chapter two revealed that data accessibility and re-use are not far from sharing research data with research outputs due to data security, privacy, acknowledgment and rewards. Seven researchers were in support of making their research data available for validation and use by other researchers.

The findings showed that data accessibility is crucial in avoiding duplicating data collection efforts because the data will be available to use, interpret and suggest new models that one

might have overlooked or not thought about during your own study. Some researchers indicated that they benefited from other researcher's data which was available through similar ideologies and initiatives, thus enhancing their own position towards making data available for re-use by other researchers.

The position of researchers on data accessibility and re-use can be attributed to their level of awareness and understanding of RDM and exposure through publishing and publishers' requirements.

5.7 RDM framework

Literature confirms that RDM challenges are a global phenomenon with major developments relating to technical orientation services such as data cataloguing, curation skills and capabilities, support services, buy-in from researchers and organisational leadership (Tang & Hu, 2019, p. 88). Such findings were also reported by Faniel and Connaway (2018) in which the findings revealed that factors such as human resources, communication, coordination, collaboration, technical resources, leadership support and researchers' perceptions were some of the aspects that hinder RDM development and implementation. Mosha et al. (2020, p. 484) depicts that RDM policy and guidelines is key to knowledge creation and institutional development, hence the demand for a policy that is able to guide capturing research data from its inception, use, storage, archiving and retrieval.

The finding showed that researchers at UNAM have RDM challenges that the institution is not able to address. Some of the researchers believe that RDM should be an organisational initiative that is developed and implemented through the research processes at UNAM and channelled to all the staff through a policy or framework. Researchers indicated that the

absence of a framework and RDM infrastructure at UNAM contributes to some of the challenges currently on the ground.

The RDM framework should be able to iron out all pertinent issues relating to data management such as privacy, security and outlining of all the relevant stakeholders at UNAM. Researchers stated that they are subscribing to cloud services using their own resources, whereas UNAM was supposed to avail data storage platforms and guidelines for data management and preservation. The implication of the findings is that an RDM framework is long overdue at UNAM and the institution should do a pilot study with other leading universities in the SADC region to acquaint themselves with what is needed in order to implement RDM successfully.

5.9 Chapter summary

This chapter discussed the findings and interpretation of data about RDM. Implications of the study were discoursed. The study established that RDM awareness is important to researchers as it enables researchers to integrate it as part of their research processes. The study also revealed that present RDM challenges are as a result of no existing RDM policy. The next chapter (Chapter six) provides the summary, conclusion, and recommendations for UNAM on how they should solve the existing RDM challenges. It also proposes an appropriate RDM framework for guiding the implementation of RDM at UNAM.

CHAPTER SIX

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS OF THE STUDY

6.1 Introduction

The previous chapter presented the data interpretation and discussion of the research findings. The researcher gave a meaningful and clear understanding of the findings so that the reader would be able to understand the data collected in accordance with the objectives of the study. Chapter six presents the summary, conclusions and recommendations of the study on RDM. The summary, conclusions and recommendations of the study are based on research objectives presented and interpreted in the previous chapters (4 and 5).

Recommendations are provided to address the shortcomings identified in RDM. A framework for guiding RDM is presented. It is hoped that such a framework will enhance RDM capabilities at UNAM. If implemented, the framework may assist UNAM researchers to address their current and persisting RDM challenges.

This chapter also makes suggestions for future work arising out of the study. The summary, conclusions and recommendation discuss the following:

- Awareness and data creation,
- Data processing and analysis,
- Data preservation,
- Data accessibility and re-use, and
- RDM framework.

6.2.1 Awareness and data creation

The first objective of the study was to ascertain the RDM awareness and establish if UNAM researchers create research data during their research processes. In this data-driven society, RDM is at the fore-front of guiding researchers through different stages of the research data lifecycle. It is inclusive of processes that deal with data creation / collection, implementation of infrastructure and services to preserve, and ensuring the accessibility and re-use of research data (Manu, 2021).

Empirical literature in Chapter two depicted that researchers' awareness about RDM plays a significant role in its implementation. The following is a synopsis of the objective findings:

- The study established that researchers were aware of RDM, some from a rhetoric orientation while others have integrated RDM practices as part of their research processes at UNAM;
- The study also discovered that UNAM researchers create diverse research data which is dictated by the different research methodologies and research tools adopted to address the objectives of any particular study. This practice presents researchers with data created in different formats;
- Data storage and storage infrastructure is a key component of data management. The study established that although researchers create a lot of data during their research processes, storage depends on the researcher's level of expertise, the digital divide and connection to available technologies and platforms; and
- Metadata is essential to research data as much as water is essential to human beings. Metadata help researchers to retrieve their data easily. The study discovered that although some researchers apply metadata to their research data, some researchers do not bother attaching metadata to their research data.

The implication emanating from the above findings are that although researchers at UNAM have realised the importance of RDM, UNAM should implement RDM through a policy to guide researchers who are at this juncture striving on their own and using their own means to store research data.

6.2.2 Data processing and analysis

The second objective of the study looked at whether researchers at UNAM undertake data processing and analysis during their research processes. Empirical literature in Chapter two depicted that data processing is vital in research as it helps with cleansing data to avoid any possible errors that would jeopardise the quality and integrity of research (Rughunathan, 2013). It is through data processing and analysis that data is converted into meaningful information and knowledge. Through data processing and analysis, researchers ensure that data anonymity is addressed and the identity of research participants is protected. The findings disclose that issues of data publishing show sceptical responses among researchers at UNAM which is resulting from data security, privacy and credit.

The implication is that UNAM researchers have access to credible and quality data that is processed and analysed. It implies that UNAM should work on an RDM framework that is able to guide researchers on concepts of data publishing and why it is important to the researcher and UNAM at large.

6.2.3 Data preservation

The third objective of the study looked at the data preservation strategies by UNAM researchers. Empirical data details that data preservation advocates for research data to continue to be accessible over time even when the current technology becomes obsolete (Rahmanto & Riasetiawan, 2018). The following is a synopsis of the objective findings:

The study discovered that researchers at UNAM have understood and realised the importance of data preservation. Although done at individual level, numerous researchers have adopted various technologies and platforms to preserve their research data such as subscribing to Dropbox, using Good Drive, external hard drives, disks, emails and physical hard copies, depending on the level of expertise, digital divide, technological penetration and level of know-how. Different researchers preserve their data according to what works for them.

This implies that UNAM should implement RDM, and avail storage infrastructure that is able to preserve research data using current formats and technologies and ensuring that technology obsolescence is addressed.

6.2.4 Data accessibility and re-use

The fourth objective of the study was concerned with UNAM researchers adopting data accessibility and re-use by other researchers. Data sharing is significant to the improvement and augmentation of scientific inquiry through a variety of other activities such as testing theories and hypotheses and methodologies used for particular studies (van Wyk, 2018, p. 245). The findings established that researchers at UNAM have diverse opinions towards availing their research data for use by other scholars. This process is essential in situations where researchers find it difficult or expensive to collect research data (Corti et al., 2014, p. 169). This implies that researchers at UNAM need a policy to guide how their research data will be shared and most importantly, taking into consideration issues of consent. The top researchers indexed in Scopus and Web of Science could be used to champion research at UNAM.

6.2.5 RDM framework

The fifth and final objective of the study looked at possible RDM challenges that UNAM staff encounter during their research processes and at the same time proposing an RDM framework that could address RDM challenges at UNAM. Empirical literature in Chapter two indicated that RDM policies are important in identifying key stakeholders in the research process and outlining different key elements pertaining to data management from the time it is created until the end of the cycle. In addition, RDM policies are critical to ensuring that research data created will be managed appropriately throughout its entire lifecycle and will be able to be discoverable and made available to allow re-use (Erway, 2013). RDM is complex in nature as it involves key stakeholders to ensure successful implementation and integration by researchers. At UNAM, key stakeholders include among many: the VC, PVCs (RID & Academic Affairs), Centre for Research Services, Centre for Grants Management and Resource Mobilisation, Information and Communication Technology Services, Faculty Deans and Heads of Schools, Research coordinators, Researchers, UNAM Press, UNAM Library and others.

The study established that much of the RDM challenges at UNAM can be attributed to a lack of an RDM framework. This is justified by literature as according to Patel (2016, p.128), RDM frameworks are fundamental to mandating data sharing in an institution by spelling out the purpose, scope, applicability and guidelines to the data contributors relating to data submission, licensing, metadata entry, data classification, copyright agreements and conditions under which the data withdrawal requests, if any, will be considered, the terms and conditions of the use of data, the protection of confidentiality of sensitive data, the protection of data against security breaches, and Intellectual Property (IP) concerns.

The implication here is that UNAM is challenged to develop and implement and an RDM framework.

6.2.5.1 Proposed UNAM RDM framework

The proposed framework (see Figure 6.1) is based on the study's findings, which are detailed in Chapters four and five, as well as the literature review, which is detailed in Chapter two.

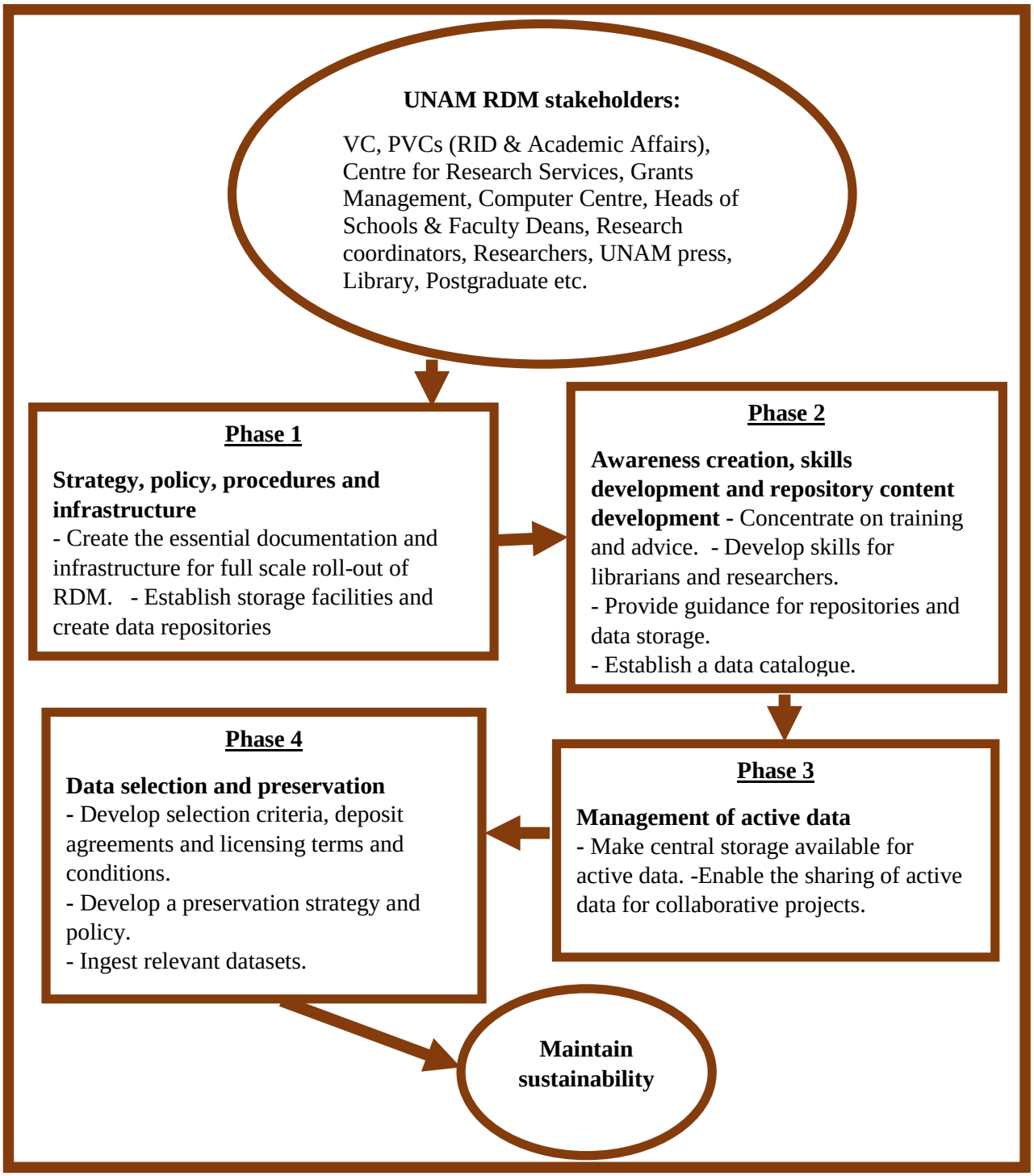


Figure: 6.1 proposed framework for research data management at UNAM (Researcher, 2022)

6.2.5.2 Explanation of the framework

RDM requires a significant investment in Information Communication Technology (ICT) infrastructure, therefore this requires inputs from all relevant stakeholders at UNAM. It is critical that UNAM management approves and supports RDM integration and implementation as part of UNAM research initiatives. This will help UNAM to meet research funders' requirements and provide better internal tracking of investments in research. Just like other initiatives, the proposed RDM framework can be implemented in stages:

a) **Strategy, policy, procedures and infrastructure (Stage One):**

This stage involves all UNAM RDM stakeholders as guided by the Centre for Research Services as a custodian to work on an RDM framework that can be rolled out through the entire UNAM community and the Centre for Communication Technology Services at UNAM to provide data storage and management services to researchers.

b) **Awareness creation, skills development and repository content development (Stage Two):**

As depicted by Patterton et al. (2018, p. 20), raising awareness about RDM and promoting good RDM practices in the institution is important for the success of an implementation plan. Although word-of-mouth and ad hoc interventions are valuable, the importance of planned promotion and awareness campaigns should be stressed.

c) **Management of active data: (Stage Three):**

When RDM infrastructure are fully developed, data emanating from collaborative research with other universities can be stored and shared with other scholars.

d) **Data selection and preservation: (Stage Four):** The UNAM Library as an RDM stakeholder develops selection requirements and appraisal criteria to guide the preservation of data and the ingestion of data sets in different formats.

6.3 Conclusions of the study from the findings

The study achieved its purpose of investigating and developing RDM capabilities at UNAM. The conclusions of the study as per objectives are presented below:

6.3.1 Awareness and data creation

Researchers at UNAM are aware of RDM and many have integrated it as part of their research processes which involve data creation, storing and attaching metadata. The level of awareness proved to be the determining factor in how researchers tackle RDM at UNAM.

6.3.2 Data processing and analysis

Data processing and analysis is a key component as part of the research process and UNAM researchers' value and understand the importance of data processing and analysis which is key to data quality and integrity.

6.3.3 Data preservation

Data analysed and processed requires preservation, whether short-term or long-term preservation. It has been found out that researchers use diverse strategies and techniques to preserve research data. Experiences, exposure, previous affiliation and digital divide inform how a researcher would deal with issues of data preservation.

6.3.4 Data accessibility and re-use

Data accessibility and re-use are at the fore-front of developments in research, especially research conducted using tax payers' money. Although sensitive in nature, researchers at UNAM raised support and concerns about availing their research data.

6.3.5 RDM challenges and framework

The study concludes that without an RDM policy, the success of any RDM initiative is limited and therefore UNAM should look into implementing an RDM policy that can guide RDM implementation at UNAM. A draft RDM policy is presented in this study under Appendix E.

6.4 Recommendations

Arising from the findings of the study, the following are the recommendations as per the objectives of the study:

6.4.1 Awareness and data creation

The study recommends that RDM awareness is key to RDM practices among researchers, therefore, UNAM through an RDM policy should raise the awareness of all researchers to integrate RDM as part of their research processes.

6.4.2 Data processing and analysis

They study recommends that although researchers at UNAM undertake data processing and analysis, UNAM should subscribe to data analysis software for both qualitative and quantitative research annually as a way to enhance research processes.

6.4.3 Data preservation

The study recommends that data preservation among researchers should not be an individual task but rather an institutional responsibility whereby preservation platforms are available for use by UNAM researchers.

6.4.4 Data accessibility and re-use

The study recommends that an RDM policy will be able to address concerns by some researchers on the issues of availing their data for re-use by other scholars and the benefits attached.

6.4.5 RDM framework

An RDM policy / framework is an urgent requirement at UNAM to be developed and implemented by all research stakeholders at UNAM. It is through a framework that RDM can become a success at UNAM. The university management should take the lead on developing an RDM framework by engaging all relevant internal stakeholders.

5.5 Contribution of study

The findings of this study contribute to the limited literature that is available regarding RDM at UNAM and Namibia at large. The results highlight some RDM positivity and the challenges that UNAM researchers encounter within UNAM. Researchers accepted their fate in relation to RDM at UNAM and they acknowledged that the current practices are supposed to be supported by UNAM through the adopted RDM systems and platforms. International funding agencies have incorporated RDM requirements on funding for certain research projects along with international RDM standards. This poses a serious challenge to UNAM researchers due to the absence of RDM support services that are available at UNAM. The proposed RDM framework can only be a starting point to RDM at UNAM with a view of proper investigation and formulation of an RDM policy. The objective to propose a preliminary framework could only be identified as a small contribution to the wide field of the RDM concept at UNAM.

6.6 Recommendation for future studies

The findings of the study provide some implications for future studies. The following suggestions are made for consideration for future and further research:

- Due to the growing importance of RDM in the research process, UNAM should urgently conduct research on promoting awareness and the implementation of RDM.

- A pilot study can be carried out between UNAM and a few South African universities and research intuitions that have development and implemented RDM as part of their research processes.

6.7 Chapter summary

The chapter presented the results as well as a summary of the chapters. It also drew some conclusions from the research findings and made some recommendations based on them. The research was divided into six sections. The purpose of this research was to investigate RDM capabilities at UNAM. The study was guided by a mixed method research design whereby a bibliometric analysis was carried out in phase one which informed the participants for the second qualitative phase. Data for the qualitative phase was collected through semi-structured interviews. The study established that RDM at UNAM is a challenge because of lack of institutional support which demands UNAM to develop and implement an RDM framework that can raise UNAM's reputation in terms of RDM.

REFERENCES

Ahlfeldt, J. & Johnsson, M. (2015). *Research libraries and research data management within the humanities and social sciences*. https://lup.lub.lu.se/search/files/6286782/5050466.pdf

Alamri, W. A. (2019). Effectiveness of qualitative research methods: Interviews and diaries. *International Journal of English and Cultural Studies, 2(1*), 65-70. https://core.ac.uk/download/pdf/228085022.pdf

Alsaawi, A. (2014). A critical review of qualitative interviews. *European Journal of Business and Social Sciences, 3*(4), 149 – 156. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2819536

Ale Ebrahim, N., Salehi, H., Embi, M. A., Habibi, F., Gholizadeh, H., & Motahar, S. M. (2014). Visibility and citation impact. *International Education Studies, 7(4),* 120-125. https://eprints.um.edu.my/9813/1/Visibility_and_Citation_Impact-33984-120606-1-PB.pdf

Ani, O. E., & Onyancha, O. B. (2011). A measure of research productivity of Nigerian universities: A bibliometric analysis. In *Proceedings of the ISSI 2011 Conference: 13th International Conference of International Society for Scientometrics & Informetrics, Durban, South Africa* (pp. 04-07). https://www.issisociety.org/proceedings/issi_2011/ISSI_2011_Proceedings_Vol1_08.pdf

Apuke, O. D. (2017). Quantitative research methods: A synopsis approach. *Arabian Journal of Business and Management Review (Kuwait Chapter), 6*(10), 40-47, DOI 10.12816/0040336

Austin, C. C., Bloom, T., Dallmeier-Tiessen, S., Khodiyar, V. K., Murphy, F., Nurnberger, A., & Whyte, A. (2017). Key components of data publishing: Using current best practices to develop a reference model for data publishing. *International Journal on Digital Libraries, 18(2),* 77-92. https://doi.org/10.1007/s00799-016-0178-2

Avuglah, B. K. (2016). *Developing an implementation plan for Research Data Management (RDM) at the University of Ghana.* [Master's thesis, University of Pretoria]. http://hdl.handle.net/2263/62100

Ball, A. (2012). *Review of data management lifecycle models*. University of Bath. https://purehost.bath.ac.uk/ws/portalfiles/portal/206543/redm1rep120110ab10.pdf

Bailey, C., Binda, L., Clayford, M., Gcora, N., Khan, F., Kondlo, L., & Zulu, T. (2019). South African National Survey of Research and Experimental Development: Statistical report 2017/18. Centre for Science, Technology and Innovation Indicators. https://repository.hsrc.ac.za/bitstream/handle/20.500.11910/4396/11191.pdf?sequence=1

Baral, U. N. (2017). Research data in social science methods. *Journal of Political Science, 17,* 82-104. https://www.nepjol.info/index.php/JPS/article/view/20515

Bester, T. V. (2018). *A research data management framework for a South African university-based research entity*. [Master thesis, North-West University]. https://repository.nwu.ac.za/handle/10394/35615

Bhardwaj, P. (2019). Types of sampling in research. *Journal of the Practice of Cardiovascular Sciences, 5(3),* 1-57. https://www.jpcs.org/temp/JPractCardiovascSci53157-3504729_094407.pdf

Blumzon, C. F. I., & Pănescu, A. T. (2019). Data storage. In Bespalov, A.,Michel, M.C., Steckler, T. (Eds.). *Good research practice in non-clinical pharmacology and biomedicine* (pp. 277-297). Springer Open.

Biddle, C., & Schafft, K. A. (2015). Axiology and anomaly in the practice of mixed methods work: Pragmatism, valuation, and the transformative paradigm. *Journal of Mixed Methods Research, 9(4),* 320-334. https://journals.sagepub.com/doi/epub/10.1177/1558689814533157

Boaduo, N. A. P. (2011). Systematic analysis and interpretation of collected data for a research study: A practical methodological framework for writing research report. *Educational Research and Reviews, 6(2),* 140-146. https://academicjournals.org/article/article1379670097_Boadou.pdf

Bolderston, A. (2012). Conducting a research interview. *Journal of Medical Imaging and Radiation Sciences*, *43*(1), 66-76. https://www.researchgate.net/publication/257735505_Conducting_a_Research_Interview

Borglund, E., & Engvall, T. (2014). Open data?: Data, information, document or record?. *Records Management Journal*, *24*(2), 163-180. https://doi.org/10.1108/RMJ-01-2014-_0012

Briney, K. (2015). *Data Management for Researchers: Organize, maintain and share your data for research success*. Pelagic.

Buys, C.M., & Shaw, P.L. (2015). Data management practices across an institution: Survey and report. *Journal of Librarianship and Scholarly Communication*, *3*(2), 12-25. https://iastatedigitalpress.com/jlsc/article/id/12748/

Byl, L., Carson, J., Feltracco, A., Gooch, S., Gordon, S., Kenyon, T., & Stirling, P. (2016). *White paper: Measuring research outputs through bibliometrics*. https://uwspace.uwaterloo.ca/handle/10012/10323

Callaghan, S., Murphy, F., Tedds, J., Allan, R., Kunze, J., Lawrence, R., & Whyte, A. (2013). *Connecting data repositories and publishers for data publication*. http://cedadocs.ceda.ac.uk/951/

Candela, L., Castelli, D., Manghi, P., & Callaghan, S. (2017). On research data publishing. *International Journal on Digital Libraries*, *18*(2), 73-75. https://link.springer.com/article/10.1007/s00799-017-0213-y

Carpenter, C. R., Cone, D. C., & Sarli, C. C. (2014). Using publication metrics to highlight academic productivity and research impact. *Academic Emergency Medicine*, *21*(10), 1160-1172. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4987709/pdf/nihms807788.pdf

Chawinga, W. D., & Zinn, S. (2020). Research data management at a public university in Malawi: The role of "three hands". *Library Management*, *41*(6/7), 467-485. https://www.emerald.com/insight/content/doi/10.1108/LM-03-2020-0042/full/html

Chigwada, J. P. (2021). Management and maintenance of research data by researchers in Zimbabwe. *Global Knowledge, Memory and Communication, 1-15.* http://ir.cut.ac.zw:8080/xmlui/handle/123456789/50?show=full

Chigwada, J., Chiparausha, B., & Kasiroori, J. (2017). Research data management in research institutions in Zimbabwe. *Data Science Journal, 16(31),* 1-9. https://datascience.codata.org/articles/10.5334/dsj-2017-031/

Chiware, E. R., & Becker, D. A. (2018). Research data management services in southern Africa: A readiness survey of academic and research libraries. *African Journal of Library Archives and Information Science*, *28*(1), 1-16. https://www.ajol.info/index.php/ajlais/article/view/174148

Chiware, E., & Mathe, Z. (2015). Academic libraries' role in research data management services: A South African perspective. *South African Journal of Libraries and Information Science*, *81*(2), 1-10. https://www.researchgate.net/publication/294426382

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

Collis, J., & Hussey, R. (2013). *Business research: A practical guide for undergraduate and postgraduate students*. Macmillan International Higher Education.

Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., & McGillivray, B. (2020). The citation advantage of linking publications to research data. *PloS One*, *15*(4), e0230416. https://doi.org/10.1371/journal.pone.0230416

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative and mixed methods approaches.* (5th ed.). Sage.

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

Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2019). *Managing and sharing research data: a guide to good practice*. Sage.

Curty, R., Yoon, A., Jeng, W., & Qin, J. (2016). Untangling data sharing and reuse in social sciences. *Proceedings of the Association for Information Science and Technology*, *53*(1), 1-5. https://asistdl.onlinelibrary.wiley.com/doi/pdf/10.1002/pra2.2016.14505301025

Dashrath, V.B. (2014). Role of metadata in digital resource management. *International Journal of Digital Library Services, 209-217.* *http://www.ijodls.in/uploads/3/6/0/3/3603729/19434.pdf*

David, M., & Sutton, C. D. (2011). *Social research: An introduction*. (2nd ed.). Sage.

De Lisle, J. (2011). The benefits and challenges of mixing methods and methodologies: Lessons learnt from implementing qualitatively led mixed methods research designs in Trinidad and Tobago. *Caribbean Curriculum Journal, 18,* 87-120. https://uwispace.sta.uwi.edu/dspace/bitstream/handle/2139/11308/Mixing%20Methods%20and%20Methodologies%20De%20Lisle.pdf

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives*, *38*(1), 105-124. https://www3.nd.edu/~ggoertz/sgameth/Drost2011.pdf

Du Plooy-Cilliers, F., Davis, C., & Bezuidenhout, R. (2014). *Research matters*. Juta.

Dube, L., Mhlongo, M., & Ngulube, P. (2014). The ethics of anonymity and confidentiality: Reading from the University of South Africa policy on research ethics. *Indilinga African Journal of Indigenous Knowledge Systems*, *13*(2), 201-214. https://uir.unisa.ac.za/handle/10500/22229

Ellegaard, O., & Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*, *105*(3), 1809-1831. https://link.springer.com/article/10.1007/s11192-015-1645-z

Erway, R. (2013). *Starting the conversation: University-wide research data management policy*. Dublin, Ohio: OCLC Research https://www.oclc.org/content/dam/research/publications/library/2013/2013-08.pdf

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4. DOI: 10.11648/j.ajtas.20160501.11

Faniel, I. M. & Connaway, L. S. (2018). Librarians' on the factors influencing research data management programs. *College & Research Libraries*, *79*(1), 100-119. https://crl.acrl.org/index.php/crl/article/view/16610/18056

Farnel, S., & Shiri, A. (2014). Metadata for research data: Current practices and trends. *International conference on Dublin core and metadata applications*, 74-82. https://dcpapers.dublincore.org/pubs/article/view/3714

Federer, L. (2016). Research data management in the age of big data: Roles and opportunities for librarians. *Information Services & Use*, *36*(1-2), 35-43. https://content.iospress.com/articles/information-services-and-use/isu797

Flick, U. (Ed.). (2013). *The SAGE handbook of qualitative data analysis*. Sage.

Fulton, S. & Krainovich-Miller, B. (2010). Gathering and appraising the literature. In G., LoBiondo-Wood, & J. Haber (Eds.). *Nursing research: Methods and critical appraisal for evidence-based practice* (7th ed.). Mosby Elsevier.

Garaba, F. (2010). *An investigation into the management of the records and archives of former liberation movements in east and southern Africa held by national and private archival institutions.* [Doctoral dissertation, University of KwaZulu-Natal]. https://researchspace.ukzn.ac.za/handle/10413/10545

Gentles, S. J., Charles, C., Ploeg, J., & McKibbon, K. A. (2015). Sampling in qualitative research: Insights from an overview of the methods literature. *The Qualitative Report*, *20*(11), 1772-1789. https://nsuworks.nova.edu/tqr/vol20/iss11/5/

Ghotkar, M., & Rokde, P. (2016). Big data: How it is generated and its importance. *IOSR Journal of Computer Engineering, 1-5*. https://www.iosrjournals.org/iosr-jce/papers/conf.15013/Volume%202/1.%2001-05.pdf?id=7556

Grant, C., & Osanloo, A. (2014). Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for 'house'. *Administrative Issues Journal: Connecting Education, Practice and Research*, *4*(2), 12-22. https://files.eric.ed.gov/fulltext/EJ1058505.pdf

Haustein, S., Larivière, V. (2015). The Use of Bibliometrics for Assessing Research: Possibilities, Limitations and Adverse Effects. In: Welpe, I., Wollersheim, J., Ringelhan, S., & Osterloh, M. (eds). *Incentives and Performance,* 121-139. Springer, Cham. https://doi.org/10.1007/978-3-319-09785-5_8

Hennink, M. M., Kaiser, B. N., & Weber, M. B. (2019). What influences saturation? Estimating sample sizes in focus group research. *Qualitative Health Research*, *29*(10), 1483-1496. https://journals.sagepub.com/doi/epub/10.1177/1049732318821692

Higman, R., & Pinfield, S. (2015). Research data management and openness: the role of data sharing in developing institutional policies and practices. *Program: Electronic Library and Information System*, *49*(4), 364-381. https://www.semanticscholar.org/paper/Research-data-management-and-openness%3A-The-role-of-Higman-Pinfield/37e88452147d3415bf7cee6600b6222b68c911d5

Honorene, J. (2017). Understanding the role of triangulation in research. *Scholarly research Journal for Interdisciplinary Studies*, *4(31*), 91-95. http://www.srjis.com/pages/pdfFiles/149544238718.%20HONORENO%20JOHNSON.pdf

Hottenrott, H., & Lawson, C. (2017). A first look at multiple institutional affiliations: A study of authors in Germany, Japan and the UK. *Scientometrics*, *111*(1), 285-295. https://link.springer.com/article/10.1007/s11192-017-2257-6

Hughes, P. (2010). Paradigms, methods and knowledge. In G. MacNaughton, S. Rolfe & I. Siraj-Blatchford (Eds.), *Doing early childhood research,* (2nd ed.) Open University Press.

Kabir, S. M. S. (2016). Basic guidelines for research. *An introductory approach for all disciplines*, *4*(2), 168-180.

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. https://www.tandfonline.com/doi/full/10.1080/00048623.2014.951910

Kamal, S. S. (2018). Research paradigm and the philosophical foundations of a qualitative study. *PEOPLE: International Journal of Social Sciences*, *4(3)*, 1386-1394. https://www.researchgate.net/publication/330798108

Khan, R. (2014). *Quantitative data analysis*. In Bezuidenhout, R., Davis, C., & Du Plooy-Cilliers, F. (Eds.), *Research matters* (pp. 228-251). Juta.

Kalusopa, T. (2011). *Developing an e-records readiness framework for labour organisation in Botswana.* [Doctoral Dissertation, University of South Africa] http://hdl.handle.net/10500/5690

Kalusopa, T., Zibani, P., Kanguti, R., & Leonard, A. (2020). Research productivity, visibility, and impact at the University of Namibia: Building a framework for open access and research data management. In *Handbook of research on connecting research methods for information science research*, 489-516. IGI Global. https://www.igi-global.com/chapter/research-productivity-visibility-and-impact-at-the-university-of-namibia/240541

Krier, L., & Strasser, C. A. (2014). *Data management for libraries: A LITA guide*. American Library Association.

Kennan, M. A., Williamson, K., & Johanson, G. (2012). Wild data: Collaborative e-research and university libraries. *Australian Academic & Research Libraries*, *43*(1), 56-79. https://www.tandfonline.com/doi/abs/10.1080/00048623.2012.10700623

Kivunja, C., & Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education*, *6*(5), 26-41.

https://files.eric.ed.gov/fulltext/EJ1154775.pdf

Koopman, M., & De Jager, K. (2016). Archiving South African research data: How ready are we? *South African Journal of Science*, *112*(7), 1-7. https://www.sajs.co.za/article/view/3572

Kremser, J., Kovácová, B., Pavlovská, M., Hejtmánek, L., & Antoš, D. (2012). *Long term preservation of digital data - background research*. CESNET.https://www.cesnet.cz/wp-content/uploads/2013/01/digital-data-long-term-preservation.pdf

Layne, R., Capel, A., Cook, N., Wheatley, M., & Contributors, J. E. (2012). Long-term preservation of scientific data: Lessons from jet and other domains. *Fusion Engineering and Design*, *87(12)*, 2209-2212. https://www.sciencedirect.com/science/article/pii/S0920379612003225

Louw, M. (2014). Ethics in research. In F., Du Plooy-Cilliers, C., Davis & R.M. Bezuidenhout (Eds.), *Research matters* (pp. 262-263). Juta and Company.

Luthuli, L. (2021). *A framework for web archiving for municipalities in the KwaZulu-Natal Province in South Africa*. [Unpublished doctoral dissertation, University of South].

Maarouf, H. (2019). Pragmatism as a supportive paradigm for the mixed research approach: Conceptualizing the ontological, epistemological, and axiological stances of pragmatism. *International Business Research*, *12(9)*, 1-12. https://www.ccsenet.org/journal/index.php/ibr/article/view/0/40468

Mabhiza, C. (2016). *An investigation of the information seeking behaviours of veterinary scientists in Namibia* [Masters thesis, University of Namibia]. https://repository.unam.edu.na/handle/11070/1643

MacLeod, W. B., & Urquiola, M. (2021). Why does the United States have the best research universities? Incentives, resources, and virtuous circles. *Journal of Economic Perspectives, 35(1)*, 185-206. https://doi.org/10.1257/jep.35

Manu, T. R., & Bhakti, G. (2021). Growth and development of research data management in India: A visual analysis of published literature indexed in Scopus. *Library Philosophy and Practice*, 1-21. https://digitalcommons.unl.edu/libphilprac/6651/

Manu, T.R. (2018). *Researchers' perceptions on research data management: A survey*. https://www.researchgate.net/publication/326082669

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., & Spies, J. R. (2016). Point of view: How open science helps researchers succeed. https://elifesciences.org/articles/16800

Majeed, A., & Lee, S. (2020). Anonymization techniques for privacy preserving data publishing: A comprehensive survey. *IEEE access*, *9*, 8512-8545. https://ieeexplore.ieee.org/document/9298747

Majid, U. (2018). Research fundamentals: Study design, population, and sample size. *Undergraduate Research in Natural and Clinical Science and Technology Journal*, *2*, 1-7. https://urncst.com/index.php/urncst/article/view/16/7

Mills, M.S. (2013). Facilitating multimodal literacy instruction through digital curation. In J. Whittingham, S. Huffman, W. Rickman, & C Wiedmaier Hershey (Eds.), Technological tools for the literacy classroom, pp. 46–63, Hershery, PA: Idea Group Inc (IGI).

Matlatse, R. (2016). *An evaluation of a structured training event aimed at enhancing the research data management knowledge and skills of library and information science professionals in South African Higher Education Institutions* [Master's thesis, University of Pretoria]. https://repository.up.ac.za/handle/2263/58393?show=full

Mitchell, A. (2018). A review of mixed methods, pragmatism and abduction techniques. *The Electronic Journal of Business Research Methods*, 16(3), 103-116. https://www.researchgate.net/publication/328343822

Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University. Economic Series*, *17*(4), 59-82. https://mpra.ub.uni-muenchen.de/83458/

Mosha, N. F., Luhanga, E. T., Mosha, M. V., & Marwa, J. J. (2020). Research data management among researchers in higher learning institutions of Sub-Saharan Africa. In *Handbook of research on connecting research methods for information science research* (pp. 479-488). IGI Global. DOI: 10.4018/978-1-7998-1471-9.ch024

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*. https://www.researchgate.net/publication/330853802

Najar, J. K., & Wani, J. A. (2019). Digital preservation: An overview. *Library Philosophy and Practice (e-journal), 1-16.* https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=6113&context=libphilprac

Neuroth, H., Strathmann, S., Oßwald, A., & Ludwig, J. (Eds.). (2013). *Digital curation of research data: Experiences of a baseline study in Germany*. Universitätsverlag Göttingen. https://univerlag.uni-goettingen.de/bitstream/handle/3/isbn-978-3-86488-054-_4/Digital_Curation_SUB.pdf?sequence=1&isAllowed=y

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. https://journals.co.za/doi/pdf/10.10520/EJC-13d7be91a2

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.

Nilesh, S., & Verma, S. (2012). *Digital preservation, issues and their strategies for libraries: A modern era.* *https://www.academia.edu/28653621/Digital_Preservation_Issues_and_Their_Strategies_for_Libraries_A_Modern_Era*

Ocholla, D. N., & Le Roux, J. (2011). Conceptions and misconceptions of theoretical frameworks in library and information science research: A case study of selected theses and dissertations from eastern and Southern African universities. *Mousaion, 29*(2), 61-74. http://www.lis.uzulu.ac.za/2011/Ocholla%20and%20Le%20Roux%20prolissa%20conference%202011%20revised%2016%20March%202011.pdf

Onyancha, O. B. (2016). Open research data in Sub-Saharan Africa: A bibliometric study using the data citation index. *Publishing Research Quarterly*, *32*(3), 227-246. https://uir.unisa.ac.za/handle/10500/22031

Onwuegbuzie, A. J., & Combs, J. P. (2011). Data analysis in mixed research: A primer. *International Journal of Education, 3(1),* 1-25. https://www.researchgate.net/publication/267387779_Data_Analysis_in_Mixed_Research_A_Primer

Oruc, A. (2021). Requirements for Productivity in the Academic Environment. *Higher Education Studies, 11*(4), 40-47. https://files.eric.ed.gov/fulltext/EJ1322311.pdf

Pal, J. K. (2020). Understanding research productivity in the realm of evaluative scientometrics. *Annals of Library and Information Studies (ALIS), 67*(1), 67-69. http://eprints.rclis.org/41739/1/understanding%20research%20productivity.pdf

Pal, S. K. (2012). *Digital preservation methods and technology in digital era*. https://www.academia.edu/13486465/Digital_Preservation_Methods_and_Technology_In_Digital_Era

Partlo, K., Symons, D., & Carlson, J.D. (2015). Revolutionary or evolutionary? Making research data management manageable. In B.L. Eden (Ed.), *Creating research infrastructures in the 21st century academic library: Conceiving, funding, and building new facilities and staff* (pp. 175-201). Rowman and Littlefield. https://core.ac.uk/download/pdf/229036206.pdf

Pasquetto, I.V., Randles, B.M., & Borgman, C.L. (2017). On the reuse of scientific data. *Data Science Journal*, 16(8), 1–9. https://doi.org/10.5334/dsj-2017-008

Patel, D. (2016). Research data management: A conceptual framework. *Library Review, 65*(4/5), 226-241. https://www.emerald.com/insight/content/doi/10.1108/LR-01-2016-_0001/full/html

Patterton, L. H. (2016). *Research data management practices of emerging researchers at a South African research council* [Doctoral dissertation, University of Pretoria]. https://repository.up.ac.za/handle/2263/59502

Patterton, L., Bothma, T. J., & Van Deventer, M. J. (2018). From planning to practice: An action plan for the implementation of research data management services in resource-constrained institutions. *South African Journal of Libraries and Information Science*, *84*(2), 14-26. https://journals.co.za/doi/epdf/10.7553/84-2-1761

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd Ed.). Newbury Park, CA: Sage.

Phiri-Chigwada, J. (2021). *Research Data management services in tertiary institutions in Zimbabwe.* IGI Global. http://ir.cut.ac.zw:8080/xmlui/handle/123456789/66

Pienaar, H. (2011). *Survey of research data management practices at the University of Pretoria: Undertaken by the Department of Library Services in order to improve research practices at the University*. https://repository.up.ac.za/handle/2263/17475

Pinfield, S., Cox, A.M., & Smith, J. (2014). Research data management and libraries: Relationship, activities, drivers and influences. *PLoS ONE*, *9*(12), 1-28. https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0114734&type=printable

Poole, A. H. (2016). The conceptual landscape of digital curation. *Journal of Documentation*, *75*(5), 961-986. https://www-emerald com.ezproxy.unam.edu.na/insight/content/doi/10.1108/JD-10-2015-0123/full/html

Pouchard, L. (2015). Revisiting the data lifecycle with big data curation. *International Journal of Digital Curation, 10(2),* 176—192. https://www.researchgate.net/publication/305095078

Pryor, G. (2012). *Managing Research Data*. Facet Publishing, Cambridge. https://doi.org/10.29085/9781856048910

Connaway, L. S., & Powell, R. R. (2010). *Basic research methods for librarians*. ABC-CLIO. https://repo.iainbatusangkar.ac.id/xmlui/bitstream/handle/123456789/8159/150909655696 6_[Lynn_Silipigni_Connaway,_Ronald_R._Powell]_Basic_(BookFi.org).pdf?sequence=1

Purkayastha, A., Palmaro, E., Falk-Krzesinski, H. J., & Baas, J. (2019). Comparison of two article-level, field-independent citation metrics: Field-Weighted Citation Impact (FWCI) and Relative Citation Ratio (RCR). *Journal of Informetrics*, *13*(2), 635-642. https://www.sciencedirect.com/science/article/pii/S1751157718303559

Qasim, U., Davis, C., Garnett, A., Marks, S., & Moosberger, M. (2018). *Research data preservation in Canada: A white paper.* https://alliancecan.ca/en/services/research-data-_management

Quinlan, C., Zikmund, W., Babin, B., Carr, J., & Griffin, M. (2015). *Business research methods*. South Western Cengage.

Rafiq, M., & Ameen, K. (2022). Research data management and sharing awareness, attitude, and behavior of academic researchers. *Information Development*, *38(3)*, 391-405.https://www.sciencegate.app/document/10.1177/02666669211048491

Sanjeeva, M. (2018). *Research data management: A new role for academic/ research librarians*. https://www.researchgate.net/publication/323604761

Raghunathan, B. (2013). *The complete book of data anonymization: From planning to implementation*. CRC Press.

Rahmanto, K. N., & Riasetiawan, M. (2018, August). Data preservation process in big data environment using open archival information system. In *2018 4th International Conference on Science and Technology (ICST)* (pp. 1-5), Yogyakarta, Indonesia, 2018, pp. 1-5. https://ieeexplore.ieee.org/document/8528669

Rafiq, M., & Ameen, K. (2022). Research data management and sharing awareness, attitude, and behavior of academic researchers. *Information Development*, *38*(3), 391-405. https://journals.sagepub.com/doi/full/10.1177/02666669211048491

Rahul, K., Banyal, R. K., & Goswami, P. (2020). Analysis and processing aspects of data in big data applications. *Journal of Discrete Mathematical Sciences and Cryptography*, *23*(2), 385-393. https://www.tandfonline.com/doi/pdf/10.1080/09720529.2020.1721869

Ray, J. M. (2014). Introduction to research data management. In. J. M. Ray. (Ed.). *Research data management: Practical strategies for information professionals*. Purdue University Press.

Rehman, A. A., & Alharthi, K. (2016). An introduction to research paradigms. *International Journal of Educational Investigations*, *3(8)*, 51-59. http://www.ijeionline.com/attachments/article/57/IJEI.Vol.3.No.8.05.pdf

Rice, R., Ekmekcioglu, Ç., Haywood, J., Jones, S., Lewis, S., Macdonald, S., & Weir, T. (2013). Implementing the research data management policy: University of Edinburgh roadmap. *International Journal of Digital Curation*, *8*(2), 194-204.

Rorty, R. (1990). Pragmatism as anti-representationalism. In J. P. Murphy. (Ed.), *Pragmatism: From Peirce to Davidson* (pp. 1-6). Boulder, CO: Westview.

Salmi, J. (2016). Excellence strategies and the creation of world-class universities. In N.C. Liu, Y., Cheng, & Q., Wang, (Ed.), *Matching visibility and performance* (pp. 13-48). Brill.

Sadiku, M., Shadare, A., & Musa, S. (2017). Digital preservation. *International Journal of Advanced Research in Computer Science and Software Engineering, 7*, 5-6. DOI:10.23956/ijarcsse/V7I4/01408

Samupwa, A. N. (2019). *Adopting research data management (RDM) practices at the University of Namibia (UNAM): A view from researchers.* [Master's thesis, University of Cape] http://hdl.handle.net/11427/31205

Sanjeeva, M. (2018). *Research data management: A new role for academic/ research librarians.* https://www.researchgate.net/publication/323604761

Sanmorino, A., & Rini, D. P. (2021). Building research productivity framework in higher education institution. *International Journal of Advanced Computer Science and Applications*, *12*(6), 184-191. https://thesai.org/Downloads/Volume12No6/Paper_20Building_Research_Productivity_Framework.pdf

Saunders, B., Kitzinger, J., & Kitzinger, C. (2015). Anonymising interview data: Challenges and compromise in practice. *Qualitative Research*, *15*(5), 616-632. https://journals.sagepub.com/doi/full/10.1177/1468794114550439

Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design?. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *69*(2), 107-131. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5602001/

Schurink, E. (2009). Qualitative research design as tool for trustworthy research. *Journal of Public Administration*, *44*, 803-823. https://journals.co.za/doi/abs/10.10520/EJC51748

Shaambeni, F. N. (2019). *The use of print library materials in the era of digital resources in institutions of higher learning in Namibia: A case study of the University of Namibia (UNAM) and Namibia University of Science and Technology (NUST)* [Master thesis, University of Namibia]. http://repository.unam.edu.na/handle/11070/2587

Shimray, S. R., & Ramaiah, C. K. (2018). Digital preservation strategies: An overview. In *11th National Conference on Recent Advances in Information Technology (READIT). IGCAR, Kalpakam, Tamilnadu* (pp. 46-54). https://www.researchgate.net/publication/327221006_Digital_Preservation_Strategies_An_Overview

Showkat, N., & Parveen, H. (2017). Non-probability and probability sampling. *Media and Communications Study*, 1-9.

Shukla, T. (2018). *Data processing*. 10.13140/RG.2.2.35660.10889. DOI:10.13140/RG.2.2.35660.10889

Singh, N. K., Monu, H., & Dhingra, N. (2018, February). Research data management policy and institutional framework. In *2018 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS)* (pp. 111-115). IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8485259

Smith, T.L., Redd, K., Nusser, S., Samors, R., & Miller, E.R. (2021). *AAU APLU guide to accelerate public access to research data*. https://www.sciencegate.app/document/10.31219/osf.io/tjybn

Strasser, C., Cook, R., Michener, W., & Budden, A. (2012). *Primer on data management: What you always wanted to know.* https://escholarship.org/uc/item/7tf5q7n3

Subaveerapandiyan, A., & Maurya, A. (2022). Research data preservation practices of library and information science faculties. DESIDOC Journal of Library & Information Technology, 42(4), 259-264. https://philarchive.org/archive/SUBRDP

Sürücü, L., & Maslakci, A. (2020). Validity and reliability in quantitative research. *Business & Management Studies: An International Journal*, *8(3*), 2694-2726. https://www.bmij.org/index.php/1/article/view/1540

Swan, A. (2012). *Policy guidelines for the development and promotion of open access*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000215863

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. https://www.sciencedirect.com/science/article/pii/S2543925122000675

Tenopir, C., Suzie, A., Kimberly, D., Arsev, U.A., Lei W., Eleanor, R., Maribeth, M., & Mike, F. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE*, *6*(6), 1-21. http://dx.plos.org/10.1371/journal.pone.0021101.

Thanos, C. (2017). Research data reusability: Conceptual foundations, barriers and enabling technologies. *Publications*, *5*(1), 1-19. https://www.mdpi.com/2304-6775/5/1/2

Triasih, H., Rahmi, R., & Devi, K. S. (2020). Providing research data management services and practices at PDDI-LIPI: Preparedness, roles, challenges, and training. *Baca: Jurnal Dokumentasi Dan Informasi*, *41*(2), 169-178.

Tripathi, D. P., & Pandy, S. R. (2018, February). Developing a conceptual framework of research data management for higher educational institutions. In *2018 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS)* (pp. 105-110). IEEE. https://ieeexplore.ieee.org/document/8485193

Tripathi, M., Chand, M., Sonkar, S.K., & Jeevan, V.K.J. (2017). A brief assessment of researchers' perceptions towards research data in India. *IFLA Journal, 43*(1), 22-39. https://journals.sagepub.com/doi/10.1177/0340035216686984

Ungerer, L. M. (2016). Digital curation as a core competency in current learning and literacy: A higher education perspective. *The International Review of Research in Open and Distributed Learning, 17*(5), 1-27. https://files.eric.ed.gov/fulltext/EJ1117357.pdf

University of Namibia. (2013). *Research Ethics Policy, Regulations and Guidelines.* http://unamintranet.unam.na/documents/Research_Ethics_Policy.pdf

University of Namibia (UNAM). (2019). *Statistics.* http://www.unam.edu.na/about-unam/statistics

Umana, S. P. (2020). *Digital preservation in institutional repositories in Namibian academic libraries: A case study of the Namibia University of Science and Technology (NUST) and the University of Namibia (UNAM)* (Master thesis, University of Namibia, Windhoek, Namibia). http://repository.unam.edu.na/handle/11070/2908

Umar, H. S., & Madugu, U. (2015). The imperative of population sampling in social science research. *Global Journal of Political and Science and Administration, 3(3),* 49-57. https://www.eajournals.org/wp-content/uploads/The-Imperative-of-Population-Sampling-In-Social-Science-Research.pdf

Van de Sandt, S., Lavasa, A., Dallmeier-Tiessen, S., & Petras, V. (2019). Submitter: The definition of reuse. *Data Science Journal, 18*(22), 1-19. https://cds.cern.ch/record/2690554/files/DSJ%2018%2022.pdf

Van Deventer, M., & Pienaar, H. 2015. Research data management in a developing country: A personal journey. *International Journal of Digital Curation, 10*(2), 33-47. http://www.ijdc.net/article/view/10.2.33/406

Van Wyk, B. J. (2018). *The relationship between research data management and virtual research environments.* [Doctoral dissertation, University of Pretoria]. https://repository.up.ac.za/handle/2263/64292

Velmurugan, C. (2013). Digital preservation: Issues and challenges on libraries and information resource centers in India. e-*Library Science Research Journal, 1(8),* 1-9.

Verhaar, P., Schoots, F., Sesink, L., & Frederiks, F. (2017). Fostering effective data management practices at Leiden University. *Liber Quarterly: The Journal of European Research Libraries*, *27*(1).

Waltman, L., & Noyons, E. (2018). *Bibliometrics for research management and research evaluation.* Leiden Centre for Science and Technology Studies.

Wang, M. (2013). *Academic library, e-science/e-research, and data services in a broader context.* https://rucore.libraries.rutgers.edu/rutgers-lib/42186/PDF/1/play/

Wang, X., Duan, Q., & Liang, M. (2021). Understanding the process of data reuse: An extensive review. *Journal of the Association for Information Science and Technology*, *72*(9), 1161-1182. https://doi.org/10.1002/asi.24483

Wessels, B., Finn, R. L., Linde, P., Mazzetti, P., Nativi, S., Riley, S., & Wyatt, S. (2014). Issues in the development of open access to research data. *Prometheus, 32(1),* 49-66. DOI:10.1080/08109028.2014.956505

Whyte, A., & Tedds, J. (2011). *Making the case for research data management: A digital curation centre briefing paper*. https://www.dcc.ac.uk/guidance/briefing-papers/making-case-rdm

Wiggins, A., Bonney, R., Graham, E., Henderson, S., Kelling, S., LeBuhn, G., & ‘Newman, G. (2013). Data management guide for public participation in scientific n research. *DataOne Working Group*, 1-41. https://www.researchgate.net/publication/260389831_Data_managment_guide_for_public_participation_in_scientific_research

Winn, J. (2013). *Open data and the academy: An evaluation of CKAN for research data management.* http://eprints.lincoln.ac.uk/id/eprint/9778/1/CKANEvaluation.pdf

Wissik, T., & Ďurčo, M. (2016, April). Research data workflows: From research data lifecycle models to institutional solutions. In *Selected Papers from the CLARIN Annual Conference 2015, October 14–16, 2015, Wroclaw, Poland* (No. 123, pp. 94-107). Linköping University Electronic Press.

Zaman, A. N. K., & Obimbo, C. (2014). Privacy preserving data publishing: A classification perspective. *International Journal of Advanced Computer Science and Applications*, *5(9),* 129-134. DOI: 10.14569/IJACSA.2014.050919

Zikmund, W. G., & Babin, B. J. (2013). *Essentials of marketing research* (5th ed). Cengage Learning.

APPENDIX A

UNAM
UNIVERSITY OF NAMIBIA

ETHICAL CLEARANCE CERTIFICATE

Ethical Clearance Reference Number: HREC-NH/14/10/2020 **Date:** 26-10-2020

This Ethical Clearance Certificate is issued by the University of Namibia Research Ethics Committee (UREC) in accordance with the University of Namibia's Research Ethics Policy and Guidelines. Ethical approval is given in respect of undertakings contained in the Research Project outlined below. This Certificate is issued on the recommendations of the ethical evaluation done by the Faculty/Centre/Campus Research & Publications Committee sitting with the Postgraduate Studies Committee.

Title of Project: DEVELOPING A FRAMEWORK FOR RESEARCH DATA MANAGEMENT CAPABILITIES AT THE UNIVERSITY OF NAMIBIA

Nature/Level of Project: M.A. (NON-HEALTH) (NQF9)

Researcher: RONALD NGUATUMO KANGUTI

Student Number: 200746065

Faculty: HUMANITIES AND SOCIAL SCIENCES

Supervisor(s): PROF. T. KALUSOPA

Take note of the following:

(a) Any significant changes in the conditions or undertakings outlined in the approved Proposal must be communicated to the UREC. An application to make amendments may be necessary.
(b) Any breaches of ethical undertakings or practices that have an impact on ethical conduct of the research must be reported to the UREC.
(c) The Principal Researcher must report issues of ethical compliance to the UREC (through the Chairperson of the Faculty/Centre/Campus Research & Publications Committee) at the end of the Project or as may be requested by UREC.
(d) The UREC retains the right to:
(i) Withdraw or amend this Ethical Clearance if any unethical practices (as outlined in the Research Ethics Policy) have been detected or suspected,
(ii) Request for an ethical compliance report at any point during the course of the research.

REC wishes you the best in your research.

Prof. H.L. Beyer, Deputy Chair: HREC-NH
pp Chair: HREC-NH

APPENDIX B

CENTRE FOR POSTGRADUATE STUDIES

University of Namibia, Private Bag 13301, Windhoek, Namibia
340 Mandume Ndemufayo Avenue, Pioneers Park
☎ +264 61 206 3275/4662; Fax +264 61 206 3290; URL: http//www.unam.edu.na

UNAM
UNIVERSITY OF NAMIBIA

30 November 2020

Student Name: RONALD NGUATUMO KANGUTI

Student number: 200746065

Programme: Master of Arts

Approved research title: DEVELOPING A FRAMEWORK FOR RESEARCH DATA MANAGEMENT CAPABILITIES AT THE UNIVERSITY OF NAMIBIA

TO WHOM IT MAY CONCERN

I hereby confirm that the above mentioned student is registered at the University of Namibia for the programme indicated. The proposed study met all the requirements as stipulated in the University guidelines and has been approved by the relevant committees.

Permission is hereby granted to carry out the research as described in the approved proposal.

Best Regards

30/11/2020

Dr. Seth J. Eiseb
Acting Director: Centre for Postgraduate Studies
Tel: +264 61 2063414
E-mail: directorpgs@unam.na

Date

APPENDIX C

Research Data Management Interview guide

Dear Respondent,

I am Ronald N. Kanguti, a master's student at UNAM currently studying for a Masters of Arts in Library and Information Science. I am gathering data to enable me to complete my research titled: "**Developing a framework for research data management capabilities at the University of Namibia"**

Research Data Management (RDM) implies the process of controlling and organising the data generated during a research project, and covers the entire data lifecycle, which includes the planning of the investigation, conducting the investigation, storage and backing up of the data as it is created, preserving the data long-term, after the research investigation has concluded, and making the data accessible for future use.

Please be assured that participant's identity will be handled with strict confidentiality. Thank you very much for accepting to participate in this interview, your time is appreciated!!

Yours Sincerely,

Ronald N. Kanguti (MSc Candidate)

SECTION A: BACKGROUND INFORMATION

1. Gender

……

2. What is your position in UNAM?

……

3. In which faculty/ department/ unit are you?

……

4. What is your highest qualification?

……

5. How long have you been with UNAM?

……

6. Are you aware of the RDM concept?

……

SECTION B

DATA CREATION

1. What type of research data do you create or generate?

……

2. In what format is your research data?

……

3. How do you capture or collect research data?

……

4. How do you store your research data?

……

5. Do you attach metadata to your research data?

……

DATA PROCESSING AND ANALYSIS

1. Do you undertake data cleansing and validation?

……

2. Have you ever done data anonymization (Removing personal info)

……

3. Have you ever done data interpretation and analysis?

……

4. Do you publish your research data?

……

DATA PRESERVATION

1. Do you consider data archiving and have you ever preserved your research data?

……

2. Did you consider long-term data preservation?

……

3. Do you link your data to research outputs?

……

DATA ACCESSIBILITY AND RE-USE

1. Do you support data sharing and re-use by other scholars?

RESEARCH DATA MANAGEMENT CHALLENGES AND FRAMEWORK

1. What challenges do you encounter in reference to RDM?

……

2. What appropriate RDM framework can be developed at UNAM to ensure effective management of research data in the whole research life cycle?

……

Thank you very much for your time!!

APPENDIX D

INFORMED CONSENT INFORMATION

Dear participant

My name is Ronald Kanguti, studying towards a Masters of Arts in Library and Information Science. My study is titled: **"Developing a framework for research data management capabilities at the University of Namibia"**. Based on a bibliometric analysis of UNAM research productivity, visibility and impact on Scopus and Web of Science, you were established to be one of the top publishers from your faculty. Therefore, you were purposefully selected based on your relevance to the objectives of the study.

Purpose of the study

This study seeks to enlighten UNAM management about Research Data Management (RDM) and the importance of an RDM framework. The study further aims to develop an RDM framework that takes into account the whole research process at UNAM. The RDM framework will give a detailed plan on how to implement RDM at UNAM and specifying the role of each stakeholder.

Your responses to the study will be anonymous and your identity will not be revealed. If you choose to participate, no one will be able to identify you, therefore no one will know whether you participated in this study.

You are free to withdraw at any time after signing the consent form. If you have any question regarding the research, please feel free to contact me at ronaldkanguti@gmail.com or my supervisor Prof Kalusopa at tkalusopa@unam.na

Thanking you in advance

Ronald N. Kanguti,

CONSENT FORM

I have read and understand the information provided and had an opportunity to ask questions. I understand that my participation is voluntary and that I am free to withdraw from the study at any time without giving any reason and without cost.

Voice recording

I understand that my voice will be recorded and I voluntarily give consent for my voice to be recorded during face to face interview, telephonic interview or through other platform such as Zoom or WhatsApp.

I understand that anonymity will be ensured in the write-up by disguising my identity.

Participants signature……. Date……

Researchers signature……. Date……

APPINDEX E

DRAFT RDM POLICY FOR UNAM CONSIDERATION

Purpose

The purpose of this framework is to provide guidance to UNAM staff, researchers and students on how they should manage and share research data that they produce in order to maximize the potential for research produced at UNAM.

Background

UNAM recognises that as a large public funded institution, it has an obligation to share its research findings and scholarly outputs with all stakeholders and the wider society and has identified that the open access model of scholarly communication is a means to advance research.

Definition of keywords

Research data

The evidence that underpins the answer to the research question, and can be used to validate findings regardless of its form (e.g. print, digital, or physical). These might be quantitative information or qualitative statements collected by researchers (staff, students or those supporting research) in the course of their work by experimentation, observation, modelling, interview or other methods, or information derived from existing evidence. Data may be raw or primary (e.g. direct from measurement or collection) or derived from primary data for subsequent analysis or interpretation (e.g. cleaned up or as an extract from a larger data set), or derived from existing sources where the rights may be held by others.

Data set

Is an interpreted presentation of information suitable for communication, interpretation or processing such as bits, table of numbers, sound recordings etc.

Data management plan (DMP)

Is a document outlining how data will be handled throughout the research process?

Format

Refers to a specific method in which information is encrypted and stored on a computer disk.

Metadata

Discusses the structured information used to describe research data and associated documentation attributes. Metadata aids in the discovery and retrieval of data and is used for citation and referencing.

Researcher

Refers to the creator or the person who originated or gave existence to the research data such as UNAM researchers, staff and students.

Data repository

Is a digital archive responsible for collecting, preserving and disseminating digital objects of research data?

Open science

Also known as open access, is a movement primarily concerned with making scientific research and data accessible to all, for public use. It includes practices such as publishing

open scientific research, campaigning for open access and generally making it easier to publish and communicate scientific knowledge.

Open data

Refers to research data that is freely available permitting any user to download, copy, analyse, re-process, pass to software or use without financial, legal or technical barriers other than the requirement for attribution and share alike.

UNAM RDM stakeholders

RDM in an organisation involves a number of key role players to ensure success: Major stakeholders at UNAM involves:

UNAM management

Research data can be seen as a UNAM asset based on UNAM's vision, which supports research. Responsible data management strategies at UNAM, and the resulting access to research data, can contribute to an improved public understanding of UNAM research and thereby of its contributions to the public good which can enhance future research fundings.

UNAM has pronounced itself on research data management through its research ethics policy whereby there is a highlight on data sharing and storage.

Centre for Research Services

UNAM through the Centre for Research and Publications (CRP) provide coordination and support services aimed at fostering research activities among staff members for them to effectively conduct and publish their scholarly work. CRP is also at the centre of linking UNAM with other institutions of higher learning, research institutions and strategic collaborators globally.

Centre Information Communication Technology Services

Technology enhances the reach of research through cyberinfrastructure. UNAM subscribes to Dropbox and available technologies should ensure support for data acquisition, storage, security and preservation. Existing Information Communication Technology (ICT) infrastructure should be coordinated to support the management of research data and addressing any possible gaps.

UNAM Library

The Library is a key player in data management, curation and preservation based on expertise on selection, metadata description, institutional repositories, preservation strategies and access platforms.

UNAM researchers

Researchers are the focal point as they are producers of research data. Researchers are presented by different RDM and open access requirements mandated by funding agencies, national laws and universities laws. They need to be well vested with RDM capabilities so that they are able to negotiate agreements to determine ownership of data.

Benefits of RDM to UNAM

- Ensure compliance with research funders requirements
- Increase citations and recognition
- Facilitates sharing and re-use of research data
- Promotes research integrity and validation of research results
- Ensures compliance with research data expectations of funders and publishers

Data sets and preparation

- Research data supporting the findings of UNAM research outputs is curated and managed for data sharing and preservation
- Preparation of data for deposition into the data repository and submitted with embargos
- The data needed to validate the results presented in scientific publications and which cannot be re-generated
- Data underpinning a master's thesis and doctoral dissertation.

Restrictions/ embargoes

Publicly funded research data is in the public domain, with free and open access requirements. However, although the research is publicly funded, there are a number of criteria to withhold research data:

- When authentic reasons are provided on why data should not be in an open repository
- Restrictions to protect confidential data
- Data that can harm the interest of community
- When it is impossible to archive such data in the data repository (e.g. big data)
- Sensitive or health related data

How long should data be kept?

Research data should be kept for long-term use and must be described so that it is retrievable, accessible and re-usable.

- The duration can vary depending on the discipline and research type.
- Research data must be kept for 10 years; after which it will be appraised for further retention.

- Preserved data must be stored securely to protect against theft, misuse, damage or loss and such data should be preserved in appropriate facilities.

Types of data

- Textual materials
- Databases
- Datasets
- Spread-sheets
- Software e.g. programming languages, operating systems, etc.
- Curriculum materials
- Digital images
- Audio and video formats
- Physical collections
- Models
- Experimental and observational data.

Data formats

Data formats are critical in the preservation of research data. Data should be preserved in current and available formats that guarantee long-term access and that are interoperable.

Type of data	**Durable formats (list not comprehensive)**
Documentation	Rich Text Format (.rtf) PDF/A or PDF, HTML, Open Document Text (.odf) Plain text (.txt) XHMTL 1.0 Widely-used proprietary formats such as Word (.doc) and Excel (.xls)

Quantitative/tabular data	SPSS portable (.por), Stata, SAS , SPSS DDI XML comma-delimited (.csv), tab-delimited (.tab), MSExcel (.xls/.xlsx), MSAccess (.mdb/.accdb), dBase (.dbf) OpenDocument Spreadsheet (.ods)
Qualitative data (Text)	Extensible mark-up language (.xml), Rich Text Format (.rtf), plain text ASCII (.txt), hypertext mark-up language (.html) NUD*IST, NVivo and ATLAS.ti
Digital video data	JPEG 2000
Digital image data	TIFF (version 6) uncompressed JPEG (.jpeg, .jpg) TIFF (other versions) Adobe Portable Document Format (PDF/A or PDF) raw image format (.RAW)
Digital audio data	Free lossless audio Codec (.flac) WAV (.wav) MPEG-1 Audio Layer 3 (.mp3) also OK
GIS and CAD (vector & raster)	ESRI Shapefile (.shp), GeoTIFF (geo-referenced TIFF) CAD data (.dwg) Binary formats of GIS and CAD also suitable

Table 5.1: Data formats

Copyright Licensing

Copyright entails that the copyright owner has to grant the rights to others in order to reproduce, migrate or communicate the dataset to the public. Consequently, the copyright owners must grant the data managers with rights to store their datasets on a repository. In a scenario whereby the dataset is open to the public, the copyright owners should grant the

rights to end users to use their data. This licence is usually attached to the dataset in the repository so end-users can determine limits to use granted by the copyright owner.

Quality assurance and quality control

Any ethical, privacy, copyright issues and intellectual property will be addressed. To exchange unpublished data during the runtime of the project, all data sets may be password protected. The original creator of the requested data will approve consent before such data can be shared.

Priority areas	**Initiatives**	**Action**	**Milestones**
Introducing RDM cultural change	1.1 Establish an RDM task force at UNAM	Institute a team consisting of all RDM stakeholders inclusive of APVC-RID, Research and Publications office, Computer Centre, Library, Grants Office Clearly specify re the role of each stakeholder	A task force assembled to overlook RDM development and establishment at UNAM
	1.2 Integrate RDM into ethical clearance at UNAM	Work with existing ethical committees to explore existing structures Propose new ideas and processes	Demonstrate understanding of RDM for ethical clearance
	1.3 Promote re-use of UNAM research data	Promote data management planning, publishing and re-use	Growing number of new research outcomes
	1.4 Incorporate data management planning into internal grant application	Explore RDM costing models for grant applications	Clarification of RDM importance to researchers

RDM Institutional policy framework	2.1 Develop a harmonised RDM policy framework	Engage all relevant stakeholders at UNAM Simplify each stakeholder's responsibilities Seek approval from UNAM Council	A tangible UNAM RDM policy approved and supported
	2.2 Awareness raising of RDM policy across UNAM	Strategize on how to best market the policy to all UNAM staff and students Establish strategies for monitoring compliance	A clear framework for monitoring compliance
Technical infrastructure for RDM	3.1 Align current technology and storage infrastructure with RDM requirements	Explore current and available technologies and integrate RDM requirements Identify RDM requirements in terms of storage capabilities and align current technologies to accommodate RDM	Institutional infrastructure integrated with RDM capabilities and functionality
	3.2 Work towards developing a coordinated RDM Infrastructure	Investigate institutional RDM resources required at UNAM Explore solutions to support the preservation and management of born-digital and	IT infrastructure to facilitate RDM at UNAM is developed
		physical research data at UNAM	
	3.3 Develop or acquire platforms to facilitate data deposit, access, storage etc.	Guarantee a data registry equipped with all the required processes at UNAM	Develop UNAM registry of institutional data

RDM skills and knowledge development	4.1 Develop and enhance RDM skills and knowledge of UNAM support staff	Introduce RDM training programmes at UNAM Offer free RDM online training	Capability to support and guide UNAM support staff on RDM related matters Digital curation capabilities
	4.2 Develop and enhance UNAM researchers and students' RDM skills and knowledge	Collaborate with faculties and departments to entrench RDM know-how into researchers and students Avail RDM opportunities for researchers and students to learn	Roll out skills and knowledge development programme for UNAM researchers and students
RDM support services at UNAM	5.1 Enact and promote RDM support services at UNAM	Establish online RDM platforms accessible to all UNAM users Enable engagement between those with the know-how to assist the needy.	The availability of an RDM secure platform
	5.2 Establish new and simple RDM services at UNAM	Ensure storage services for RDM at UNAM Provide dedicated institutional email address for contact	Services available for training and data management planning support
		persons for RDM support	
Funding	6.1 Introduce RDM programme into institutional planning at UNAM	Make a compelling business case for RDM that is aligned with UNAM strategy Develop an RDM business plan Identify and include departmental requirements for RDM institutional planning	Draft business plan for RDM Institutional requirements for RDM is included in institutional planning

	6.2 Explore strategies for cost recovery and RDM financing	Investigate RDM cost and available funding models to support long-term preservation and RDM curation	Clearly outlined cost variable for RDM Information on RDM costing included funding applications

Table 5.2