

**NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY**  
**FACULTY OF COMMUNICATION AND INFORMATION SCIENCE**  
**DEPARTMENT OF LIBRARY AND INFORMATION SCIENCE**

**EXPLORING THE FEASIBILITY OF OFFERING RESEARCH DATA  
MANAGEMENT SERVICES AT THE UNIVERSITY OF ZIMBABWE LIBRARY**

**BY**

**NANCY NHENDODZASHE  
(N01521888L)**

**A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS OF THE MASTER OF SCIENCE DEGREE**

**IN**

**LIBRARY AND INFORMATION SCIENCE**

**ACADEMIC SUPERVISOR**

**MR N. PASIPAMIRE**

**BULAWAYO, ZIMBABWE**

**YEAR 2017**

**NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY**

**COPYRIGHT DECLARATION**

**NAME OF AUTHOR:** Nancy Nhendodzashe

**TITLE OF PROJECT:** Exploring the feasibility of offering research data management services at the University of Zimbabwe Library

**DEGREE PROGRAMME:** Master of Science Degree in Library and Information Science

**YEAR THIS DEGREE GRANTED:** 2017

Permission is hereby granted to the National University of Science and Technology Library to reproduce copies of this project and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights and neither the project nor extensive extracts from it may be reprinted or otherwise reproduced without the author's written permission.

**SIGNATURE:** .....

**RESIDENTIAL ADDRESS:** 843 Manresa Park  
Greendale, Harare

**REGULAR EMAIL ADDRESS** [nancie1987@gmail.com](mailto:nancie1987@gmail.com)

**DATE:** .....

**NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY**

**DECLARATION ON PLAGIARISM**

I am aware of and understand the University's policy on plagiarism. I hereby declare that this research project is the result of my own independent scholarly work, and that in all cases material from the work of others is properly acknowledged. Quotations and paraphrases are clearly and properly referenced. Referencing in this project was done according to the Department of Library and Information Science citation guidelines.

I have not sought or used the services of professional agencies to produce this research project. This written work has not previously been used as examination material at this or any other university and the work has not yet been published. In addition, I understand that any false claim in respect of this work will result in disciplinary action in accordance with University regulations

**NAME**

**NANCY NHENDODZASHE**

**SIGNATURE .....**

**DATE.....**

**NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY**

**RELEASE FROM**

I certify that I have read and recommended to the National University of Science and Technology for acceptance; a research project entitled “**Exploring the feasibility of offering research data management services at the University of Zimbabwe Library**”, submitted by Nancy Nhendodzashe, in partial fulfilment of the requirements of the Master of Science Degree in Library and Information Science.

I further certify that she attended all meetings that were scheduled with me and that all requirements were fulfilled. It is my professional judgment that the project is of a suitable standard as to be submitted with my name attached to it as the supervisor.

I therefore release the student to submit her dissertation for marking.

**Mr N. Pasipamire**

**Supervisor**

.....

**Signature**

.....

**Date**

## **DEDICATION**

To my husband Talent Kwangwa, our son Brandon Kwangwa and to my late parents.

## **ACKNOWLEDGEMENTS**

I would like to thank all the people who made this Masters programme come to fruition. Firstly, I am grateful to my supervisor, Mr Notice Pasipamire, for his help, insightful comments and optimism throughout this project. Many thanks to the lecturers and staff from the department of Library and Information Science at NUST for their contribution in the completion of this Masters programme and for their generous comments during the conception of this study.

I would like to extend my sincere gratitude to the UZ Library and Community for taking their time to participate in this study. Their contribution is forever cherished.

I am grateful to my husband Talent Kwangwa for the support and encouragement.

I am highly indebted to my family and friends who supported and encouraged me throughout this study. Mbuya mai Tanya and uncle Sydney, thank you for being there for Brandon when I was attending classes. To my friend Grace Msauki, thank you for the help you rendered in this study.

Last but not least, I am grateful to God for giving me the strength and provisions to accomplish this Masters programme.

## **ABSTRACT**

Research data management services are being established in academic libraries in response to the changing scientific research landscape. To this end, this study explored the feasibility of offering research data management (RDM) services at the University of Zimbabwe (UZ) library. The study was guided by the Technological, Economic, Legal, Organisational and Schedule feasibility model and the UK Data Archive Research Data Lifecycle model. The study adopted the mixed methods research design and the case study research strategy. Data was collected through a questionnaire from 104 researchers who were selected using stratified random sampling. Semi-structured interviews were conducted with nine (9) faculty librarians, the university librarian and the library ICT manager who were selected using purposive sampling. The study findings showed that the UZ library was not prepared to offer RDM services. It was also revealed that in terms of the technological infrastructure, the library can expand existing institutional repositories to incorporate RDM services. The study showed that the economic needs for setting up RDM services include staff training, costs of acquiring ICT resources and advocacy costs. The legal obligations for setting up RDM services include institutional RDM policy, copyright and data protection policies. It was also found that researchers at UZ need RDM services and the faculty librarians were optimistic to take up the new role given that they receive the necessary training. The researcher concluded that given the current technological, economic, legal and organisational capability, the library can begin with research data storage services. The researcher recommended that before implementing RDM services, the UZ library should ensure that faculty librarians are trained on RDM, necessary technology is in place, sustainable economic resources should be secured and enabling policies need to be developed and all the stakeholders at UZ need to be involved from the conception of the project.

## TABLE OF CONTENTS

<b>TITLE.....</b>	<b>i</b>
<b>COPYRIGHT DECLARATION.....</b>	<b>ii</b>
<b>DECLARATION ON PLAGIARISM.....</b>	<b>iii</b>
<b>RELEASE FORM.....</b>	<b>iv</b>
<b>DEDICATION.....</b>	<b>v</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>vi</b>
<b>ABSTRACT.....</b>	<b>vii</b>
<b>LIST OF TABLES .....</b>	<b>xi</b>
<b>LIST OF FIGURES .....</b>	<b>xii</b>
<b>LIST OF APPENDICES .....</b>	<b>xiii</b>
<b>LIST OF ACRONYMS AND ABBREVIATIONS .....</b>	<b>xiv</b>
<b>CHAPTER 1 .....</b>	<b>1</b>
<b>INTRODUCTION.....</b>	<b>1</b>
1.1 Background of the Study .....	1
1.2 Theoretical Framework.....	5
1.2.1 The TELOS Feasibility Framework .....	5
1.2.2 The UK Data Archive Research Data Lifecycle Model .....	6
1.2.3 Conceptual Framework.....	6
1.3 Statement of the Problem .....	8
1.4 Purpose of the Study.....	8
1.5 Research Objectives .....	8
1.6 Research Questions.....	8
1.7 Significance of the Study.....	9
1.8 Assumption of the Study .....	10
1.9 Definitions of Terms.....	10
1.10 Delimitation of the Study .....	11
1.11 Limitations of the Study .....	11
1.12 Summary.....	12
<b>CHAPTER 2.....</b>	<b>13</b>
<b>LITERATURE REVIEW .....</b>	<b>13</b>
2.1 Introduction .....	13
2.2 Definition of Research Data Management .....	13
2.3 Research Data Management in Academic Libraries .....	14
2.3.1 Attitudes of Librarians Towards RDM.....	15
2.3.2 Areas of RDM Responsibility .....	16
2.3.3 The Significance of Research Data Management in Academic Institutions .....	17



2.4 Theoretical Framework.....	19
2.4.1 The TELOS Feasibility Framework .....	19
2.4.2 The UK Data Archive Research Data Lifecycle model .....	20
2.4.3 Conceptual Framework.....	22
2.5 Technological Requirements for Research Data Management .....	23
2.5.1 Hardware and Software Requirements for RDM.....	24
2.5.1.1 Research Data Types Generated by Researchers .....	24
2.5.1.2 Quantity of Research Data Generated by Researchers .....	25
2.5.1.3 Research Data Storage Devices .....	25
2.6 Economic Needs for RDM .....	27
2.6.1 Cost/Benefit analysis for RDM .....	27
2.6.2 Costs for setting up RDM infrastructure.....	29
2.7 Legal Obligations for RDM.....	32
2.7.1 Ownership of Research Data .....	34
2.7.2 Ethical Considerations for RDM .....	34
2.7.3 Institutional Policies for RDM.....	35
2.8 Organisational Capabilities for RDM.....	36
2.8.1 Skills and Knowledge Requirements for RDM .....	36
2.8.2 Researcher’s Attitudes Towards RDM .....	40
2.8.3 Institutional Responsibility for RDM .....	41
2.8.4 Management Support for RDM .....	42
2.8.5 Stakeholders in RDM .....	43
2.9 Scheduling RDM Services.....	44
2.10 Summary.....	46
<b>CHAPTER 3 .....</b>	<b>48</b>
<b>RESEARCH METHODOLOGY .....</b>	<b>48</b>
3.1 Introduction .....	48
3.2 Research Philosophy.....	48
3.3 Research Approach.....	49
3.4 Research Strategy .....	51
3.5 Study Population.....	53
3.6 Sampling and Sampling Procedure.....	53
3.7 Data Collection Instruments .....	56
3.7.1 Questionnaires .....	56
3.7.2 Interviews .....	58
3.8 Reliability and Validity .....	59
3.9 Ethical Considerations.....	60
3.10 Data Collection, Presentation and Analysis Procedures.....	61
3.11 Summary.....	61
<b>CHAPTER 4 .....</b>	<b>63</b>
<b>DATA PRESENTATION AND INTEPRETATION .....</b>	<b>63</b>
4.1 Introduction .....	63

4.2 Background of Respondents and Response Rate .....	63
4.4 Technological Requirements for RDM .....	65
4.4.1 Types of Data Generated .....	65
4.4.2 Data Storage Devices Used .....	66
4.4.3 Quantity of Research Data Generated .....	67
4.4.4 Quality of ICT Resources Provided by the UZ.....	68
4.5 Economic Needs for RDM .....	70
4.5.2 Costs Incurred in RDM.....	70
4.5.3 Benefits of Research Data Management.....	71
4.5.5 Risks in RDM .....	73
4.6 Legal Obligations for RDM.....	75
4.6.1 Research Data Ownership .....	75
4.6.2 Research Data Management Policies.....	76
4.7 Organisational Capabilities for RDM.....	77
4.7.1 Knowledge of Research Data Management.....	77
4.7.2 Researchers' Attitudes Towards RDM .....	78
4.7.3 Responsibility for RDM .....	82
4.7.3 Skills and Competences Needed for RDM .....	84
4.7.4 Roles that Librarians can Take Up in RDM .....	84
4.7.5 Management Support to RDM Services .....	85
4.8 Scheduling RDM services at the UZ .....	85
4.9 Data Interpretation and Discussion.....	87
4.9.2 Economic Needs for RDM .....	89
4.9.3 Legal Obligations for RDM.....	91
4.9.4 Organisational Capabilities for RDM.....	92
4.9.5 Scheduling RDM Services.....	95
4.10 Summary.....	95
<b>CHAPTER 5 .....</b>	<b>97</b>
<b>SUMMARY, CONCLUSIONS AND RECOMMENDATIONS .....</b>	<b>97</b>
5.1 Summary.....	97
5.2 Conclusions .....	98
5.3 Recommendations .....	100
<b>REFERENCES.....</b>	<b>103</b>

## LIST OF TABLES

<b>TABLE</b>	<b>DESCRIPTION</b>	<b>PAGE</b>
Table 1.1	Subcomponents of the TELOS Framework	6
Table 2.1	Librarians' roles in RDM and required competencies mapped to existing roles	38
Table 3.1	Sampling frame for researchers	56
Table 3.2	Sample sizes from each department	56
Table 4.1	Response rate by faculty, highest qualification, period of service and gender	65
Table 4.2	Knowledge of RDM by researchers at the UZ	78
Table 4.3	Researchers who require RDM Services	79
Table 4.4	Data re-use by researchers	83
Table 4.5	How RDM services can be implemented	87

## LIST OF FIGURES

<b>FIGURE</b>	<b>TITLE</b>	<b>PAGE</b>
Figure 1.1	Conceptual Framework: Application of the TELOS framework to the revised UK Data Archive Research Data Lifecycle	7
Figure 2.1	UK Data Archive Research Data Lifecycle	21
Figure 2.2	The KRDS Benefits Framework	30
Figure 2.3	Costing Research Data Management	33
Figure 2.4	Core skills for data management	41
Figure 2.5	Stakeholders in RDM development	45
Figure 4.1	Research data formats generated by researchers at the University of Zimbabwe	66
Figure 4.2	Data storage devices used by researchers at UZ	67
Figure 4.3	Estimated average volume of research data generated by UZ researchers	68
Figure 4.4	Rating of UZ IT resources by researchers	69
Figure 4.5	Benefits of RDM	73
Figure 4.6	Research data loss	75
Figure 4.7	Consequences of research data loss	76
Figure 4.8	Ownership of researchers' research data at the UZ	77
Figure 4.9	Researchers' interests in RDM services	80
Figure 4.10	Research data sharing	81
Figure 4.11	Reasons for not sharing research data	82
Figure 4.12	Primary responsibility for RDM at the UZ	84
Figure 4.13	When researchers require RDM services	86

## **LIST OF APPENDICES**

- Appendix A – Questionnaire for researchers
- Appendix B – Interview guide for faculty librarians
- Appendix C – Interview guide for university librarian
- Appendix D – Interview guide for the library ICT manager
- Appendix E – Letter of authorisation to undertake research  
from the University of Zimbabwe Librarian

## **LIST OF ACRONYMS AND ABBREVIATIONS**

ACRL	Association of College and Research Libraries
ARL	Association of Research Libraries
ICT	Information and Communication Technology
IT	Information Technology
KRDS	Keeping Research Data Safe
LIS	Library and Information Science
RDM	Research Data Management
RDS	Research Data Services
TELOS	Technology, Economic, Legal, Organisational and Schedule
UK	United Kingdom
UZ	University of Zimbabwe

# CHAPTER 1

## INTRODUCTION

### 1.1 Background of the Study

Academic libraries are increasingly interested in supporting the activities of researchers and assessing the quality of what they produce, since research performance is a key determinant of each university's reputation and financial success (Jubb, 2016). Research is identified as one of the strategic thrusts of the University of Zimbabwe (UZ) in its 2016-2020 strategic plan (UZ, 2016). Data are the currency of research; but analogue and digital research data generated within academia have largely been an invisible resource utilised within the research unit and shared with a select group of trusted colleagues, and consequently their management is poorly understood (Koopman and Jager, 2016). Research data management (RDM) refers to the collection, storage, access and preservation of data produced in particular investigations or research projects (Chiwere and Mathe, 2015). The "data deluge" generates a need to develop policies, infrastructures and services in institutions to manage data, with the aim of assisting researchers in creating, collecting, manipulating, analysing, transporting, storing and preserving datasets (Tenopir et al., 2014). Libraries need to provide support for the complete research cycle, and therefore, they need to analyse what researchers require to manage their data from creation or compilation to archive and preservation. In this context, many university libraries are seeking to enhance their support for research (Digital Curation Centre, 2016). Hence they are rethinking the roles of their collection-based services; changing the roles of liaison librarians; and developing new services for researchers including advice on scholarly communications and open access, bibliometrics services, research data management and library-led publishing services.

The University of Zimbabwe grew tremendously since its inception in terms of research and academic excellence. The university started with the faculties of Arts, Education, Science and Social Studies but currently there are six additional faculties namely Commerce, Engineering, Law, and Veterinary Science and the College of Health Sciences (University of Zimbabwe, 2015). As a result of these developments in academic programmes some of which use

sophisticated and specialised computing programmes, vast quantities of born-digital research data are being produced in a wide variety of forms and at a rapid rate which Pinfield et al. (2014) referred to as creating the so-called “volume”, “variety” and “velocity” challenges of data. Apart from research which is conducted for academic purposes, there is commissioned research by different funding organisations which is done at the UZ. Speaking at the 2015 UZ graduation ceremony, the UZ Vice Chancellor Professor Nyagura reported that the institution holds many research leadership positions where it is trusted by partner institutions and donors to lead multi-million dollar research projects. Most of these donors mandate grantees to have data management plans. Tenopir et al. (2015) noted that the growth of data intensive science, coupled with funding mandates for data management plans and government open data, has led to a growing emphasis on data management across all academic disciplines. Henderson and Knott (2015) emphasised that as the institutions and granting agencies began to understand the value of all the computer-generated data for reuse or discovery and the need for a standardised way to store and manage data became apparent.

Research data management services are being established in response to the changing scientific research landscape and being supported both by the existence of cyber infrastructures and data sharing mandates by funding agencies and researchers who are committed to open science, which advocates for open verification and reproduction of research data (Naum, 2014). The last several decades of network- and computer-enabled work in science have produced untold amounts of data, leading to the challenge of developing practices to manage and provide access to this data (Gold, 2007). Science is said to have entered a “fourth paradigm” which is more collaborative, more computational and more data intensive than the previous experimental, theoretical, and computational paradigms (Tenopir et al. 2014). This changing nature of scientific research has prompted research and academic institutions to respond by investing in systems to leverage the research data being generated at an exponential rate.

The researcher works at the University of Zimbabwe library and has observed that the institution has been producing a lot of research in different academic disciplines. The UZ library has set up systems to capture and preserve research outputs from the institution. These



include the institutional repository which collects and preserves open access research output from the University of Zimbabwe, the UZ digital collections which contains digitised publications about the UZ, the UZ Online Maps Repository which contains scanned maps and the UZ Exams database which contains past examination papers. However, there is no system for research data management. The researcher received requests for research data from library users. One researcher complained that there are no research data management services at the institution and identifying trends was very difficult without making reference to the original data sets. The researcher observed a case where data was lost before it was processed. These circumstances triggered the researcher to conduct a study on research data management.

Starting a new data management service is much like starting any new service; there is need for staff who are willing to take on new roles and responsibilities, need to identify the members of the community who could benefit from the service, and there is need to be sure that what is being offered meets the needs of the users. Strasser (2014) pointed that the struggle for resources in libraries is a well-known issue, so investing time, energy, and money into something new must be done thoughtfully and with a good balance of caution and initiative. A study was conducted by Sykes (2009) to establish the feasibility of a national shared service for managing research data in the United Kingdom. This study was conducted to find out the technological requirements for the entire data lifecycle. Chiware and Mathe (2015) conducted a study to establish the requirements for setting up RDM services at the Cape Peninsula University of Technology (CPUT) Library. Der (2015) explored the academic libraries' readiness for research data management from Hungary and Estonia academic libraries. Regrettably, a few studies on RDM were done in Zimbabwe with one which was done by Nera in 2015 where he looked at the management of research records at Africa University and Boriwondo in 2016 looked at research data management services at the Women's University in Africa. Nera (2015) concluded that research records were not systematically managed at Africa University and there was need to come up with a strategy of managing research records. Ndhlovu (2016) conducted a study on the preparedness for digital curation and preservation at the National University of Science and Technology. He concluded that there were inconsistencies in digital curation and ICT competences were low among some library staff. No studies were conducted at the UZ on the feasibility of offering RDM services.

Therefore, there is need for a study that will not only address this gap, but which would act as a archetypal for assessing the feasibility of offering RDM services in Zimbabwe.

Research data management is a complex issue involving multiple activities carried out by various actors addressing a range of drivers and influenced by a large set of factors (Pinfield et al., 2014). Therefore, Lotter (2014) opined that before setting up any RDM services, the infrastructure and investment need to be carefully assessed by each institution in line with their own mission, objectives and strategic aims. Establishing a research data management programme has become a pressing imperative for many research libraries, but relatively few have a programme in place. The challenges are many; these include learning about RDM principles and issues, assessing the local institution's greatest needs, selecting and implementing a repository environment, working with researchers to convey the importance of RDM, preparing training materials, building expertise among library staff, and establishing metadata guidelines (Pinfield et al., 2014). The Association of Research Libraries (ARL) E-Science Working Group conducted a survey in 2010 to determine how ARL libraries have been approaching this task of providing RDM to their patrons (Soehner et al., 2010). Although the majority of institutions lacked designated departments for providing research data services (RDS), over one-third had conducted assessments of their researchers' needs for data services. The DataONE 2011 baseline assessment of ACRL library directors identified that well under half of libraries surveyed offered some form of RDM services and they had to do a thorough feasibility study before introducing the services (Tenopir et al., 2013). Reporting on how institutions can start up RDM services, Henderson (2014) highlighted that a needs assessment survey is typically an important part of starting a data programme, although it can be time consuming and expensive.

Offering research data management services is an emerging area in most academic institutions in developing countries particularly in Zimbabwe. The mission of the UZ library is "To provide access to robust high quality scholarly information resources, services and facilities that support cutting edge and innovative research, teaching, learning and knowledge creation." As one of the library's activities to achieve the mission, the University of Zimbabwe library is planning to offer RDM services to capture the ever-increasing research data from the

institution and to help researchers meet the funding mandates (UZ Library, 2016). Just like any other intervention, introducing RDM requires organisations to investigate the needs of researchers, readiness and capabilities of on-campus partners, and available infrastructure on a particular campus. It is imperative to conduct a feasibility study before implementing RDM (Wilkinson et al., 2016). The researcher observed a number of projects which have failed at the UZ library because they did not do a thorough feasibility study. Worthy mentioning is the agricultural information kiosk which never materialised, the e-farming project which did not successfully take off and the mobile library project failed due to technological challenges.

## **1.2 Theoretical Framework**

The study adopted the Technology, Economic, Legal, Organisational and Schedule (TELOS) feasibility model which was first presented by Hall in 2007 (Hall, 2007). The study took RDM stages from the UK Data Archive Research Data Lifecycle model which provides a data lifecycle model as an aid to researchers in identifying activities that are involved in RDM.

### **1.2.1 The TELOS Feasibility Framework**

The framework looks at the Technological, Economic, Legal, Organisational and Schedule aspects when conducting feasibility studies. Bause et al. (2014) pointed that feasibility studies focus on five subjects: technical, economic, legal, organisational and scheduling feasibility studies. The TELOS framework provides a generic approach to feasibility studies therefore, the constructs of the framework were qualified with RDM activities from the UK Data Archive Research Data Lifecycle model. Table 1.1 below shows the subcomponents of each of the constructs of the TELOS framework that were used in this study. The choice of the TELOS framework was inspired by its simplicity and traceable application in the area of RDM by different scholars. The Oxford University used the framework when they were developing the infrastructure for RDM (Wilson et al, 2010). At Griffith University, an operational feasibility was done where they looked at librarians as partners in research data service (Cox et al, 2015).

**Table 1.1 Subcomponents of the TELOS Framework (Researcher)**

Feasibility	Components to be covered
Technological	<ul style="list-style-type: none"> <li>• Hardware and software requirements</li> <li>• Data formats</li> <li>• Quantity of research data produced</li> <li>• Data repository versus licences for data services</li> </ul>
Economic	<ul style="list-style-type: none"> <li>• RDM Cost/Benefit analysis</li> <li>• Risks in RDM</li> </ul>
Legal	<ul style="list-style-type: none"> <li>• Research data ownership</li> <li>• RDM policies</li> </ul>
Organisational	<ul style="list-style-type: none"> <li>• Skills and knowledge required for RDM</li> <li>• Researcher’s attitudes towards RDM</li> <li>• Institutional responsibility for RDM</li> <li>• Management support</li> <li>• Stakeholders for RDM</li> </ul>
Schedule	<ul style="list-style-type: none"> <li>• Implementation of RDM services</li> <li>• When should RDM services provided</li> </ul>

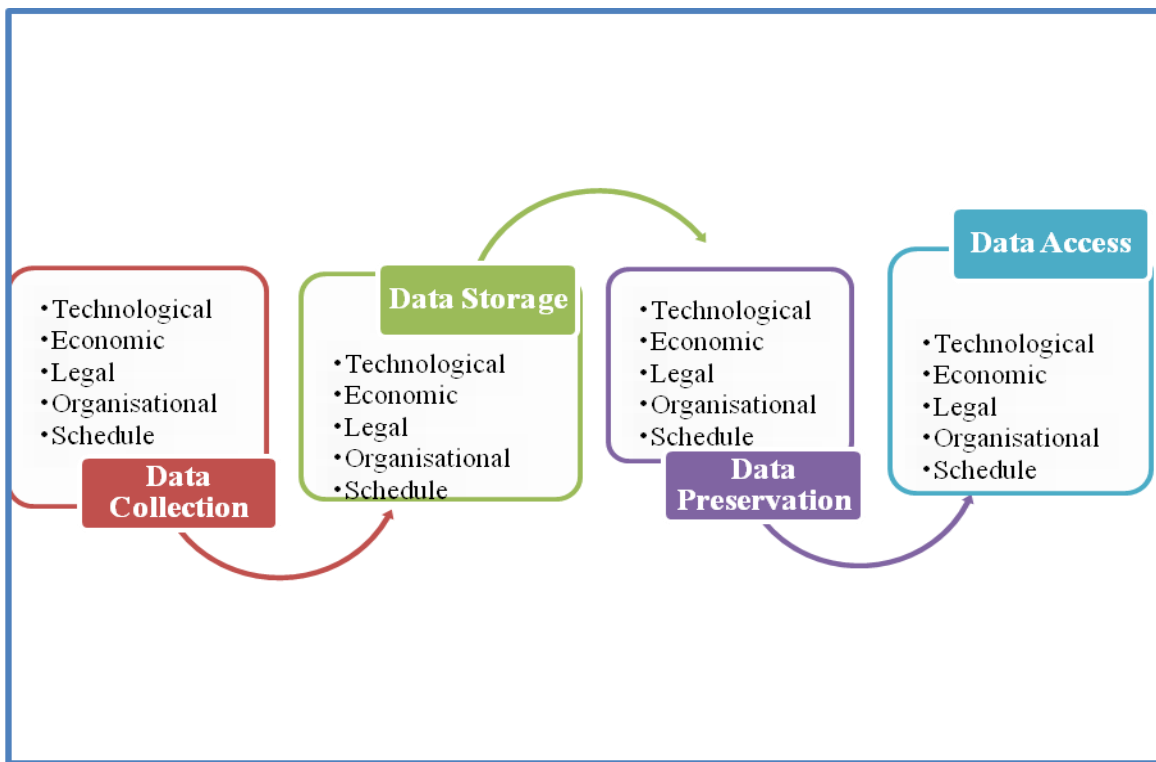
### **1.2.2 The UK Data Archive Research Data Lifecycle Model**

Together with the TELOS model, the researcher used the UK Data Archive Research Data Lifecycle model as a guide to where research data management services can be applied by the UZ library. The model is illustrated herein after in Figure 2.1. The UK Data Archive Research Data Lifecycle provides a data lifecycle model as an aid to researchers considering how data management relates to the lifecycle of a research project. A research data lifecycle describes the stages through which research data is collected, recorded, processed and results are published (UK Data Archive, 2015).

### **1.2.3 Conceptual Framework**

The TELOS feasibility framework is a general feasibility model that can be applied in any discipline. To contextualise the framework, the researcher applied the TELOS framework to selected stages of the research data lifecycle model where the library can assume

responsibility. Research data management services are seen as platforms that bring together the socio-technical elements (Pinfield et al., 2014). The social element refers to the interaction of the people who manage and develop the RDM platform and those who contribute contents in the data repository. The technical aspects refer to those dealing with the technology used to make the contents available and accessible to users. The objectives of this study were informed by the constructs of the framework which will be defined herein after. Figure 1.1 below shows how the TELOS framework was applied to the revised UK Data Archive Research Data Lifecycle model. RDM in academic libraries can be conceptualised as comprising of the following activities; data collection, storage, preservation and access (Wilson, 2010). The researcher picked four stages from the UK Data Archive Research Data Lifecycle model which are; data collection, data storage, data preservation and data access.



**Figure 1.1 Conceptual Framework: Application of the TELOS framework to the revised UK Data Archive Research Data Lifecycle (Researcher)**

### **1.3 Statement of the Problem**

The emerging need for research data management is prompting library directors to plan for RDM services to be offered by their libraries (Tenopir et al., 2014). The UZ library's strategic focus document (2017-2020) states that, one area that is lagging behind is the area of research data management. The library will need to take a leading role at the institution in ensuring that research data is systematically collected, managed and curated for long term preservation and storage. However, is the UZ library ready to offer such services?

### **1.4 Purpose of the Study**

The aim of the study was to explore the feasibility of offering research data management services at the UZ library. The researcher used the TELOS feasibility model guided by the UK Data Archive Research Data Lifecycle Model to bring out the key activities involved in RDM.

### **1.5 Research Objectives**

The study sought to principally fulfill the following objectives:

1. To find out the technological requirements for establishing RDM services at the UZ library;
2. To measure if the UZ library had the necessary economic resources for setting up RDM services;
3. To assess if the UZ library had an enabling legal environment for providing RDM services;
4. To review the organisational capabilities of the UZ library in offering RDM services; and,
5. To find out how RDM services can be scheduled at the UZ library.

### **1.6 Research Questions**

The following research questions were posed:

1. What are the technological requirements for establishing research data management services at the University of Zimbabwe library?
2. What are the economic needs for setting up RDM services at the UZ library?

3. What are the legal obligations for the establishment of RDM services at the University of Zimbabwe library?
4. What are the organisational capabilities of the UZ library in providing RDM services?
5. How can research data management services be scheduled at the UZ library?

### **1.7 Significance of the Study**

Research data management services are multidisciplinary and interdisciplinary in an academic institution therefore, its significance is multifaceted. It is hoped that the results from the study might help the UZ library management in establishing the requirements for setting up RDM services. The researcher will forward the findings and recommendations from the study to UZ library management for consideration in the planning and design of RDM services. RDM is still in its infancy in Zimbabwe so the findings from the study might inform the establishment of research data management services at other institutions of higher education in the country.

It was hoped that the findings of the study might help the University of Zimbabwe in the development of enabling RDM policies.

From a library and information science (LIS) practice standpoint, this study is relevant and timely for academic libraries, which are currently in the midst of scaling up their research support activities. It is hoped that the study findings would bring out the core RDM competences that academic librarians should possess in order to take up the new roles.

It was hoped that the study might contribute to the existing literature in the field of research data management especially in academic libraries. It was observed that there were very few studies in the area of RDM in Zimbabwe; therefore, this study will contribute the body of knowledge. Future scholars might use this research as a basis for further research in the area of research data management services in academic libraries. This need for broadening the horizon by investigating different communities has been articulated by several previous publications, such as Cox and Pinfield (2013), who called for further work “to capture a sense of the pattern in different countries” in the changing context of research data management and its implications to academic libraries.

It was hoped that the study may contribute to the acceptance and use of research data management models by studying the authenticity and pragmatic applicability of the UK Data Archive Research Data Lifecycle model as well as the TELOS feasibility framework because of their use in this study. The adopted conceptual framework might act as a source of reference for those aiming to do further research on this topic.

### **1.8 Assumptions of the Study**

The researcher assumed that research data management is a critical research support activity at the UZ Library.

### **1.9 Definitions of Terms**

In the context of this study:

- i. **Conceptual framework** means a general orientation to a topic using a mix of published literature, personal knowledge and speculations on the kind of relationships that might emerge in the main study (Hammond and Wellington, 2014).
- ii. **Digital curation** means maintaining, preserving and adding value to digital research data throughout its lifecycle (Digital Curation centre, 2016).
- iii. **Faculty librarian** means faculty liaison who work with academics, researchers and students to help them achieve their research, learning and teaching goals, (Lindstrom and Shonrock, 2006).
- iv. **Feasibility** refers to the viability of offering RDM services at the University of Zimbabwe library in terms of technological, economic, legal, and organisational and schedule capabilities for the creation, storage, preservation of and access to research data.
- v. **Feasibility study** refers to a tool for determining whether an institution has what it takes to embark on an intervention (Lowa State University, 2009).



- vi. **Metadata** means a subset of core standardised and structured data documentation that explains the origin, purpose, time reference, geographic location, creator, access conditions and terms of use of a data collection (UK Data Archive, 2011).
- vii. **Researcher** means any staff member, person awarded honorary or academic status by the university, student, or person otherwise associated with the university, who conducts research in the course of employment, study or a formal research affiliation (Charles Darwin University, 2010).
- viii. **Research data** means recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings (Burnham, 2012)
- ix. **Research Data Management** refers to the storage, access and preservation of data produced in particular investigations or research projects (Chiwere and Mathe, 2015).

### **1.10 Delimitation of the Study**

The study was limited to the technological, economic, legal, and organisational and schedule feasibility of offering research data management services at the University of Zimbabwe library. The study looked at the feasibility of offering digital research data storage, preservation and access at the UZ library. The depth of the study was guided by the research objectives. The conceptual framework was drawn from the TELOS feasibility framework and UK Data Archive Research Data Lifecycle model. Due to limited time and resources, participants were drawn from a selected sample of researchers, faculty librarians, university librarian and library ICT manager. The study was conducted at the UZ main campus located at number 630 Churchill Avenue Mount Pleasant Harare.

### **1.11 Limitations of the Study**

Ideally, it would have been proper to conduct a study involving all researchers at the institution. However, due to the limited timeframe and resources the researcher focused on a selected sample from the University of Zimbabwe. Research data management is a complex

issue involving multiple activities carried out by various actors addressing a range of drivers and influenced by a large set of factors (Pinfield et al., 2014). However, the researcher narrowed the scope of RDM services by developing a conceptual framework highlighted in Figure 1.1. There were few studies on RDM in Zimbabwe, so finding literature to refer to at local level was challenging. To counter this limitation, the researcher made reference to literature from a regional perspective.

### **1.12 Summary**

The study was conducted under the background that the UZ library is planning to offer research data management services an area which the library acknowledged was lagging behind. The main aim of the study was to explore the feasibility of offering research data management services at the UZ Library. The background of the study brought to the fore the rationale for conducting feasibility studies for new library projects and for offering RDM at institutions of higher learning. The chapter highlighted the objectives of the study which guided the formulation of research questions. The importance of the study was highlighted where it was hoped that the study might benefit the UZ Library in crafting policies related to RDM, it was also hoped that the study will add to the body of literature in the library and information science discipline particularly research data management. The chapter also reviewed and justified the use of the TELOS feasibility framework and the UK Data Archive Research Data Lifecycle model which were insightful in the development of research objectives. The limitations of the study were highlighted and measures of mitigating them were outlined. The chapter delineated the scope of the study.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Introduction

This chapter provides a review of literature around the themes of research data management in academic libraries thereby laying a foundation for this study. Hammond and Wellington (2014) pointed that a literature review gives an overview of what has been written about a particular field or topic. It covers what has been said, who has said it, and sets out prevailing theories and methodologies. The review of the literature was guided by objectives of the study and the conceptualised TELOS feasibility framework and the UK Data Archive Research Data Lifecycle model highlighted in the previous chapter. The constructs of the models were used in reviewing literature. The TELOS framework informed the main themes of the reviewed literature and the selected stages of the UK Data Archive Research Data Lifecycle guided the RDM activities under technological, economic, legal, organisational and schedule requirements. From the review of literature, the researcher managed to identify trends, gaps, weaknesses in the prevailing literature on research data management thereby finding justification why further study in the field was necessary. The researcher made reference to scholarly peer reviewed electronic journals, textbooks and the World Wide Web.

#### 2.2 Definition of Research Data Management

Before defining RDM, it is essential to define what research data is in an academic environment. On the simplest level, data can be defined as “facts and statistics collected together for reference and analysis” (Oxford University, 2014). Pryor (2012) defines research data as, “...the output from any systematic investigation that involves a process of observation, experiment or the testing of a hypothesis.” Burnham (2012) defined research data as recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings. From these definitions of research data, the commonality is that whenever researchers embark on a research project, there is some form of output that comes out of the research process in form of data.

Research data comes in different formats (Koltay 2016; Surkis, 2015; Burnham, 2012 and National Science Board, 2005). This is supported by Kennan and Markauskaite (2015) who provided that data are heterogeneous because they can take many forms depending on their origins, the research problem being addressed and the discipline of the researcher. Borgman (2007) added that data comes in several varieties, so it can be observational, computational, and experimental. Data also comes from works of art and literature, as well as from artifacts of cultural heritage (Nielsen and Hjørland, 2012). Borgman (2007) provides some examples of data from different scientific disciplines: X-rays in medicine, protein structures in chemistry, spectral surveys in astronomy, specimens in biology, and events and objects in physics. In the life and physical sciences, data are generally gathered or produced by researchers through observations, experiments or by computer modelling. In the social sciences researchers may gather or produce their own data from, for example, interviews, surveys and questionnaires, and observations. It is perhaps most useful to think of data as everything that would be needed to produce a given scientific output.

Research data management can be defined as the collection and organisation of data produced in particular investigations or research projects from its entry to the research cycle taking into account processes and activities involving the design and creation of data, storage, security, preservation, retrieval, sharing, and reuse; bearing in mind technical capabilities, ethical considerations, legal issues and governance frameworks (Chiwari and Mathe, 2015; Tenopir et al, 2015; Pinfield, Cox and Smith, 2014; Whyte and Tedds, 2011; Wilson 2010). Therefore, this means that RDM involves processes from creation to the dissemination of research data hence it requires proper planning in terms of resources and requirements.

### **2.3 Research Data Management in Academic Libraries**

Management of research data is a service area of increasing interest to academic libraries (Chiwari and Mathe, 2015; Tenopir, Birch, and Allard, 2012; Tenopir et al., 2014). Several studies have been done on the library's involvement in offering RDM services (Cox, Pinfield, and Smith, 2014; Corral, Kennan, and Afzal, 2013; Wilson, 2010). The IFLA Journal in December 2016 ran a special issue to gather the latest theory, research, and state-of-the-art practices from libraries that are informing and innovating effective data services from an

international perspective (Witt and Horstmann, 2016). Early contributions, such as those by Delserone (2008), Henty (2008) and Lewis (2010), set out the case for library involvement. Later works, such as Corral (2012) and Cox (2013) further discuss the range of possible roles of librarians in RDM. Lyon (2012) identified a number of opportunities for libraries but also major challenges in developing the capacity and capabilities to carry out RDM. A recent study by Cox et al. (2017) looked at research data service maturity in academic libraries in some UK and Australian institutions revealed that maturity in RDM was becoming evident. Shen and Varvel (2013) described the development of data services at Johns Hopkins University libraries, which provide storage, archiving, preservation, and curation of data. In South Africa, the Cape Peninsula University of Technology (CPUT) library, took a leading role in creating platforms, systems and processes for the management of research data (Chiwere and Mathe, 2015). In light of the growing interest in RDM in South Africa and the need for capacity building in this area, the Library and Information Association of South Africa's Higher Education Libraries Interest Group hosted a RDM workshop, facilitated by the Digital Curation Centre (DCC), UK, in Cape Town in March 2014 (Kahn et al., 2014).

### **2.3.1 Attitudes of Librarians Towards RDM**

One of the early surveys to study the preparation and attitude of librarians towards research data service was undertaken by Tenopir et al (2012). The survey was conducted among 223 librarians of the Association of Research Libraries (ARL) and the findings indicated that although there was very low percentage of libraries involved in RDM offerings, the librarians believed that this was an important service for academic research libraries to render. Similar findings were reported from the survey conducted by Corral, Kennan and Afzal (2013) among 140 libraries in Australia, New Zealand, Ireland and United Kingdom. They also found that RDM service represents a relatively new development in library service offering though there was an interest among the libraries to offer RDM services, with a high proportion of libraries in the process of planning to offer RDM services support. Similarly, in South Africa, a study by Kahn et al. (2014) revealed that 20 of the 41 participants rated their understanding and confidence in supporting the concept of RDM' as 'average'. Van Wyk and Van der Walt (2014) conducted a study on RDM developments at the University of Pretoria where they found that RDM does not exist in any formal system at the institution and data storage was

done in an ad hoc manner. The University of Cape Town (UCT) has established the UCT eResearch Centre to work and partner with researchers to find appropriate IT solutions for their research data activities. The University of South Africa (UNISA) has completed an investigation into RDM practices at UNISA as part of a project to establish data management services (Macanda, Rammutloa and Bezuidenhout, 2015).

### **2.3.2 Areas of RDM Responsibility**

In order to fully understand the library's involvement in RDM, it is important to highlight the areas in which the library can take up responsibility. Corral (2012) argues that managing research data continues to be an emergent area of activity where responsibilities and practices within libraries are generally not yet firmly established. A survey by Richardson et al. (2012) on library research support in Queensland, Australia found that 17% of respondents reported having an active role in mapping research collections and participation in eResearch projects, 17% reported having an advisory role and another 25% are developing strategies, 41% libraries reported having either a limited role or no role at all.

A study by Tenopir et al. (2017) on RDM services in European libraries revealed that European academic research libraries are more likely to offer consultative-type RDM services than hands-on/technological services. Consultative services frequently involve a personal client-librarian relationship and inform the client about such things as how to find information on data management plans, metadata standards, or data citation practices. Similarly, a content analysis of RDM services by USA academic libraries conducted by Yoon and Schultz (2016) revealed that most libraries provided basic services, such as "help/ask librarians" (79.5%), on their data management web pages through online chat, web inquiry forms, or email. The activities currently conducted by the greatest number of libraries (76.8%) are "discussing RDM with others on campus" and involvement in "policy development/planning related to RDM" (66.3%), which may indicate that many libraries are still in the planning stages or that Similarly, Chiware and Mathe (2015) reported that at the CPUT, the library was working with other stakeholders to develop RDM policies. Lotter (2014) highlights RDM roles of libraries in South Africa. The study showed that there was engagement with researchers and students about RDM and RDM training was provided by faculties and library at the Witwatersrand

University, planned RDM capacity building, advocacy, training at UCT and UNISA. From these empirical studies, it can be noted that the lower and slower up-take of technical services compared to consultative services may reflect the fact that these services require a substantial investment in time, resources, and new technical knowledge.

Furthermore, overall responsibilities of academic libraries with regards to research data were described as maintenance and curation, as mentioned by Hey and Trefethen (2003); Swan and Brown acknowledged the need to archive and preserve research data. Brandt clearly pointed out that librarians need to get involved in data management and related activities: it has become obvious that there is an urgent need for the knowledge that librarians have: the ability to collect, organise, describe, curate, archive, and disseminate data and information (Brandt, 2007). More specifically, Carlson (2012) observed that librarians can further engage with researchers as data interviewers in order to investigate the research process and develop understanding of data practices and workflows

### **2.3.3 The Significance of Research Data Management in Academic Institutions**

Properly managed and shared data have the potential to yield manifold benefits throughout and beyond the life of a research project when reused in primary research, follow-up, and synthesis studies, as well as in interdisciplinary and data-intensive research (Jones, Pryor and Whyte, 2013; Pryor 2012; Heidorn, 2011). The authors further stated that well-managed data lead to higher-quality research, increased visibility and the consequent benefits of enhanced citation rates. Justifying the need for RDM services, Jones, Pryor and Whyte (2013) postulate that researchers need to exchange data across diverse platforms and demand effective systems to store, access and share data securely across different institutions.

Accurate and complete research data are an essential part of the evidence necessary for evaluating and validating research results and for reconstructing the events and processes leading to them (University College Dublin Library, 2015). Additionally, RDM enables research continuity through secondary data use. Good research data management permits new and innovative research to be built on existing information. So the importance of research data quality and provenance is paramount, particularly when data sharing and re-use is becoming

increasingly important within and across disciplines. Sharing well-managed research data and enabling others to use it will also help to prevent duplication of effort (Tenopir et al., 2015).

Lewis (2010) provided a list of the rewards of managing research data. These include:

1. The ability to share research data, minimising the need to repeat work in the laboratory, field or library
2. Ensuring that research data gathered at considerable cost is not lost or inadvertently destroyed
3. The retrieval, comparison and co-analysis of data from multiple sources can lead to powerful new insights
4. The ability to check or repeat experiments and verify findings, particularly important amid growing national and international concern about research integrity
5. New research themes – and in particular cross-disciplinary themes – can merge from re-analysis of existing data or comparisons with new data: increasingly data may become the starting point for new research as well as representing an output from current research.

As the currency of science and essential to scientific productivity, collaboration, and to discovery itself, the sharing of research data has long been a commonly acknowledged concept among many research communities (Gold, 2007). Borgman (2007) provides four rationales for sharing research data, namely “to (a) reproduce or verify research, (b) make results of publicly funded research available to the public, (c) enable others to ask new questions of extant data, and (d) advance the state of research and innovation.”

Despite these varied arguments in favour of RDM, there are relatively low levels of data sharing currently occurring (Pryor, Jones and Whyte, 2014). There are also marked disciplinary differences in the level of sharing, partially caused by the diverse types of materials used in different disciplines and significantly, given the costs involved in preparing data for public release and the availability of funding in different disciplinary areas (Borgman, 2012).



## **2.4 Theoretical Framework**

There is a range of assessment methodologies that can be used by academic libraries to capture RDM requirements associated with data and current research practice (Lyon, 2012). One of the early tools to be developed was the Data Asset/Audit Framework (DAF) which was developed by HATII at the University of Glasgow, in 2008 to provide organisations with the means to identify, locate, describe and assess how they are managing their research data assets. The UK Digital Curation Centre has developed the Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO) which provide a maturity-based assessment of research data collections and practices. The three-legged stool digital asset management model from Anne R. Kenney at Cornell University has been used with foundational elements of technology, organisation and resources. These models provide RDM maturity-based assessment and they do not provide the capabilities to assess institutional readiness for RDM. Therefore, the study adopted the Technology, Economic, Legal, Organisational and Schedule (TELOS) feasibility model which was first presented by Hall in 2007. The study also took RDM stages from the UK Data Archive Research Data Lifecycle model which provides a data lifecycle model as an aid to researchers in identifying activities that are involved in the RDM.

### **2.4.1 The TELOS Feasibility Framework**

The TELOS feasibility framework looks at the Technological, Economic, Legal, Organisational and Schedule aspects when conducting feasibility studies. Bause et al. (2014) pointed that feasibility studies focus on five subjects: technical, economic, legal, organisational and scheduling feasibility studies. The constructs of the TELOS framework will be explained in detail herein after in the context of RDM.

The choice of the TELOS framework was inspired by its simplicity and traceable application in the area of RDM by different scholars. The Oxford University used the framework when they were developing the infrastructure for RDM (Wilson et al., 2010). At Griffith University, an operational feasibility was done where they looked at librarians as partners in research data service (Cox et al., 2015). The TELOS framework was used during the building of research

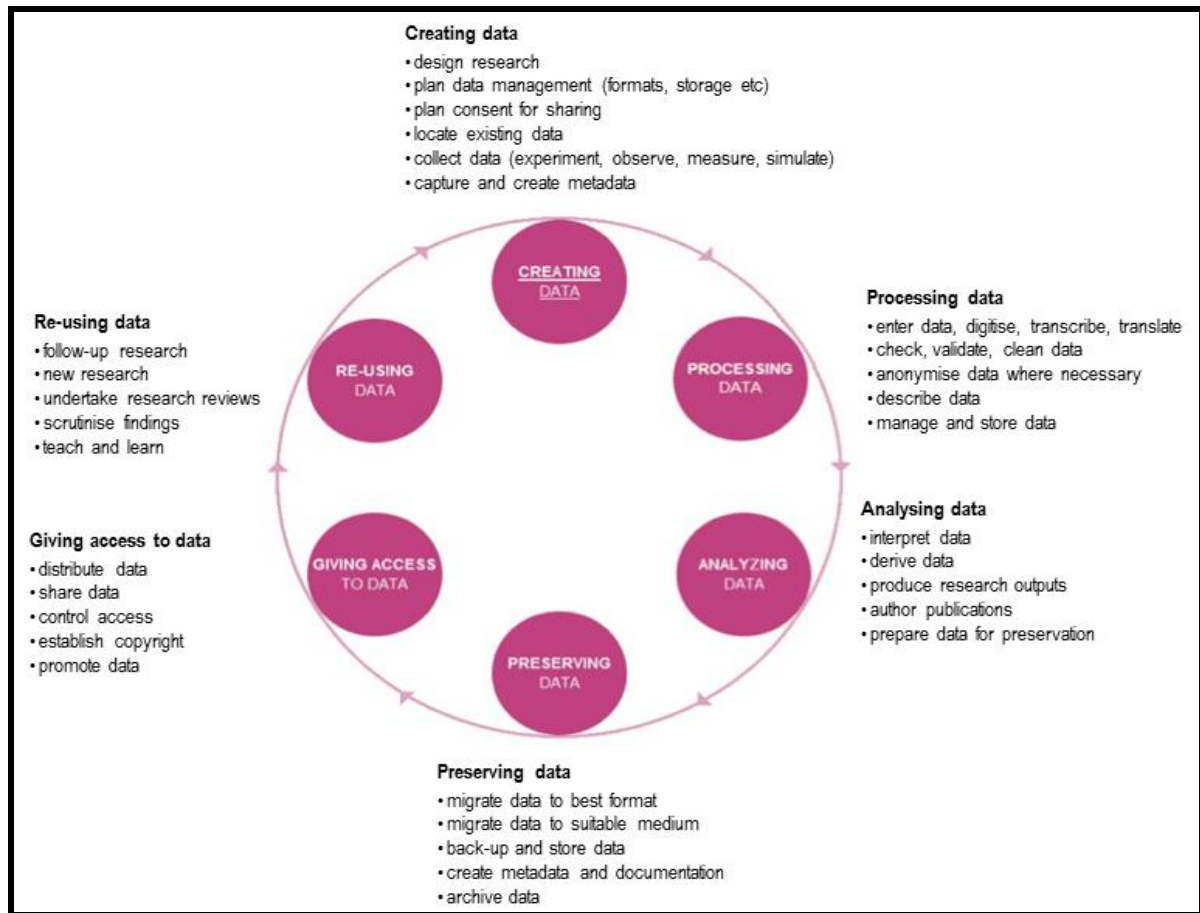
data management services for the London School of Hygiene and Tropical Medicine (Cole, 2016). Sykes (2009) reported that in the United Kingdom, a feasibility study aimed at addressing the issue of the entire data management lifecycle and providing a roadmap for RDM infrastructure development was conducted. In South Africa at the Cape Peninsula University of Technology library, RDM services are being developed on the premise that, within an RDM environment, several components exist including: infrastructure development; information flow and management; communication with researchers; development of tools related to the full research lifecycle and the means to store, curate and retrieve data for further use; and the training of researchers (Chiwane and Mathe, 2015).

Despite of its popularity, TELOS feasibility framework is not without its shortcomings. Bause et al. (2014) argue that in product development processes the term “feasibility study” is also used in a technical sense, whereby different organisations use it in different ways. Frequently, it is unclear, which are the pursued aims or which activities are necessary, resulting in uncertain methods.

#### **2.4.2 The UK Data Archive Research Data Lifecycle model**

Together with the TELOS model, the researcher used the UK Data Archive Research Data Lifecycle model as a guide to where research data management services can be applied. The UK Data Archive Research Data Lifecycle provides a data lifecycle model as an aid to researchers considering how data management relates to the lifecycle of a research project. A research data lifecycle describes the stages through which research data is collected, recorded, processed and results are published (UK Data Archive, 2015). A lifecycle approach is helpful for effective data management as it encourages planning ahead, including awareness of pre-existing data and how the data will complement this, thinking carefully about how data and metadata will be gathered and processed, and what type of consent is needed to deposit research data for reuse beyond the original project (Ray, 2014). Strasser (2014) explained that thinking about the data lifecycle can be a useful tool for beginning to sketch out the types of data services the institution might want to offer. Each stage of the lifecycle requires different services, and this model can be a helpful way to think about what the organisation can and

want to do. Figure 2.1 shows the UK Data Archive Research Data Lifecycle model showing the stages involved in RDM.



**Figure 2.1 UK Data Archive Research Data Lifecycle (UK Data Archive, 2015)**

There are quite a number of research data lifecycle models that exist to support RDM activities (Ball, 2012). The UK Data Archive Research Data Lifecycle model was adopted because it explicitly aims to provide a relatively detailed and comprehensive list of the processes and practices necessary for a functioning and mature data management infrastructure, and assembles this list through reference to data science, data curation and data management literature (UK Data Archive, 2015; Ball, 2012). This being so, the model is well suited to act as a checklist of data management practices that merit attention in a data management plan, and which therefore should be borne in mind when laying out the requirements for setting up RDM services.

### **2.4.3 Conceptual Framework**

The TELOS feasibility framework is a general model that can be applied in any discipline and by any organisation to assess the viability of starting new projects or interventions. To contextualise the framework, the researcher applied the TELOS framework to selected stages of the research data lifecycle model where the library can assume responsibility. The objectives of this study were informed by the constructs of the framework which will be defined herein after. Figure 1.1 highlighted in the previous Chapter shows how the TELOS framework was applied to the revised UK Data Archive Research Data Lifecycle model. The researcher picked four stages from the data lifecycle model which are; data collection, data storage, data preservation and data access.

The first stage from the conceptual framework of the research data lifecycle model is data collection. This involves gathering research data from researchers using various media. At this stage, research libraries should be in a position to offer advice on the appropriate structure, storage of research data (Stuart, 2014). Librarians need to devise strategies for collecting research data which in the traditional library roles can be equated to collection development. Witt et al. (2014) referred to this as collection development in the context of research data.

Data storage is the next stage in which libraries can provide RDM services. After data has been acquired from different researchers, there has to be mechanisms in place to store the data. The University of Queensland (2017) highlights that librarians must ensure that all research data, regardless of format, is stored securely and backed up or copied regularly. A data storage strategy is important because digital storage media are inherently unreliable and all file formats and physical storage media will ultimately become obsolete. This stage is crucial because the accessibility of any data is dependent on the quality of the storage media on which they are stored and the availability of the relevant data-reading equipment (UK Data Archive, 2015).

Data preservation is the next stage and it involves migrating data to suitable formats and media for preservation, creating backups, and creating any additional metadata that is necessary for preservation (Strasser, 2014). Stanford University Libraries (2016) pointed that

data preservation means more than just making a backup copy of the data; it means protecting data in a secure environment for long-term access and reuse. Hitchcock (2007) explained preservation as an activity in which specific items of data are maintained over time so that they can still be accessed and understood through changes in technology.

Data access is whereby an institution distributes data, share data, control access, establish copyright and promote data (UK Research Data Archive, 2015). Providing access to the data not only requires that the data is available, but that it can be found and appropriate rights are provided for its reuse.

## **2.5 Technological Requirements for Research Data Management**

Technological feasibility focuses on gaining an understanding of the present technical resources of the organisation and their applicability to the expected needs of the proposed project (Marakas and O'Brian, 2014). Pinfield et al. (2014) highlighted that RDM is underpinned by processes with technical implementations including data repositories, hardware and software allowing for storage and transport of different types and quantities of data. Steeleworthy (2014) posits that RDM is dependent on information technology and its ability to collect, store, preserve and provide access to research. Chronicling the development of RDM services at the University of Pretoria, Van Deventer and Pienaar (2015) highlighted that a task team comprised of the library and Information Technology (IT) staff was set to investigate infrastructure and technology needed for RDM across the university. Additionally, Van Wyk and Van der (2014) as one of their recommendations for the establishment of RDM services at the University of Cape Town state that it is necessary to investigate IT infrastructure that can accommodate small and big data sets as well as high performance computing.

Steeleworthy (2014) posits that technology plays a central role in RDM activities therefore, it is vital to assess the technological landscape before embarking on the RDM programme. This is supported by a study by Pinfield, Cox and Smith (2014) which revealed that RDM require significant levels of technology infrastructure. In the data lifecycle, technology is used to collect, process, store preserve and disseminate research data. The infrastructure required for

RDM covers a wide spectrum from the relatively small and unsophisticated to the large and complex systems (Henty and Kingsley, 2007).

### **2.5.1 Hardware and Software Requirements for RDM**

Computer hardware refers to the physical components of a computer that include input devices such as keyboards, storage devices such as hard drives and output devices such as speakers and monitors (Rajiv, 2016). Research data collection, storage, preservation and access is highly dependent on hardware and software used (Parsons (2013). The same author defines software is a set of programs, procedures, algorithms and its documentation concerned with the operation of a data processing system. In order to understand hardware and software requirements and to come to a meaningful conclusion about the technological requirements of offering RDM, it is important to highlight the research data formats, data storage devices and quantities of data produced. The choice of format for research data will determine how that data may be used, analysed, backed up, stored and potentially reused in the future thereby impacting on the choice of hardware and software requirements (University of Queensland, 2017; Eynden et al., 2010). Since digital research data may be endangered by the obsolescence of the hardware and software environment on which access to data depends; it is best practice to use standard and open data formats for long-term preservation, avoiding proprietary formats, which may be readable only using particular software.

#### **2.5.1.1 Research Data Types Generated by Researchers**

Van Tuyl and Michalek (2015) conducted a study at Carnegie Mellon University to assess the faculty RDM practices. Queried about the data formats that they produce in their various research projects, respondents indicated that they use many types of files for their data, with the five most common formats being: data tables (Microsoft Excel, comma delimited text), documents (Microsoft Word, PDF), code (python, R, MATLAB), text files, and image formats. Across colleges, uptake of various file formats varies slightly, though major format categories show some uniformity from college to college. The major data formats at Carnegie Mellon University were numerical data (78%), documents (71%) and images (50%) (Van Tuyl and Michalek, 2015). The findings are similar to Parsons (2013) who conducted a study at the University of Nottingham on the creation of RDM services. The study revealed that the most

common data types produced at the institution in order of popularity, included documents (text), spreadsheets, raw data (from software or specialist equipment), notebooks, databases and slides or specimens. Several studies have shown that top on the list of types of research data produced were documents (Van Tuyl and Michalek, 2015; Parsons, 2013).

### **2.5.1.2 Quantity of Research Data Generated by Researchers**

The quantity of research data produced during research endeavours varies from one project to the other and it has an impact on the choice of hardware and software. In order to have an in-depth understanding of technological requirements, it is worthy investigating the amount of data produced by researchers in their projects. Kennan and Markauskaite (2015) conducted a study which sought to provide an insight into the research data management and sharing practices of academics at ten universities in New South Wales, Australia. The study revealed that most participants (40.2%) indicated that they stored less than 10 Gigabytes (GB) of research data, and another large group (38.4%) answered that they have less than one Terabyte (TB) stored. Less than 20% of respondents had data sets larger than 1TB. Yet, the majority of them (13.2%) had less than 10TB and only 4.9% of respondents had more than 10TB. The findings are similar to those of Van Tuyl and Michalek (2015) who conducted a study at Carnegie Mellon University. The study revealed that at the project-level, researchers at Carnegie Mellon University produce varying amounts of research data ranging from just a few MBs to many TBs. The frequency distribution of data produced at Carnegie Mellon University range from 50MB to 50TB. These findings indicate that when an institution intends to start an RDM programme, there is need to invest in high storage servers and other storage devices that accommodate large files.

### **2.5.1.3 Research Data Storage Devices**

Additionally, in order to understand technological requirements, it is important to highlight the storage devices for research data. Researchers store their research data in different media. According to a study conducted by Parsons (2013) at the University of Nottingham, the common storage media in order of importance were university managed computers and laptops, networked university drives, external hard drives, Universal Serial Bus (USB) pen drives and web based services on paper. Similarly, Schumacher and Vandecreek (2014)

conducted a study for faculty members at five American universities to find out how they were managing their research data. Faculty members' responses revealed that most managed their data themselves, relying on some combination of individual devices and storage services. Again, participants were asked to name any and all digital object storage and preservation methods that they used. Thirty seven (37) (66%) relied on the hard drives of their office computer; 22 (39%) used an external hard drive; 21 (37.5%) used a hard drive as a built-in component of a personal computer; 18 (32%) used cloud-based services; 16 (28.5%) used a Flash/USB drive; ten (17.8%) used their email account(s); six (10.7%) used means or devices not mentioned in the project interview's list of storage options; and three (5.4%) relied on optical discs like CDs or DVDs.

Another study was conducted by Kennan and Markauskaite (2015). The study revealed that temporary and insecure local storage options featured highly, such as internal computer hard drives (80.8%), external hard drives (66.3%), USB sticks or flash drives (61.9%) and CDs or DVDs (28.4%). More secure external storage was indicated by far smaller number of participants and included such places as local area networks (35.7%), central IT or other university affiliated computing centres (21.4%) and departmental computing centres (10.1%). Of the 8.8% (n=48) of respondents who specified "Other" locations, nine respondents reported that their data was stored in or on the cloud during a project, another 20 respondents specifically mentioned Dropbox, 11 mentioned that their data was hardcopy and several specified "lab books", "note books" or "in locked filing cabinets." Two respondents mentioned that they used their home computer as back up.

Findings from these studies show that the most used data storage media are computer hard drives, external hard drives, cloud services. Whyte, Jones and Pryor (2014) suggest that cloud services may be considered as an option to reduce capital investment and avoid the need for expertise to establish services in house. The data storage devices all determine the backup mechanisms for research data. What this means is when considering technology requirements, the library should take into account the researchers' preferred storage media.



## **2.6 Economic Needs for RDM**

The economic feasibility evaluation purposes to determine whether the objectives of the proposed project are viable with benefit to the organisation under consideration of the costs involved (Bause et al., 2014). Furthermore, economic feasibility is an assessment to determine the extent to which the proposed project will provide positive economic benefits to the organisation (Marakas and O'Brian, 2014). This involves the identification and quantification of all benefits expected from RDM versus the costs incurred in setting up the system (cost/benefit analysis). In this study, economic feasibility looked at the costs involved in setting up RDM services and the associated benefits. The costs were aligned to selected stages of the research data lifecycle which are data collection, storage, preservation and access.

Since it might happen that developing a particular system may be technically possible but it may require huge investments and benefits may be less. In order to unearth the economic needs of an RDM project, the economic viability assessment of the proposed system should be carried out.

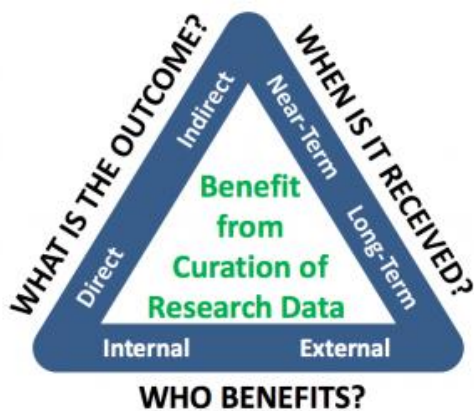
### **2.6.1 Cost/Benefit analysis for RDM**

Starting up research data management services will have significant costs, but will also enable significant research benefits (Hitchcock, 2013). Jones, Pryor and Whyte (2013) advise that when planning for RDM services, there is need to describe what are the predicted returns on investment, of which for RDM services they are likely to be expressed in terms of improved research impact, a more effective and cost-efficient research process, improved opportunities for new and more research, or increased funding. On the contrary, Buchhorn and McNamara (2010) argue that there is no administrative mechanism for bringing cost benefit analyses to a situation where costs of data management are currently being covered by grant and university funds, for the benefit of individual research groups and in some cases the general public.

During the onset of the RDM project, it is imperative to ascertain who will bear the costs of maintaining the service. Berman (2008) proposes that to ensure good data stewardship, institutions must be aware of data costs and include them in their overall IT budget. However, RDM funding is generally not (yet) seen as a part of the standard research process, nor is it part of the normal research budget, and the specifics of RDM and the budget scope for funding

data facilities are usually not clearly defined (Science Europe, 2016). This unclear position regarding funding for RDM activities presents challenges as to who should fund RDM activities within institutions of higher learning. There is need to ensure that all costs are factored in, including hardware, software, expert support, and time.

In order to ascertain the net benefits of research data management, it has to be done in a logical manner. There are models which were designed to explore the benefits and economic value of RDM. Worthy mentioning is the Keeping the Research Data Safe (KRDS) Benefits Framework (Beagrie et al., 2010; Beagrie, 2011). The KRDS Benefits Framework is a tool for identifying, assessing, and communicating the benefits from investing resources in the curation/long-term preservation of research data. The Framework organises benefits along three broad dimensions: the outcome achieved; when the outcome is achieved; and who benefits from the outcome (Beagrie and University of Victoria, 2012). Each of these dimensions can be subdivided into two categories: direct and indirect benefits, near-term and long-term benefits and internal and external benefits respectively. This is summarised graphically in Figure 2.2 below. Other approaches can be used to explore the economic value and benefits of RDM. For example, at the Economic and Social Data Service Centre of University of Victoria, approaches used include investment and use value, contingent valuation using stated preference techniques, welfare approaches to estimating consumer surplus, and a macro-economic modelling approach that seeks to explore the returns to investments in data creation and hosting (Beagrie, 2011).



**Figure 2.2 The KRDS Benefits Framework (Beagrie, 2011)**

### **2.6.2 Costs for setting up RDM infrastructure**

Hole et al. (2010) assert that predicting the costs of long-term digital collection, storage, preservation of and access to research data is a crucial yet complex task for even the largest repositories and institutions. The costs involved include costs for staff recruitment and training, consumables, equipment and maintenance. Several models have been developed to ascertain costs involved from setting up RDM services to maintenance of the repositories (Davies et al., 2007; JISC, 2008; Addis and Arkivum, 2016). The cost variables for RDM can be centred on the activities of the research data lifecycle. Several scholars have recommended that costs should be calculated based on the stages and activities involved in the research data lifecycle (UK Data Service, 2015; Jones, Pryor and Whyte, 2013; Eynden et al., 2011). Addis and Arkvum (2016) concur to the approach by saying that using a lifecycle model helps to ensure that nothing gets left out, but it also allows the boundaries of the cost model to be well defined. The Life Cycle Information for E-Literature (LIFE) model from the British Library is a good example of an empirical model for providing costing estimates for the lifecycle of digital collections and, consequently, allowing for the exploration of the practical and strategic dimensions (Hole et al., 2010). The LIFE model defines each of the stages of the lifecycle: acquisition ingests metadata, access, storage and preservation and then works out the cost of each one over time.

Davidson et al. (2014) assert that preserving research data for the long term has a cost; although the infrastructure itself is costly, more significant is the cost associated with human resources, such as personnel to manage and maintain the archive. Storage costs for digital data are decreasing, but costs related to storage, such as power, data curation and annotation, and personnel, are not decreasing (Berman 2008 as cited by Strasser, 2014). Increasing amounts of digital data, and the need to comply with regulations regarding backup and monitoring, emphasise that these costs should not be underestimated or overlooked. Strasser (2014) posits that short-term costs for data preservation are primarily those related to storing data rather than archiving it. This may include software or hardware for backing up data or personnel costs for managing and organising data for storage. Longer-term preservation costs are associated with archives. Many repositories and archives use annual pricing schemes for a set

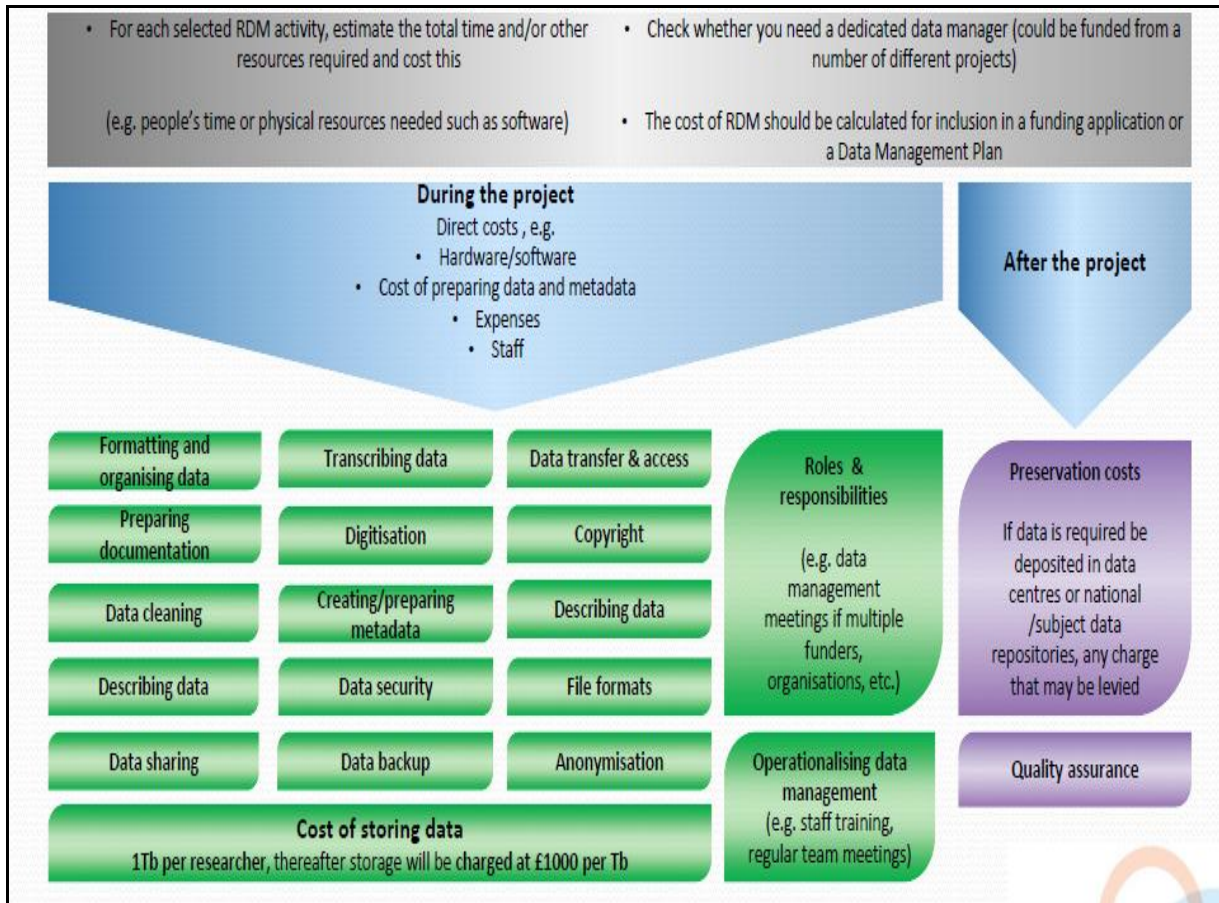
amount of data; this situation is changing, however, to better meet the needs of researchers whose costs are intertwined with grant cycles.

In their analysis of the institutional readiness and responsibility for managing research data Henty and Kingsley (2015) point out that there are considerable financial and economic implications in having increased quantities of research data being generated in electronic form and the consequent need to ensure that this is properly managed and preserved. The study further revealed that many of the university librarians interviewed expressed their concern that repositories are not embedded in the university's funding model and that there is no guarantee of future funding to sustain what is becoming a core service.

The UK Data Archive (2015) provided a data costing toolkit to help formulate RDM costs in advance of research starting , for example for inclusion in a data management plan or in preparation for funding application. Additionally, Eynden et al. (2011) propose that to cost research data management in advance of research starting, for example for inclusion in a data management plan or in preparation for a funding application, two approaches can be taken. Either all data-related activities and resources for the entire data cycle that is from data creation, through processing, analyses and storage to sharing and preservation can be priced, to calculate the total cost of data generation, data sharing and preservation. Alternatively, one can cost the additional expenses above standard research procedures and practices that are needed to make research data shareable beyond the primary research team. This can be calculated by first listing all data management activities and steps required to make data shareable, then pricing each activity in terms of people's time or physical resources needed such as hardware or software.

In 2008, Beagrie, Chruszcz and Lavoie developed a cost model for research data management under the JISC funded Keeping Research Data Safe (KRDS). The model provides for service and resource costing around the stages of the data lifecycle. The services costs are those that are involved in RDM activities such as data collection, preservation and access provision. The UK Data Service (n.d) added that costs of RDM range from data creation, through processing, analyses and storage, to sharing and long-term preservation- the activities can be priced to

calculate the total cost of all data generation, data sharing, data access and preservation activities. Resource costs include infrastructure development and cost of technology acquisition and development. Similarly, the UK Data Service provided that another approach of determining RDM costs is through identifying the resources that would be needed to preserve and make research data shareable beyond the primary research team. These resources may include: people, equipment, infrastructure and tools to manage, document, organise, store and provide access to data. The same approach was applied during the development of the Dryad data repository (Beagrie, Eakin-Richards and Vision, 2010). Wilhem, Olster and Shoulson (2014) identified four major costs that are involved in RDM which are infrastructure, administration, standardisation, human resources, and opportunity costs. Buchhorn and McNamara (2010) assert that the costs of a highly centralised data management can be estimated as some fraction of the funds being expended on labour and equipment. The aforementioned studies indicate that starting an RDM programme involves a considerable amount of financial resources which must be taken into consideration for the success of the proposed project. Figure 2.3 summarises the costing for RDM as provided by St Andrews University (2014):



**Figure 2.3 Costing Research Data Management (St Andrews University, 2014)**

## 2.7 Legal Obligations for RDM

The legal feasibility of a proposed project includes a thorough analysis of any potential legal ramifications resulting from the construction and implementation of the new project (Marakas and O'Brian, 2014). It determines whether the proposed system conflicts with legal requirements, for example a data processing system must comply with the local data protection regulations and if the proposed venture is acceptable in accordance to the laws of the land. Greenbaum and Gerstein (2003) acknowledged that it has become increasingly apparent that to achieve seamless access to data it is necessary not only to adopt appropriate technical standards, practices and architecture, but also to develop legal frameworks that facilitate access to and use of research data, whether on an inter-organisational basis or across national borders.

Understanding the legal obligations surrounding research data management is crucial as it guides the preservation and access of research data. Fitzgerald and Pappalardo (2007) assert that it is apparent that to achieve seamless access to data it is necessary not only to adopt appropriate technical standards, practices and architecture, but also to develop legal frameworks that facilitate access to and use of research data, whether on an inter-organisational basis or across national borders. The UK Data Archive (2015) emphasises that before embarking on a RDM project, it is imperative to know your legal, ethical and other obligations regarding research data, towards research participants, colleagues, research funders and institutions. RDM legal obligations broadly include intellectual property rights which encompass trademarks, design rights, patents and copyright. Trade secrets protect confidential business information. Copyright provides exclusive control over the copying, distribution, performance and display of a piece of work (Caroll, 2015).

Legal requirements for RDM services are also concerned about data governance. Data governance is the system of rights, rules, and responsibilities that specify who can do what with data. Many of the skills needed to deal with data governance are those librarians already possess: experience dealing with licensing terms and agreements, knowledge about copyright, and the ability to read contracts carefully to bring out ownership issues (Strasser, 2014). The author argues that much of the complexity of data governance is related to the array of stakeholders in the data being produced. Various stakeholders have different needs from and investments in the data and are not always clear about their rights and responsibilities in relation to that data.

A detailed understanding of legal obligations regarding research data is also centred on access and use of research data. While several authors have provided benefits of providing open access to data, there has not been agreement on whether research data should be made available without any restrictions (Strasser, 2015; Schopfel et al., 2016; Wessels et al., 2014; Wilson, 2010). The Research Council of UK (2015) argues that there are sometimes legitimate reasons to restrict access to data. Research organisations must ensure that the publication of data conforms to relevant legislation. Institutions should ensure they are aware of relevant legislation in particular regarding data protection, freedom of information and environmental

information regulations and disseminate this to their researchers as appropriate. Additionally, given this kind of complexity around data collection, processing and interpretation, the Royal Society (2012) as cited by Wessels et al. (2014) suggests that open access to data must ensure that the provenance and clarity of data and metadata are clearly understood by stakeholders and by those accessing the data.

### **2.7.1 Ownership of Research Data**

It is crucial that researchers ensure ownership of primary materials and research data is identified and documented at the start of a research project and reviewed and updated whenever appropriate (Strasser, 2015; Charles Darwin University, 2010; Fitzgerald and Pappalardo, 2007). Research Data Management Plans should detail how ownership and storage of data and materials will be affected by researchers changing institutions, or withdrawing from collaborative projects. Fitzgerald and Pappalardo (2007) suggest that reference to the “owner” of data will usually be a reference to the person who owns legal rights in relation to the data, the person who has physical possession of the data, the person who controls access to and use of the data or a combination of any of the above. Corall (2015) asserts that the question of who should own the data raises so many questions given the number of players involved in data generation. To deal with such complexities of data ownership, strategies such as developing contracts and licences, clarifying contracts of use should be developed (Corrall, 2015; Fitzgerald and Pappalardo, 2007)

### **2.7.2 Ethical Considerations for RDM**

Privacy, confidentiality and consent need to be considered whenever research data include, or are derived from, personal data. This is important not only to satisfy legal requirements but also, crucially, to maintain and build public trust in the use of personal data for research (Research Council of UK, 2015). Data which are personal, sensitive or confidential may be shared provided attention is paid from the outset to the legal and regulatory requirements and any professional standards are met.



The Charles Darwin University (2015) suggests that it is the responsibility of all researchers and research administrators to ascertain and comply with any ethics or contractual confidentiality conditions relating to the primary materials and research data and related records, such as:

1. Respecting any confidentiality agreement about stored data that has been made with participants and ensure documentation of same for the awareness of future users;
2. Considering appropriate ways of collecting, storing and accessing data, and communicating with research participants about these issues;
3. Establishing consent processes that include information about the form in which the data will be stored, indexed and the purposes for which the data will be used and/or disclosed.

### **2.7.3 Institutional Policies for RDM**

The role of RDM policies cannot be overemphasised in the success of RDM programmes (Higman and Pinfield, 2015; Nugroho et al., 2015). RDM policies respond to a number of drivers, including data collection, storage and preservation, but also data access and sharing (Pinfield et al., 2014). Policies are mostly used to provide “credentials” for those championing RDM, gain access to funding for IT infrastructure (Pryor, Jones and Whyte, 2014), clarify institutional positions and outline roles and responsibilities (Brown and White, 2014). Patel (2016) recommends that the institutional RDM policy should clearly spell out the following in the context of data sharing: 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, terms and conditions of the use of data, protection of confidentiality of sensitive data, protection of data against security breaches, intellectual property (IP) concerns.

However, the policies are not without difficulties. There are, for example, concerns that they will become unachievable statements of aspiration as opposed to intent (Jones, 2014). Even when policies are written realistically there remains a significant risk that they will be unfulfilled without both considerable cultural change (Brown and White, 2014; Pryor, 2014) and acceleration of the implementation of promised services.

Buchhorn and McNamara (2010) in their study to determine sustainability of Australian research data found that research groups and organisations rarely have formal policies for the management of data. Similarly, Tenopir et al. (2016) in their study on RDM activities among European academic libraries revealed that less than half (40.9%) of libraries say they currently have policies relating to RDM. In the same vein, a study by Cox and Pinfield (2014) revealed that an affirmative response was made by 25 (31%) of institutions to the question: ‘Does your institution have a formal research data management (RDM) policy in place?’ This compares with 17% reported by Corral et al. (2013) from earlier in 2012. An additional 35 (43%) in Cox and Pinfield’s survey stated that they expected to have a formal policy in place within the next year. Similarly, the University of Cape Town is currently drafting a Research Data Management Policy (UCT, 2017).

## **2.8 Organisational Capabilities for RDM**

The organisational feasibility assessment focuses on the degree to which the proposed project fits in with the existing environment and objectives with regard to development schedule, existing skills and corporate culture (Marakas and O’Brian, 2014). In this study, organisational feasibility covered librarian’s skills and knowledge of RDM, researchers’ attitudes towards RDM, institutional responsibility for RDM, management support and stakeholders for RDM. Henderson and Knott (2014) observed that the introduction and success of RDM services in academic libraries calls for the need to hire new staff or re-skilling and up skilling of librarians to take up new roles and responsibilities. As the creators and users of research data, researcher engagement is crucial in the development of RDM services (Jones, Pryor and Whyte, 2013; Wilson and Jeffrey 2013; Buchhorn and McNamara 2010).

### **2.8.1 Skills and Knowledge Requirements for RDM**

The success of RDM programmes partly depends on the skills and knowledge of the people involved. Henderson and Knott (2014) observe that the introduction and success of RDM services in academic libraries calls for the need to hire new staff or re-skilling and up skilling of librarians to take up new roles and responsibilities. Many academic libraries are hiring data librarians to cater for the emerging role of RDM. For example, The Virginia Commonwealth University libraries, Witwatersrand, University, University of Pretoria, have data services

librarians (Henderson and Knott, 2015; University of Pretoria, 2015; Van Wyk et al, 2014). Cox et al. (2015) in their study on upskilling liaison librarians for RDM suggested that there is need to align RDM skills to existing liaison librarians. Table 2.1 below highlights the librarians' roles and required competences in RDM as provided by Cox et al. (2015)

**Table 2.1: Librarians' roles in RDM and required competencies mapped to existing roles (Cox et al., 2015)**

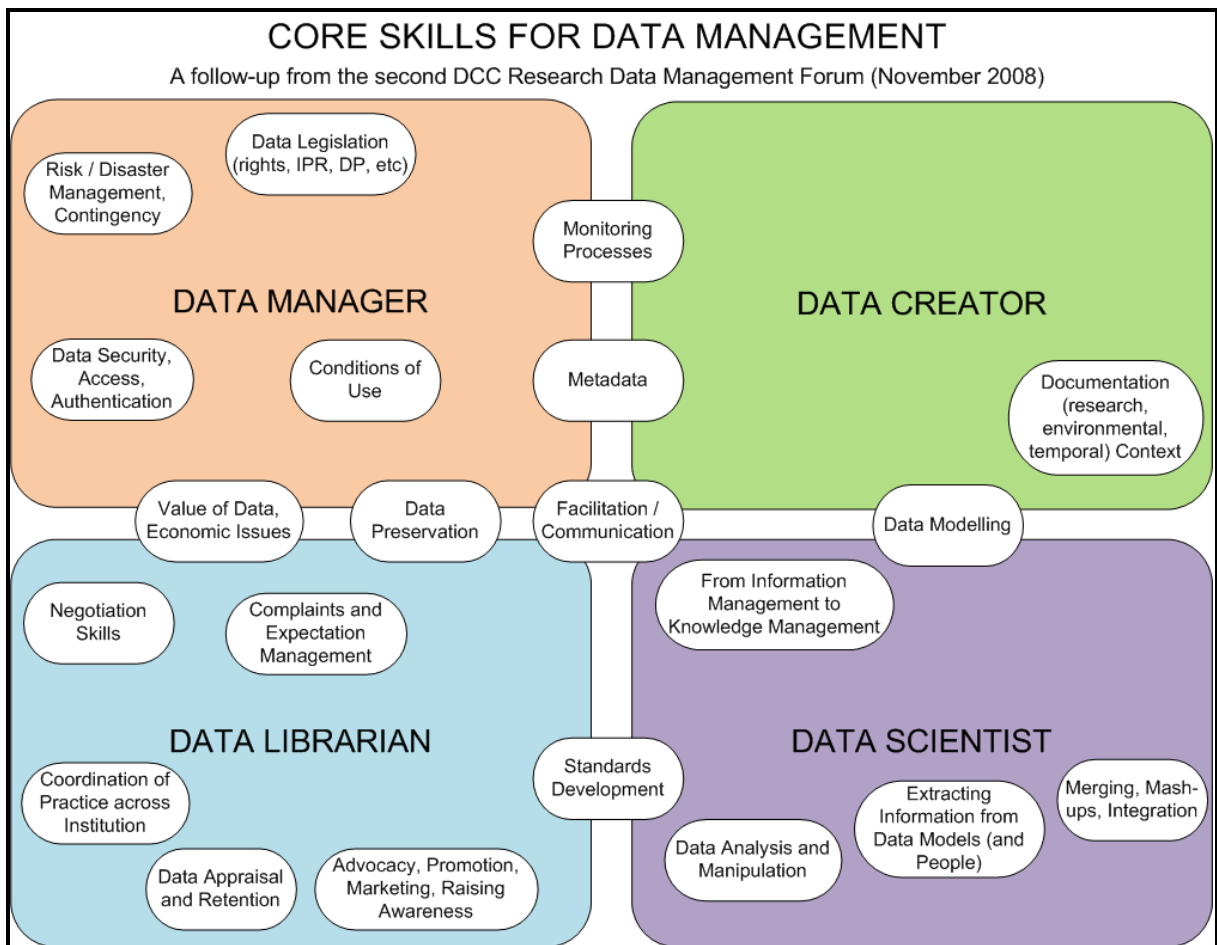
<b>Role</b>	<b>Alignment with existing roles</b>	<b>Competencies required</b>
<i>Policy and advocacy</i>		
Lead on institutional data policy	Advocacy role for example in the area of open access	Strategic understanding and influencing skills
<i>Support and training</i>		
Bring data into undergraduate research-based learning, promoting data information literacy	Information literacy training	Understanding of RDM best practices as they apply to relevant disciplines; pedagogic skills
Teach data literacy to postgraduate students		
Develop researcher data awareness		
Provide an advice service to researchers (and research administrators)  Eg on writing Data Management plans or advice on RDM within a project. Advice on licensing data. Advice on data citation. Perhaps measurement of impact of data sharing.	Reference and enquiry roles; producing print and Web-based guides; copyright advice.	Reference interview, knowledge of RDM principles
Provide advice as above through a Web portal	Library Web site	Knowledge of institutional and extra-institutional resources
Signpost who in the institution should be consulted in relation to a particular question	Role of library as point of enquiry and the reference interview	Knowledge of institution

Promote data reuse by making known what is available internally and externally; explaining data citation	Marketing of library resources	Knowledge of researchers' needs, knowledge of available material
<i>Auditing and repository management</i>		
Audit to identify data sets for archiving, create a catalogue of materials or to identify RDM needs	Metadata skills	
Develop and manage access to data collections	Collection development, digital library management and metadata management	Audit interviews, knowledge of RDM principles, metadata, licensing
Develop local data curation capacity	Open access role. Preservation role.	Knowledge of RDM principles, relevant technologies and processes, metadata

Tenopir et al. (2015) conducted a study on data intensive roles for the future in selected Association of College and Research Libraries (ACRL) library directors across the United States of America and Canada. Respondents were asked how their libraries have developed staff capacity for RDM. A little over half of the 25% of those offering RDM (54.5% or 12 of 22) indicated that they have reassigned existing staff. Only a few (6) have hired new staff. However, the 2011 baseline survey of academic librarians by Tenopir and colleagues showed that many academic librarians do not feel prepared to take on these new roles in spite of plans by their libraries to offer RDS. Similarly, Cox and Pinfield (2014) conducted a study which revealed that libraries had patches of the relevant skills but that they were not seen as widely enough spread. The study also showed that there is a wide range of skills required for research data management, and where the library is very strong in is: cataloguing including metadata, digital preservation, curation, training, academic engagement, copyright, publication process. Over 50% said the library staff did not have the right skills, but these replies were qualified too, acknowledging that they had some of the skills needed.

Similarly, Cox and Pinfield (2014) found that there was an RDM skills deficit in most academic institutions. The specific skills needed were data curation skills (mentioned by

nearly 90%), technical IT skills and knowledge of research methods. About 40% also recognised the need for disciplinary knowledge. A study by Woolfrey (2010) on skills development to assist data usage for policymaking in Africa showed that there are only a handful of RDM practitioners in Africa. An earlier study by Parsons (2013) revealed that there was a wide interest in three key areas at Nottingham University namely data management plans, storing data and cataloguing and metadata. Some additional suggested roles for libraries are to develop researchers' data-awareness, to adopt a data archiving and preservation role, and to train data librarians (Tenopir et al., 2014). Figure 2.4 below shows core skills for data management as provided by the Research Data Management Forum.



**Figure 2.4 Core skills for data management (Research Data Management Forum, 2008)**

In order to embrace RDM in academic libraries, several scholars have suggested re-skilling, up skilling liaison librarians (Auckland, 2012; Cox, Verbaan and Sen, 2013). In light of the

growing interest in RDM in South Africa and the need for capacity building in this area, the LIASA's Higher Education Libraries Interest Group hosted an RDM workshop, facilitated by the DCC, UK, in Cape Town in March 2014 (Kahn et al., 2014).

### **2.8.2 Researcher's Attitudes Towards RDM**

As the creators and users of research data, researcher engagement is crucial in the development of RDM services (Jones, Pryor and Whyte, 2013; Wilson and Jeffrey 2013; Buchhorn and McNamara 2010). Any service provision needs to be based upon a close understanding of research, its patterns and timetables, motivations and priorities. This cannot be achieved without a commitment by the research community to contribute to the definition of service requirements. Without their active involvement and support the success of an RDM service is bound to be limited. Furthermore, Sompel et al. (2004) aptly reminded "Like any technology, success will depend not only on technical soundness but on the willingness of the participants in the system that is publishers, scholars, academic institutions, funding institutions, and others to adopt new tools and develop new organisational models on top of them."

Research has shown that there are mixed views from researchers regarding their willingness to submit research data to data repositories (Wilson and Jeffrey 2013; Buchhorn and McNamara 2010). Keil (2014) in a study on research data needs from academic libraries taking a perspective of the faculty researchers revealed that sharing data can be particularly unnerving to scientists who may perceive a loss of a competitive edge for their next follow-up manuscript or grant proposal. Similarly, Kennan and Markauskaite (2015) conducted a study on research data management and sharing practices of academics at ten universities in New South Wales, Australia. Researchers were asked whether, once they had collected their own data, they would be willing to share them outside of their research team or project. While more than half (54.7%) indicated that they would not be prepared to share any of their data, 36.4% indicated they would be prepared to share some of their data and 8.9% indicated that they would be prepared to share most of their data. On the same note, a study by Buchhorn and McNamara (2010) reveals that researchers are often reluctant to share their data. Data is viewed as a personal good, created by researchers and to be exploited by them. While many

researchers feel data collections should be available to the research community there is a very strong and unanimous view that researchers should be able to exclusively exploit ‘their’ data for a period of time before it becomes available to others. On the contrary, a study by Kennan and Markauskaite (2015) in South Africa showed that the majority 73% of the researchers were willing to submit their research data to data repositories.

### **2.8.3 Institutional Responsibility for RDM**

To further explore organisational requirements for offering RDM services, it might be interesting to take note of institutional responsibility for RDM. Apparently, authors do not agree who specifically should hold the responsibility of engaging with researchers in order to develop further knowledge of the research process and research data (Tenopir et al., 2015; Erway, 2013; Jones, Pryor and Whyte, 2013). Gold (2007) argues that it is fair to say there is still a substantial amount of uncertainty about the roles libraries can play in scientific data management, reflecting an environment of ongoing experimentation and negotiation. A considerable amount of literature has been published acknowledging that librarians are well positioned to play an important role in RDM (Kennan and Markauskaite, 2015; Tenopir et al., 2015; Erway, 2013; Jones, Pryor and Whyte, 2013; Lyon, 2012; Corral, 2012; Alvaro et al., 2011, Eynden et al., 2011; Buchhorn and McNamara, 2010; Gabridge, 2009 and Henty, 2008). Lewis (2010) provides areas in which the library can take part. Similarly, MacColl (2010) advises that libraries can take on a more comprehensive and strategic role: libraries should be involved throughout the research process and need to be actively engaged in curating, advising, and preserving research outputs.

On the contrary, Potter, Cook, and Kyrillidou (2011) argue that libraries and librarians have a long way to go before realising RDM roles. The authors found that only 9 of 86 (or 10%) narrative profiles created by ARL members in the US and Canada included references to e-science/data curation and management as an important service supporting faculty success and scholarly communications. As part of a 2007 investigation into all of the types of support provided to researchers by 134 US and Canadian academic health science libraries, Cheek and Bradigan (2010) found that just 12.2% of these libraries provided support for “data curation”. About half of the respondents to ARL's 2009 North American e-science survey (Association

of Research Libraries, 2010) had on-campus support units for scientific research data; however, the Data Working Group at Cornell University Library discovered that few university libraries were actually involved in research data curation (Steinhart et al., 2008). In the same vein, Corral (2012) categorised IT experts as conduit specialists, library or information scientists as content specialists, and academics or professionals as context specialists. Applying that analysis, the activities of data scientists would traverse all three fields, although the demands for handling data in each of the fields is constantly evolving, producing a continual pressure upon everyone to up-skill. It is therefore, observed that there is need to clarify the responsibilities for RDM services within an academic institution.

#### **2.8.4 Management Support for RDM**

Any intervention without the support of management is bound to fail. Parsons (2013) acknowledges that there is little doubt that RDM is a high profile service that is driven by those who fund and support research and top management. Tenopir et al. (2015) assert that the development of an RDM programme requires both a top-down and bottom-up approach. Leadership and engagement from the librarians and staff who will provide and promote the services is essential; however, leadership and resources are also required from the top-level administrators. Campus-wide support is important for securing funding for RDM since libraries do not always have the resources to implement such a programme on their own (Tenopir et al). Jones, Pryor and Whyte (2013) highlight that the principal role of university management in the introduction of RDM services will be to ensure proposed services are desirable, achievable and sustainable, and if so, to give clear, informed and unequivocal support. Equally challenging for university managers is the need to treat RDM services as a serious investment in infrastructure within the long-term institutional business planning process. Pinfield, Cox and Smith (2014) argue that a lack of institutional support may be one of the reasons there has not been a faster library adoption rate of RDM.

Jones, Pryor and Whyte (2013) provided the principal responsibilities to be met by senior management in RDM and they are to:

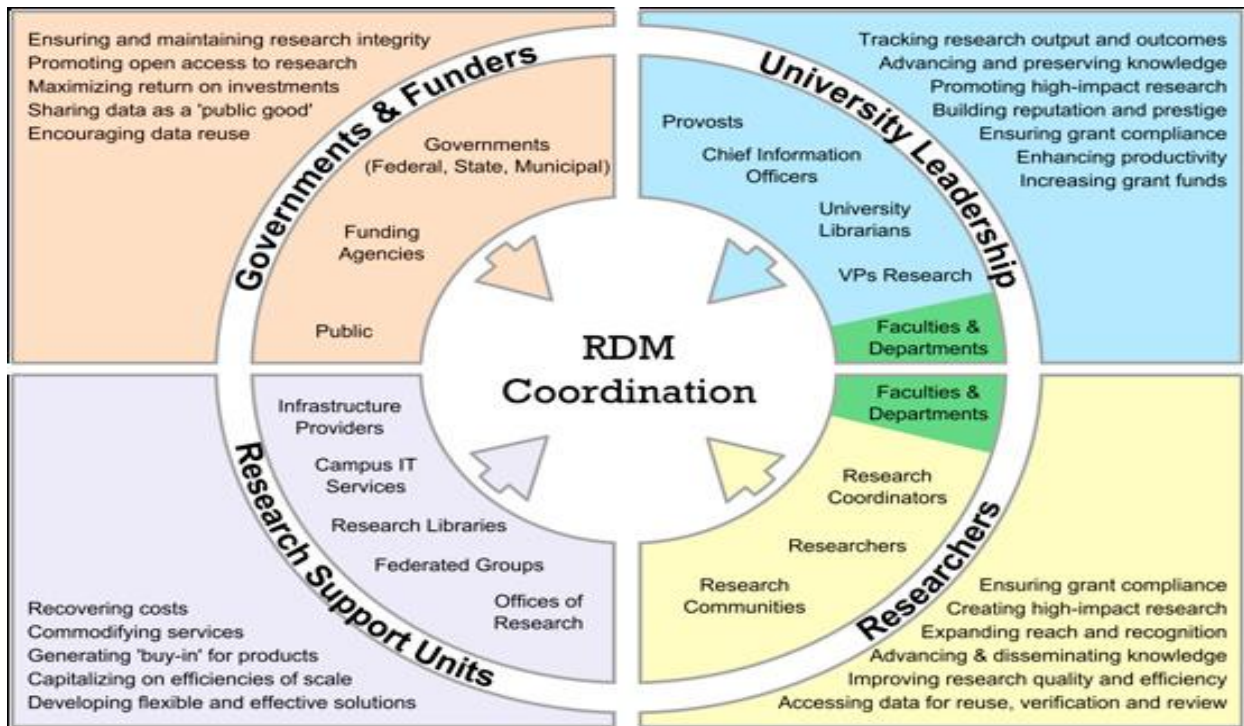
1. provide a champion, at Pro Vice-Chancellor (PVC) Research level or equivalent, to act as influential advocate and to chair steering/working group business;



2. establish a representative, balanced and appropriately equipped steering/working group that will reflect the interests of essential stakeholders;
3. consider, comment on and eventually approve proposals, plans and strategies, including the endorsement of budgets and organisational restructuring;
4. advise on the higher-level strategic issues that must be addressed during service design;
5. ratify a policy that articulates the core RDM principles and acts as a framework for guidelines and service design

### **2.8.5 Stakeholders in RDM**

Although the spectrum of RDM stakeholders has been variously categorised in the literature (Erway 2013, Jones et al. 2013, Pinfield et al. 2014; Wilson, 2010), Flores et al. (2013) assembled them into four main categories as highlighted in figure 2.3 below. RDM partners at the institutional level commonly include the library, information technology services, and the office of research, records and archives services, the institutional quality management unit, the institutional ethics committee chair, research chairs, heads of research units and centres, and the centre for postgraduate studies (Chiwere and Mathe, 2015; Jones et al. 2013). The structure is not intended to denote or prescribe segregation between groups, but rather to align them according to similar interests, roles, and responsibilities in RDM. Figure 2. below highlights some of the stakeholders in RDM.



**Figure 2.5 Stakeholders in RDM development, categorised into four general stakeholder groups. Individual stakeholder units are identified in the central ring, with general group interests listed in the adjacent boxes (Flores et al., 2013)**

From the previous studies, it can be observed that key elements of a successful RDM implementation are institutional collaboration and inter-institutional partnership with different stakeholders.

## 2.9 Scheduling RDM Services

Schedule feasibility is concerned about how and when the proposed project can be completed within its scheduled time limits, by a planned due date (Marakas and O'Brian, 2014). This means estimating how long the system will take to develop, and if it can be completed in a given time period. In this study, schedule feasibility looked at how RDM services can be implemented. Rice et al. (2013) emphasised that each new system or process developed to improve the management of research data needs to be integrated into the existing architecture of university systems, such that it maximises the benefits it can deliver whilst reducing the cost and time required to use.

There are various avenues that can be taken by an institution planning to implement RDM services. Factors such as the availability resources, skills and technology have a bearing on

whether an institution should take a campus wide approach or departmental approach to RDM. RDM services must go through the same planning as any new service as described by Shen and Varvel (2013), but luckily many of these plans have been shared in the literature for libraries to adapt to their own institution. One of the main themes that arose from these plans is the need to start with a small number of RDM services, and then look to expand those services (Toups and Hughes, 2013; Wright et al., 2014). However, starting with a small number of services may seem problematic if there is a high need at the institution, but, just as with cataloguing gifts and donations, one must perform triage. The plan outlined by Henderson and Knott (2015) explain that, the Virginia Commonwealth University triaged data services on a needs basis. They also caution that curation, such as appraisal and weeding, should perhaps not be the first plan of action. Another model of implementing RDM services is presented by Knight (2015) from the London School of Hygiene and Tropical Medicine (LSHTM) who describes the school's three-tiered triage approach: priority is given to researchers performing funded research, questions are directed to existing expertise whenever possible, and each question submitted is recorded to help plan educational support and to reduce the time needed to answer frequently-asked questions. Similar question tracking systems are used by reference services to expedite response time.

Piloting RDM services is another technique for starting small and identifying problems with low risk to the entire project. The pilot described by Knight (2015) targeted only research projects that had secured funding and did not require a data management plan in the funding process. Although the criterion meant running the risk of researchers not planning for the funding required for RDM, Knight explains that establishing research data management plans after acquiring funding allowed researchers more time to discuss the plan with their team. Since many universities are multi-disciplinary and deal with different forms of data, the pilot can focus on reflecting the diversity. The pilot described by Johnston (2014) only used five data sets, but from different disciplines, at the University of Minnesota. Piloting with interested users also ensures that the researchers are committed to the project and the pilot will produce useful results (Mclure et al., 2014; Newton et al., 2011). CPUT libraries in South Africa decided to run a pilot project with one or two research groups before rolling out the full service infrastructure to the university (Chiwere and Mathe, 2015). The advantage of piloting

is that RDM developers can utilise the knowledge of department liaisons to help find potential subjects or leverage existing relationships within the departments (Newton et al., 2011).

Another approach of implementing RDM services suggested by Creamer (2015) is to start with data from student theses and dissertations. Where students are required to submit electronic theses and dissertations, libraries can ask students to submit their data, or the institution may make submission mandatory (Doty et al., 2015). The student subjects can be used to pilot data management plan services, and their research data can be used to pilot data-specific repositories. Students make excellent pilot subjects; they are more willing to accept new practices, since they are not yet comfortable with a standard practice, and the data management skills that they learn will be brought forward into the rest of their life's work (Choudhury, 2012; Creamer, 2015). Although piloting with student data sets has many benefits, Steve Van Tuyl from the Oregon State University reminds librarians that student data sets may also need the most work to curate (Creamer, 2015). Doty et al. (2015) also note that at their university, where data submission is not mandatory, the number of electronic theses and dissertations being submitted with supplemental data files have doubled in 2014. If academic libraries are hesitant to pilot data repositories with researcher or student data, they could use the library's own data, such as electronic resource data (Ogier et al., 2014). As an added benefit, a data audit can be performed at the same time as the pilot. Another method of implementing RDM services strategically is by using pre-existing skill and infrastructure. Section 2.4.5 above highlighted how academic institutions can make use of existing repositories for research data management services.

## **2.10 Summary**

This chapter reviewed literature on RDM in academic libraries. Literature showed that libraries are not on the same level in the development of RDM services. From the review of literature, it was noted that most studies focused on already established RDM services hence they focused on assessing RDM maturity levels. Very few studies provided a holistic approach for setting up a RDM services. It was noted that there are very few studies which were done on RDM from an African perspective and a few that exist were done mostly in South Africa.

The gaps established the validity of this study and a point of departure as they offer opportunities for new research. The review was guided by the TELOS feasibility framework and the UK Data Research Data Lifecycle model to establish the feasibility of offering RDM services in academic libraries. From the literature review, it was found that in terms of technological requirements for setting up an RDM programme, there is need to have the necessary software and hardware that can accommodate different types and quantities of research data. It was also revealed that in terms of economic requirements, there is need to have financial resources to cater for staff training on RDM, advocacy, acquisition of technological resources. In terms of legal obligations, previous literature showed that there is need to have institutional RDM policies that cover all activities throughout the research data lifecycle. Previous research has shown that in terms of organisational capabilities, there is need to assess RDM skills and competencies of librarians, institutional responsibility for RDM, researcher's attitudes and management support. Literature has shown that librarians in most institutions do not possess the requisite RDM skills to take up the new roles. Regarding scheduling RDM services, previous literature showed that there are different approaches that can be adopted by organisations planning to offer RDM. These include piloting RDM services and starting with data from student theses and dissertations. The decision on which approach to take is determined by factors such as the availability of resources, skills and technology.

## CHAPTER 3

### RESEARCH METHODOLOGY

#### 3.1 Introduction

This chapter covers the research design and methodology that was used to address the research problem. Research methodology is a systematic way to solve a research problem (Kothari, 2004). Research methodology is the overall approach to a problem which could be put into practice in a research process, from the theoretical underpinning to the collection and analysis of data (Remenyi et al., 2003). On a similar note, Collis and Hussey (2009) identified methodology as the “overall approach to the entire process of the research study”. Essentially, it involves the procedures by which researchers go about their work of describing, explaining and predicting phenomena. The methodology brings out the strategies which were employed by the researcher to answer the research questions which focused on uncovering whether the UZ library was prepared to offer research data management (RDM) services by looking at the technological, economic, legal, organisational and scheduling aspects. This chapter will provide the research philosophy, research approach, research strategy, data collection methods, sampling and sampling techniques, ethical considerations and data analysis, presentation and interpretation procedures. The methodology chosen was seen befitting to answer the research questions of the study which required the use of both qualitative and quantitative techniques.

#### 3.2 Research Philosophy

The research philosophy or worldview is defined by distinct elements including epistemology (how we know what we know), ontology (nature of reality), axiology (values) and methodology (the process of research) (Remenyi et al., 2003). The pragmatic worldview was adopted for this study. Given the research problem at hand where the researcher aimed at exploring the feasibility of offering research data management services at the UZ, the pragmatic worldview was suitable because it uses all approaches available to understand the problem. The objectives of the study which were to find out the technological, economic, legal, organisational and schedule feasibility of offering RDM; require pluralistic approaches

to fulfill them and this implies more than one system of philosophy and reality. The pragmatic approach is one which takes a practical orientation to a problem and finds a solution that is fit for a particular context (Hammond and Wellington, 2013). The pragmatic worldview is based on the premise that knowledge is both socially constructed and based upon the reality of the world we experience and live in (Gray, 2014). Thus, pragmatism generates solutions which are 'fit for purpose' and these solutions will be generated in 'pragmatic ways'. Its mode of inquiry then makes use of induction (to identify patterns), deduction (testing theories) and abductions (uncovering and relying on the best explanations for understanding one's results) (Johnson and Onwuegbuzie, 2004). These qualities of pragmatism made it suitable to adopt the worldview in this study.

The researcher opted for pragmatism because it can be underpinned in mixed methods design as it emphasises more on the research problem and use of all approaches available to understand the problem. Additionally, pragmatism opens the door to multiple methods, different worldviews, as well as different forms of data collection and analysis. A pragmatic research approach is dismissive of dogmatic distinctions between quantitative and qualitative methods thus making it fit into the mixed methods. Pragmatists welcome mixed methods; numbers give meaning to narratives and narratives give meaning to numbers (Johnson and Onwuegbuzie, 2004 as cited by Hammond and Wellington, 2013). The choice of this philosophical worldview was also necessitated by its fundamental tenet that to know the meaning of a concept then we need to consider its practical consequences rather than to hold on to preconceived ideas.

### **3.3 Research Approach**

The mixed methods study was used to address the overall aim of exploring the feasibility of offering research data management services at the University of Zimbabwe. Creswell (2009) defines mixed methods research as, "...an approach to inquiry that uses two or more methods, processes and (in certain situations) philosophies in undertaking a research study. Mixed methods provide the combination of, typically, quantitative and qualitative methods in order to provide complementary and perhaps contrasting perspectives on a phenomenon (Hammond and Wellington, 2013).

The researcher adopted the concurrent mixed approach where qualitative and quantitative approaches were implemented simultaneously. The researcher applied the mixed methods in the construction of research questions where some questions-technological and economic requirements were addressed using mostly quantitative methodology and legal, organisational schedule requirements were addressed mostly in qualitative terms. Additionally, interviews were used to gather qualitative data from faculty librarians, university librarian and the library ICT manager and questionnaires were used to collect both quantitative and qualitative data from lecturers. The mixing of data was done at data collection, data analysis and data interpretation. In this study, quantitative and qualitative data was integrated by transforming the quantitative themes into counts and comparing these counts with descriptive quantitative data (Creswell, 2009). Using multiple, complementary data gathering and analysis techniques increased the meaningfulness and validity of constructs by capitalising on inherent method strengths and counteracting inherent method biases. The concurrent use of qualitative and quantitative data collection methods saved the researcher's time.

The design and establishment of research data management services is a complex issue involving multiple activities carried out by various actors addressing a range of drivers and influenced by a large set of factors (Pinfield et al., 2014). The authors added that RDM is socio-technical in nature therefore, the requirements of offering such services should be established from different angles and research methodologies. Given this complexity, the researcher found it befitting to employ the mixed methods approach to assess the viability of offering RDM services at the UZ library. Quantitative methodology was used to gain an understanding on the technological requirements, economic requirements and scheduling of RDM services. The objectives on organisational capabilities and legal obligations were of a social nature so they required the use of qualitative methodology. In this study, quantitative data was used to test the TELOS feasibility framework which predicts that its variables affect negatively or positively the collection, storage, preservation and access to research data at the UZ library. Qualitative data was used to explore the skills, knowledge and attitudes of librarians as well as



Quantitative and qualitative methods were also mixed in sampling, where the researcher used stratified random sampling (quantitative) to select researchers and judgemental sampling (qualitative) to select respondents from the library. In data analysis, the researcher mixed numerical and statistical data from the questionnaire and thematic and narrative data from the interviews in order to relate the findings from the researchers and library staff.

The researcher opted for the mixed methods because it enhances the validity of research findings by integrating information obtained from different methods of data collection for example integrating responses from researchers with what the library staff said in interviews. Additionally, mixed methods research has clear benefits in that it provides confirming, complementary and contrasting sources of data, very often as part of a strategy of triangulation. The use of both qualitative and quantitative methods enhanced the breadth and depth of inquiry results and interpretations by analysing them from the different perspectives of different methods and paradigms. Such attributes are essential in the current study because quite a number of factors are taken into consideration when assessing the viability of starting RDM services.

The study aimed to explore the feasibility of offering RDM services and is guided by multiple objectives and not all of them required to be explored with one method hence the need for mixed methods from the formulation of research questions to data collection and analysis. In this study quantitative data was collected from researchers using questionnaires and qualitative data was collected from faculty librarians, university librarian and the library ICT manager using semi-structured interviews. This provided for the enrichment of data and a complete understanding of the research problem through triangulation. The mixed methods design was also used by Van Tuyl and Michalek (2015) in their study on assessing RDM practices of faculty at Carnegie Mellon University.

### **3.4 Research Strategy**

Saunders et al. (2009) defined research strategy as the general plan of how the researcher will go about answering the research questions. On a similar note, Bryman (2008) describes research strategy as a general orientation to conduct research. The researcher adopted the case

study research strategy. A case study is an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-world context (Yin, 2014). Gray (2014) adds that a case study is “...an in depth study of a particular situation rather than a sweeping statistical survey. It is a method used to narrow down a very broad field of research into one easily researchable topic.” This approach provides tools for researchers to study complex phenomena within their contexts (Yin, 2014). In this study, adopting the case study strategy allowed the researcher to have an in-depth analysis of UZ library’s preparedness to offer RDM services by looking at the local context. The choice of using the case study strategy was triggered by the need to understand a real-world case (research data management) and assume that such an understanding is likely to involve important contextual conditions (technological, economic, legal, organisational and schedule) pertinent to the University of Zimbabwe (Yin and Davies, 2007 as cited by Yin, 2014). The case study methodology was used in this study to get in-depth information about the feasibility of offering research data management services. The University of Zimbabwe was selected as the unit of analysis so as to enable the researcher to gain in-depth insight into the feasibility of offering research data management.

Adopting case study strategy allowed the use of multiple sources of data collection and analysis, allowing the researcher to address the research objectives and answer the research questions satisfactorily. Ability to accommodate different research techniques, both qualitative and quantitative, is a salient feature of case study research. Neale, Thapa and Boyce (2006) state that the primary advantage of a case study is that it provides much more detailed information than what is available through other methods and also allows one to present data collected from multiple methods such as questionnaires and interviews. This gave the researcher greater flexibility in data gathering, since interviews and questionnaires were used in the study. The choice of the case study strategy was guided by the research questions and objectives, the extent of existing knowledge on the subject area to be researched, the amount of time and resources available, and the philosophical underpinnings of the researcher (Saunders et al., 2009). Based on the researcher’s underpinning philosophical views, the research was positioned within the philosophical viewpoint of a pragmatist which is compatible with the case study strategy.

### **3.5 Study Population**

Kurtz and Boone (2001) define study population as the total group of people that the researcher wants to study. Participants for the study were drawn from fulltime permanent lecturers (who will be referred to as researchers herein after) and the UZ Library. These were viewed as key players in the design of RDM programmes (Erway 2013, Jones et al. 2013, Pinfield et al. 2014; Wilson, 2010). The researchers were drawn from (nine) 9 faculties within the UZ. The researcher excluded the College of Health Sciences (CHS) from the study population because it is located off campus so given the limited time and financial resources, it was difficult for the researcher to include them in the study. The CHS researchers were also excluded from this study because in most cases they are in hospitals for their clinical practice. The target population of the study was 520 permanent fulltime lecturers from the selected faculties (UZ Human Resources department, UZ personal communication 24 February, 2017), nine (9) faculty librarians, the university librarian and the library ICT manager from the UZ library. There are different categories of faculty researchers at the UZ that include undergraduates, postgraduates and part time researchers. However, this study focused on fulltime permanent lecturers since they continuously conduct research as per UZ policy that all academics need to do research for tenure; most of them are the principal investigators for institutional and grant aided research projects and they are also responsible for supervising research projects for undergraduate and postgraduate students.

### **3.6 Sampling and Sampling Procedure**

Sampling is the act, process or technique of selecting a suitable sample of a representative part of a population for the purpose of determining parameters or characteristics of the whole population (Mugo, 2010). The researcher used both probability and non-probability sampling techniques to select study participants. The researcher used probability random sampling in particular stratified random sampling to select researchers. This was done to ensure that each researcher from each of the nine faculties had a known probability of being selected. Kumar (2014) defined stratified random sampling as a method in which the total study population is first classified into different subgroups based upon a characteristic that makes each subgroup more homogeneous in terms of the classificatory variable. The researchers were placed in

different strata according to faculty and thereafter respondents were randomly selected from each department. The selection of study participants from each subgroup was equal to its proportion in the total population as shown in Tables 3.1 and 3.2. Through stratified random sampling the desired representation from the various subgroups on the basis of faculty was assured. Tenopir et al. (2014) also used stratified random sampling to select participants for their study on research data management services in academic research libraries and perceptions of librarians.

The researcher used an online sample size calculator to calculate the sample size for researchers (<https://www.surveymonkey.com/mp/sample-size-calculator/>). From a population of 520 permanent fulltime lecturers, with a confidence interval of 94% and sampling error of 6%, a sample size of 168 was calculated. Table 3.1 below depicts the calculations for the sample size for each faculty. The researcher adopted proportionate stratified random sampling method which uses a sampling fraction that is proportional to the total population to ensure that each stratum contributes an evenly significant percentage towards the total sample (National Statistical Service, n.d as cited by Matingwina, 2015). The formula for the sampling fraction used is:

$$k = \frac{n}{N}$$

In this formula  $n$  is the sample size,  $N$  is the population size and  $k$  is the sampling fraction.

**Table 3.1 Sampling frame for researchers**

<b>Faculty</b>	<b>Population (Permanent fulltime lecturers)</b>	<b>Percentage contribution in the sample size of population</b>	<b>Sample Size</b>
Agriculture	52	10%	17
Arts	106	20%	34
Commerce	31	6%	10
Education	57	11%	18
Engineering	37	7%	12
Law	22	4%	7
Science	107	21%	35
Social Studies	84	16%	27
Veterinary Sciences	24	5%	8
<b>Total</b>	<b>520</b>	<b>100</b>	<b>168</b>

To ensure that there was equal representation of departments from each faculty, the researcher did further stratification by departments within faculties. For example, the faculty of Agriculture has four departments so from the proportionate sample size of 14; three researchers were randomly selected from each department. Calculations were rounded to the nearest whole number for decimal figures and remainders were allocated to departments with the highest numbers. Table 3.2 below illustrates the sample sizes from each department.

**Table 3.2 Sample Sizes from each department**

<b>Faculty</b>	<b>Population (Permanent fulltime lecturers)</b>	<b>Number of departments in the faculty</b>	<b>Sample Size</b>	<b>Participants from each department</b>
Agriculture	52	4	17	4
Arts	106	8	34	4
Commerce	31	4	10	2
Education	57	7	18	2
Engineering	37	6	12	2
Law	22	4	7	1
Science	107	10	35	3
Social Studies	84	8	27	3
Veterinary Sciences	24	3	8	2

Judgemental sampling method was used to select nine (9) faculty librarians, university librarian and the library ICT manager. Judgmental sampling also referred to as purposive sampling is as non-probability sampling method of sampling that is based on the idea that the interviewer or researcher already knows the required qualities from the interviewees thus making them the proper ones for the study (Kumar, 2011). According to Parahoo (1997), in non-probability sampling researchers use their judgment to select the subjects to be included in the study based on their knowledge of the phenomenon. Judgmental sampling was used because the researcher knew the limited number of individuals who were in the position to respond to questions on research data management. The advantage of judgemental sampling is that the researcher can identify participants who are likely to provide data that are detailed and relevant to the research questions (Olive, 2012). The sampling technique was also opted for because it is less time consuming and inexpensive. The faculty librarians, university librarian

and the library ICT manager were instrumental in this study as they would bring out organisational, economic, schedule, legal and technological requirements for setting up RDM services at the UZ Library.

### **3.7 Data Collection Instruments**

Data collection instruments are tools used in collecting research data or information needed to find solutions to the issues under investigation (Kothari, 2011). Semi-structured interviews and questionnaires were used as data collection instruments.

#### **3.7.1 Questionnaires**

The researcher used questionnaires to collect data from researchers. Gray (2014) defines a questionnaire as a series of questions which are systematically and objectively designed for the purpose of gathering information from respondents. The questionnaire developed for this study builds on previous similar research, and incorporates a wide range of recommendations and best practices from the international discourse on researcher's attitudes towards RDM. The questionnaire was designed building from previous similar studies on RDM such as those by (Kennan and Markauskaite, 2015; Linde et al., 2015; e-Infrastructure Austria, 2015). Additionally, the construction of the questionnaire was guided by the conceptualised model of the TELOS feasibility framework and the UK Data Archive Research Data Lifecycle model. The questionnaire contained both closed and open ended questions.

The questionnaire comprised of 33 questions which were related to the objectives and the conceptual framework of the study (see Appendix A). The questionnaire comprised of six broad sections which were:

**Section A Preliminary information**

**Section B Technological requirements for research data management**

**Section C Economic needs for research data management**

**Section D Legal obligations for research data management**

**Section E Organisational capabilities for and scheduling of research data management**

There was an open option which allowed respondents to add any other information which was deemed useful to the study. Each broad section had sub-questions which comprised of open and closed ended questions. Most questions offered codified answers in order to standardise assessment, including an open option (“other”).

Open ended and closed questions were used. Open-ended questions allowed the respondents to freely voice their experiences and minimise the influence of the researcher's attitudes and previous findings. The use of closed questions enhanced the comparability of answers and the responses were easy to process. Additionally, they clarified the meaning of a question to respondents. Ranking scales were also used to answer some of the questions where respondents were asked to rank their responses against given conditions. Five-point Likert scale was used where respondents were asked to rate their level of agreement to given statements. The questionnaire collected mostly quantitative data. Issues around researchers' attitudes towards RDM and the types of research data produced and storage requirements were covered in the questionnaire. Unique terms in the questionnaire were defined.

The use of the questionnaire was inspired by its ability to offer greater anonymity, thereby increasing the likelihood of obtaining accurate information. Additionally, the use of questionnaires is generally convenient and less expensive (Kumar, 2011). More so, questionnaires enabled the researcher to reach academic researchers in a geographically dispersed UZ community at a relatively low cost. Questionnaires were hand delivered to faculty researchers and follow ups were made via the departmental secretaries.

However, the use of questionnaires has its shortcomings that include lack of opportunity to clarify issues, incomplete responses and some questionnaires might not be returned. It is also difficult to ascertain if the intended respondent was the one who actually completed the questionnaire. These innate limitations of a questionnaire were mitigated by pretesting the instrument before administering it, contact email address for the researcher was provided so that those with questions could make follow ups and the researcher engaged faculty administrators to make follow-ups to increase the response rate.

### **3.7.2 Interviews**

An interview is defined by Burns and Groove (2001) as a verbal interchange in which an interviewer tries to elicit information, beliefs or opinions from another person. The researcher used semi-structured interviews to collect data from faculty librarians, university librarian and the library ICT manager. A total of 11 interviews were conducted. Nine (9) interviews were conducted with faculty librarians, one interview with University Librarian, and one interview with the library ICT manager. In a semi-structured interview the researcher asks a predetermined set of questions guided by the interview schedule (Kumar, 2011). The interview guide was developed with guidance from studies by Kennan (2016), Tenopir et al. (2015), Cox et al. (2015), Cox and Pinfield (2014) and Pinfield, Cox and Smith (2014) and Henderson and Knott (2014).

The research questions which were addressed by the interviews include organisational capabilities to offer RDM that include staff skills and knowledge about RDM, roles and activities that can be taken up by librarians in RDM and the attitudes of librarians towards RDM. The questions to address the technological requirements, schedule and legal obligations were also covered through the interview with the library ICT manager and the university librarian. Faculty librarians were asked about their skills and knowledge about RDM, the university librarian as one of the UZ administrators was asked mostly administrative questions that included legal issues, economic and technological, and the library ICT manager was asked questions mostly on technological requirements for setting up RDM services. From the interviews, mainly qualitative data was collected. The interview guides for faculty librarians, University Librarians and the library ICT manager are appended at the end of this study as Appendices B, C and D respectively.

The researcher booked for interview appointments with the interviewees. The researcher used a predetermined set of questions which contained both closed and open-ended questions which were first pre-tested for standardised wording, meaning and interpretation. The researcher conducted face-to-face interviews with faculty librarians and the library ICT manager; and rapport was established with the participants. The researcher introduced the aim and objectives of the study to the interviewees and they agreed to voluntarily participate in the interviews.



The data was recorded in the form of transcripts of conversations. The interview schedule for the university librarian was sent electronically via email as she was not available for a face-to-face interview. However, this method lacked the advantages of a face to face interview mentioned herein after.

Semi-structured interviews were chosen for this study because they are flexible, adaptable and provides direct human interaction that enable the researcher to probe and clarify answers with respondents, follow up leads, elaborate on the original response, and obtain more data with greater detail and clarity (Creswell, 2009). The interviewer followed the interview guide, but was able to follow topical trajectories in the conversation which were deemed appropriate. The researcher also obtained clues from respondent non-verbal communication on what the respondent felt, expected and thought. That actually helped in contextual interpretation of meanings and events as well as in detecting falsehoods and insincerity.

Limitations of interviews, however, included the possibility of researcher bias, coding of open-ended questions and time management during the interviews. To curb some of these limitations, the researcher conducted the interviews while maintaining an impartial and objective stance throughout the interviews. The researcher used a pre-coded interview guide to record responses and the researcher strictly adhered to the proposed timeline. The researcher had a sequence of questions and grouped them in themes that follow a logical progression.

### **3.8 Reliability and Validity**

Yin (2014) defines reliability as the extent to which results are consistent over time and an accurate representation of the total population under study. To ensure reliability of the study, the researcher chose a more 'neutral' time to distribute the questionnaires and conduct the interviews. Pretesting was done to test the reliability and content and internal validity of the data collection instruments. The questionnaire was adjusted based on the results from the pretesting exercise. Themes for the questions were established guided by the research objectives to enable flow of the questions. Creswell (2009) points that, if themes are established based on converging several sources of data or perspectives from participants, the process can be claimed as adding to the validity of the study.

Babbie (2001) defines validity as the extent to which an empirical measure adequately reflects the real meaning of the concept under study. Validity is often defined as the extent to which an instrument measures what it purports to measure. Content, internal, construct and external validity are crucial in measuring the validity of a research instrument. Content validity looks at the extent of coverage of areas under study. To ensure content validity, the researcher designed the questionnaire guided by the research objectives. Content validity was also enhanced by adopting questions from previous studies on RDM where the questions were proved to be valid (Yin, 2014). Internal validity is concerned with the presence of causal relationships between variables and results. To ensure internal validity, the researcher compared empirically observed patterns with either predicted ones or patterns established in previous studies and in different contexts. The questionnaire for researchers was created using suitable questions modified from related research and individual questions formed by the researcher. Construct validity is the extent to which a study investigates what it claims to investigate, that is, the extent to which a procedure leads to an accurate observation of reality. To enhance construct validity, the researcher used triangulation where both questionnaires and interviews were used to gather data and from different sources to look at the same phenomenon. The questionnaire and the interview guide contained both closed and open ended questions to enhance construct validity.

### **3.9 Ethical Considerations**

The researcher sought permission from the responsible UZ authorities to conduct this study. The aim of the study was explained to the interviewees and on the questionnaire it was clearly stated on the introduction. Kumar (2011) emphasised that in every discipline it is considered unethical to collect information without the knowledge of participants, and their expressed willingness and informed consent. Consent was sought from the respondents and they were informed that their participation in the study was voluntarily and that they were able to withdraw at any time. Privacy, anonymity and confidentiality were maintained by keeping researchers anonymous, and the researcher did not probe into beliefs, backgrounds, and behaviours in a way that reveals intimate private details and their responses were held in confidence. Privacy, anonymity and confidentiality were maintained in data collection, presentation and analysis where the researcher did not reveal the names of the participants and

the questions did not require respondents to provide identification information. There were no incentives which were offered to participants to take part in this study. Research findings were communicated in a clear and appropriate language which was not biased against persons because of gender, racial or any other attributes. The researcher avoided bias in reporting results by presenting the results objectively without hiding what was found. The researcher avoided plagiarism by giving credit to the works of others through referencing and citation.

### **3.10 Data Collection, Presentation and Analysis Procedures**

Data was collected using questionnaires and semi-structured interviews. The researcher made use of departmental secretaries and faculty librarians to distribute questionnaires to departments. Appointments were made for interviews. Quantitative data was presented using descriptive statistics, graphs, tables and charts to present salient information in a clear and accurate manner. Qualitative data was presented using narratives and themes were used to record responses from the interviews. The Statistical Package for Social Sciences (SPSS) was used for quantitative data analysis and Microsoft Excel was used to generate graphs for data presentation.

### **3.11 Summary**

The chapter highlighted the research philosophy, research approach, research strategy, study population and sampling techniques, data collection methods, validity and reliability and ethical considerations. The researcher opted for the pragmatic worldview as it accommodates multiple stances. The mixed methods design was adopted because of its ability to enhance the validity or credibility of research findings by comparing information obtained from different methods of data collection. The case study research strategy was used taking the UZ as the unit of analysis and research data management as the subject of inquiry. Questionnaires and semi-structured interviews were used as data collection methods. The study population was drawn from the UZ researchers (permanent fulltime lecturers) and participants were selected using proportionate stratified random sampling from 9 faculties. Faculty librarians, university librarian and the library ICT manager were selected using purposive sampling. The chapter looked at how validity and reliability of the study were enhanced through triangulation and

pretesting of data collection instruments. Ethical issues were considered during data collection, analysis and presentation. Permission to carry out the study was sought.

## CHAPTER 4

### DATA PRESENTATION AND INTERPRETATION

#### 4.1 Introduction

This chapter focuses on data presentation, analysis and interpretation of the study findings. The presentation and interpretation of the results was guided by the mixed methods methodology which was adopted by this study. Quantitative findings from the researchers were presented together with qualitative results from the UZ library staff in order to relate the results in order to come up with more coherent conclusions. Quantitative data was gathered from the closed ended questions of the questionnaire and it was analysed using SPSS. Qualitative data was collected from interviews with faculty librarians, university librarian and the library ICT manager. Data is presented in themes guided by the research objectives which were outlined in chapter one. The presentation of quantitative data is done in graphs, charts, tables and descriptive statistical narratives. Qualitative responses were grouped in themes and are presented as narratives.

#### 4.2 Background of Respondents and Response Rate

The researcher administered a total of 168 questionnaires across nine faculties and 104 questionnaires were returned which gave a 62% response rate. Table 4.1 highlights the response rate by faculty, period of service in the department, highest qualification and gender.

The results show that the highest number of respondents 34.6% (36) have been in the department for a period of 6-10 years, followed by 26.9% (28) with a period of 0-5 years. Moreover, 24.0% (25) have been in the department for a period of 11-25 years followed by 14.4% (15) of the respondents who have been in the department for over 15 years.

Respondents were asked to provide their highest qualification. The data shows that the highest number of respondents 35.6% (37) were PhD holders followed by 29.8% (31) DPhil/PhD candidates. Furthermore, 26.0% (27) of the respondents were Masters holders followed by

8.7% (9) who were Professors. The data shows that 62.5% (64) of the respondents were males and 37.5% (39) were females.

**Table 4.1 Response rate by faculty, highest qualification, period of service and gender**

<b>Item</b>	<b>Count</b>	<b>Percentage</b>
<b>Faculty</b>		
Agriculture	11	10.6
Arts	20	19.2
Commerce	6	5.6
Education	9	8.7
Engineering	9	8.7
Law	5	4.8
Science	22	21.2
Social studies	18	17.3
Veterinary sciences	4	3.8
Total	104	100
<b>Length of service in the department</b>		
0-5 years	28	26.9
6-10 years	36	34.6
11-15	25	24.0
15>	15	14.4
Total	104	100
<b>Highest qualification</b>		
Professor	9	8.7
Dr	37	35.6
DPhil/PhD candidate	31	29.8
Masters	27	26.0
Total	104	100
<b>Gender</b>		
Male	64	62.5
Female	39	37.5
Total	104	100

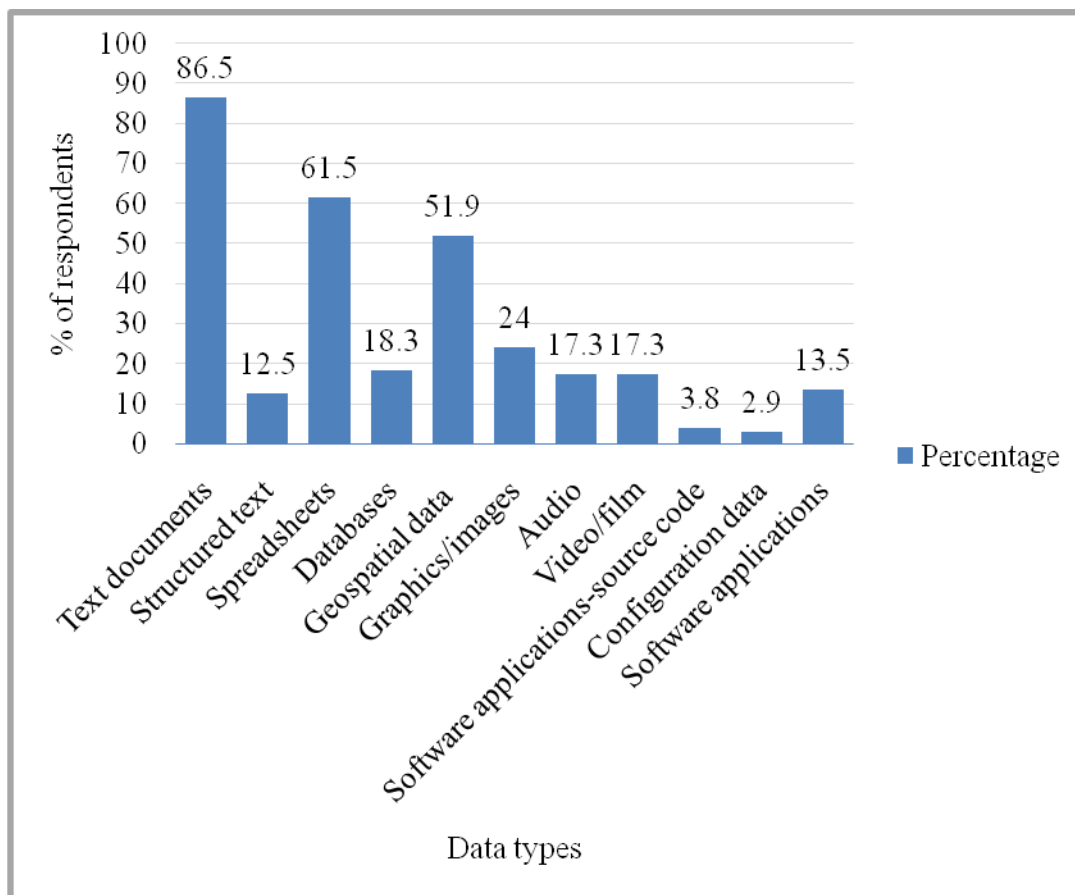
The researcher managed to conduct interviews with all the targeted interviewees who were comprised of nine (9) faculty librarians, the UZ librarian and the library ICT manager.

#### 4.4 Technological Requirements for RDM

Technological requirements for setting up RDM were addressed from four distinct angles namely types of data generated, data storage devices, quantity of research data generated and the quality of ICT infrastructure provided by the UZ.

##### 4.4.1 Types of Data Generated

Researchers were asked to provide the types of data formats generated in their research projects. Figure 4.1 shows the responses which were given on research data formats generated.



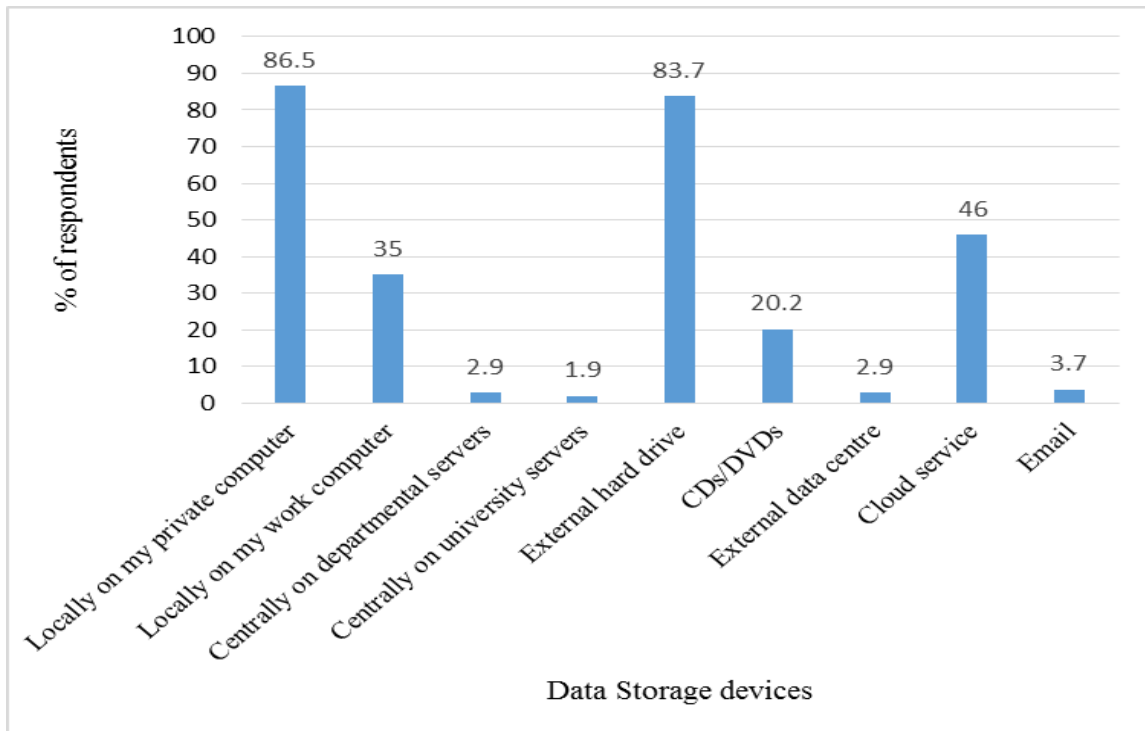
**Figure 4.1** Research data formats generated by researchers at the University of Zimbabwe

The data shows that the most common data type produced by researchers are text documents 87%, followed by spreadsheets 62%, geospatial data 52%, graphics/images 24%, databases 18%, audio and video/film were at par with 17% each, software applications 14%, structured text 13%, software applications-source code 4% and lastly configuration data at 3%.

Faculty librarians were asked if they were aware of the research data generated in their faculties. In response eight (8) out of nine (9) faculty librarians said that they were not aware of the data types generated by researchers from their faculties. One (1) faculty librarian who said was aware of the research data generated said that the information was acquired through informal means. Asked about the types of data that can be accommodated by the current ICT infrastructure, the library ICT manager said that text and structured text and to a lesser extent images.

#### 4.4.2 Data Storage Devices Used

Researchers were asked about where they regularly store their research data. Figure 4.2 below depicts the data storage devices used by researchers at the UZ.



**Figure 4.2 Data storage devices used by researchers at UZ**

The results in Figure 4.2 shows that the highest number of researchers stored their data in private computers (86.5%), 83.7% in external hard drives, 46% in cloud services, 35% made use of work computers, 20.2% said they use CDs/DVDs, 3.7% said they make use of email,

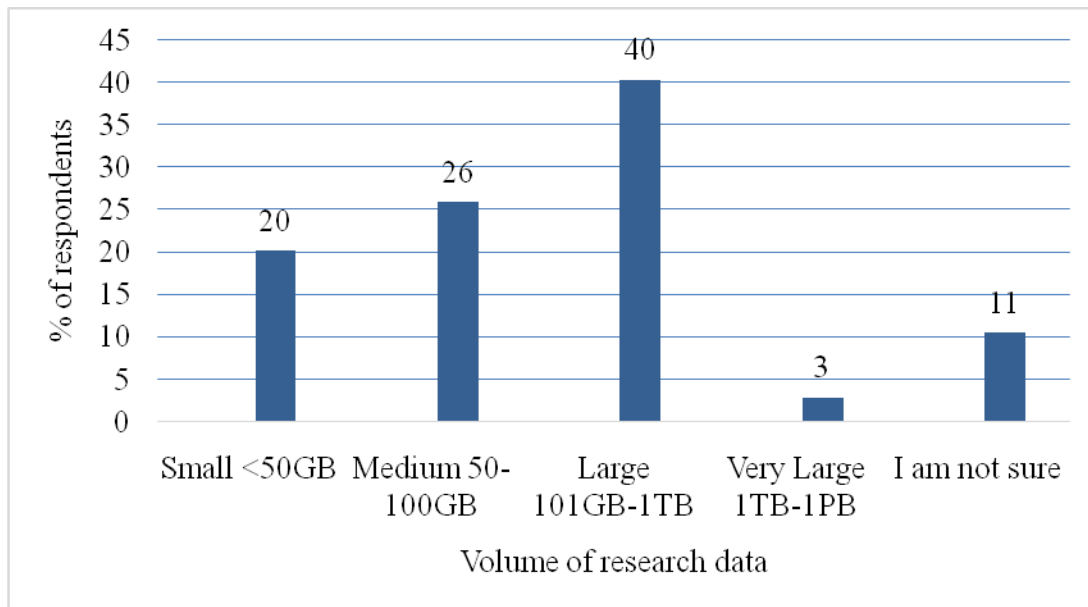


2.9% said they use an external data centre, 2.9% said they use departmental servers and lowest number 1.9% said they use university servers.

A similar question was posed to the library ICT manager about data storage devices that are used by the UZ library. The library ICT manager revealed that the library made use of distributed servers and some projects share servers that have a storage capacity of one Terabyte to store information for the repositories and other information that is relevant to the library. The manager added that for data backup, the library uses external hard drives. The ICT manager pointed that if the library is to start RDM services, the current servers can only store data but when it comes to preservation and data backups; there is need to upgrade the current servers.

#### 4.4.3 Quantity of Research Data Generated

Researchers were asked to estimate the total volume of research data on average per year. Figure 4.3 below shows the volume of research data produced by researchers at UZ per year.



**Figure 4.3 Estimated average volume of research data generated by UZ researchers**

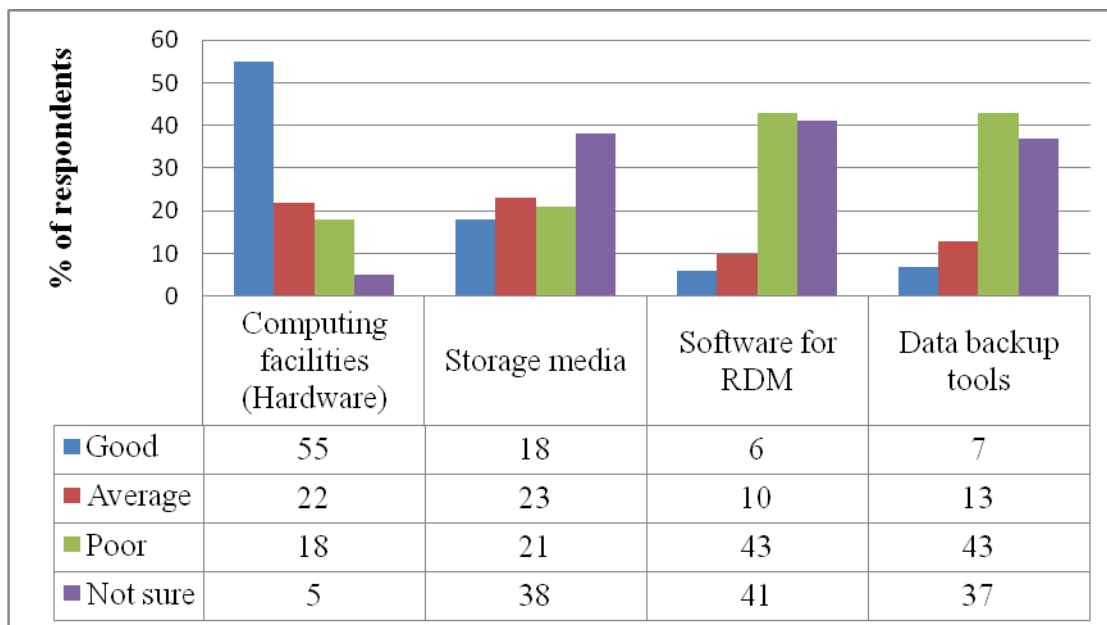
Data presented in figure 4.3 show that the highest percentage of respondents 40% generate large data sets that range from 101GB-1TB, this is followed by 26% of the respondents who

generate medium volume that ranges from 50-100GB. Furthermore, 20% of the respondents produce a small volume of research data that is less than 50GB. In the same vein 3% of the respondents generate very large data which ranges from 1TB-1PB. Lastly, 11% of the respondents were not sure about the amount of data that they generate.

Similarly, the library ICT manager was asked about the storage capacity range of storage devices. He pointed that each server had a storage capacity of one Terabyte and there were currently five servers.

#### 4.4.4 Quality of ICT Resources Provided by the UZ

Researchers were asked to rate the quality of ICT resources provided by the UZ. Figure 4.4 illustrates the responses which were given.



**Figure 4.4 Rating of UZ IT resources by researchers**

In terms of UZ computing facilities (hardware), 55% of the researchers said they were good, 22% said they were average, 18% said it was poor and 5% said they were not sure about the resources. On storage media, 18% rated it good, 23% said it was average, 21% said it was poor and the highest number of researchers 38% said they were not sure. Rating software for

RDM, 6% said it was good, 10% said it was average, 43% said it was poor and 41% said they were not sure. In terms of data backup tools, 7% of the respondents said it was good, 13% rated it as average, 43% said it was poor and 37% said they were not sure. One researcher reported that, “The UZ ICT policy gives preference to students compared to researchers. This stifles all efforts to improve RDM with a private laptop without a computer at the workplace.” Another researcher said that, “There is need to improve on Internet access to enable upload/download of data from online back-ups.”

A question was posed to the university librarian on whether the ICT infrastructure was adequate to support research data collection, storage, preservation and access. She responded that, “We are going to try and find out what we are able to do with the current ICT infrastructure. If we find the current infrastructure to be insufficient for our RDM initiatives, we will obviously invest in the infrastructure. We are going to find ways of mobilising the financial resources needed for example by engaging strategic partners.”

The library ICT manager was asked about the technological requirements for research data collection, storage, preservation and access, he pointed that there is need to have adequate hardware and software that supports data storage and preservation. He added that specialist software for data management is needed. The ICT manager added that there has to be extra storage devices for data backup if the library is to include preservation of data in the initial stages. The library ICT manager was further asked about the extent to which the current IT infrastructure was adequate to cater for RDM services. He said that in terms of software, existing open source DSpace can be used for a start but when it comes to storage and preservation of data, the current servers need to be upgraded. He was further asked about the better option for implementing RDM services between using existing institutional repositories and licences for data repositories. The ICT manager said that expanding existing repositories is cheaper but it does not come with the support that is with licenced data repositories. He added that for a start making use of existing repositories that use free and open source software is better as it minimises start-up costs. He said that the existing software and platforms that are used to manage the institutional, maps and digital collections repositories can be expanded to accommodate RDM services. The ICT manager added that the only

difference between the existing repositories and the proposed data repository is that the data repository will require extra security features to control data access for data protection since some data might be sensitive and confidential.

#### **4.5 Economic Needs for RDM**

Economic needs for setting up RDM services were addressed from four distinct angles namely funding for research projects, costs for setting up RDM, benefits of RDM and risks associated with RDM.

##### **4.5.2 Costs Incurred in RDM**

Researchers were asked if they incur any costs in managing their research data. The results showed that 71% (74) of the respondents incur costs, 29% (30) do not incur costs. For those who said they incur costs, 28% said they incur costs on data collection, 62% reported incurring costs on acquiring data storage devices, 20% reported incurring costs on acquiring software for data analysis. Researchers were asked if they had any funders for their research. The results of the study shows that the majority of the researchers do not receive funding for their research 79% (82) and 21% (22) reported that they receive funding for research. Among the 21% who said they receive funding for their research, 16% of the respondents said that they were funded through research grants from other organisations and 5% of the respondents received research grants from the UZ. Researchers were further asked if they budget for RDM in their initial research funding proposals. The responses showed that 68% (73) do not budget for RDM while 32% (31) budget for RDM.

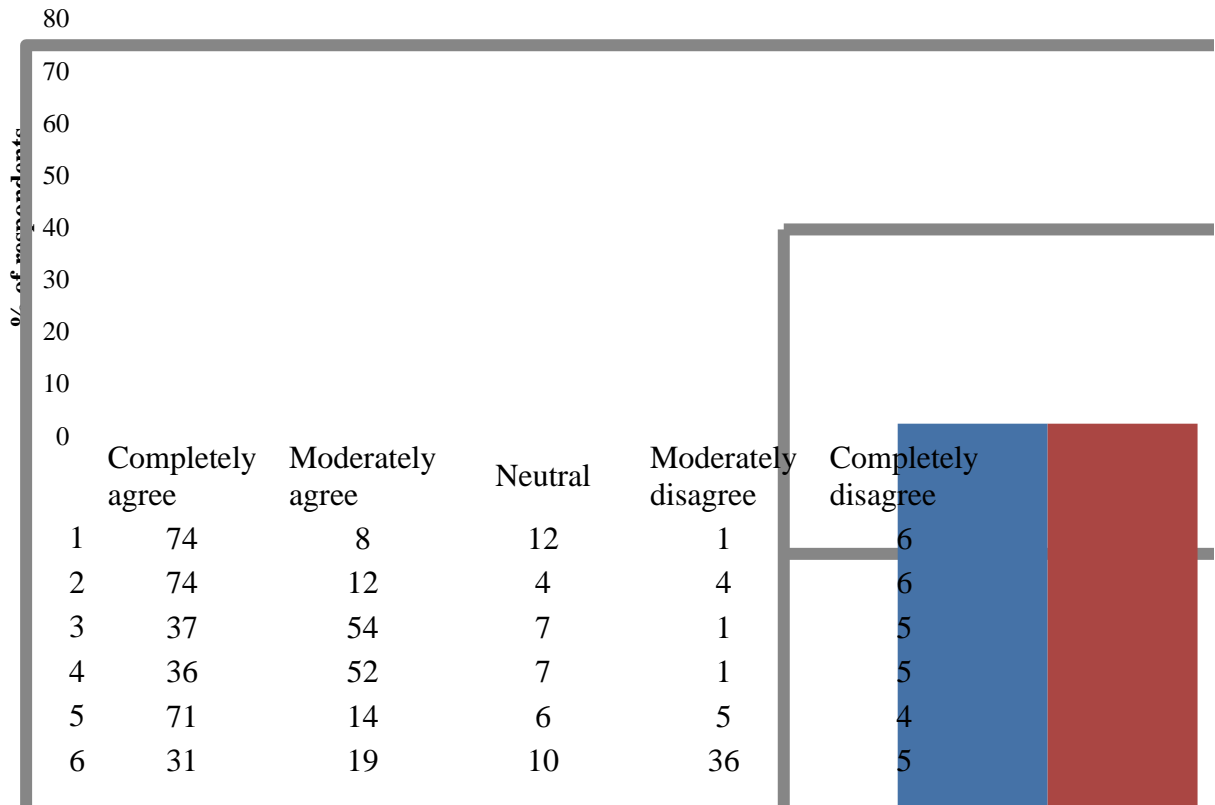
The university librarian was also asked about the anticipated costs for research data collection, storage, preservation and access services. She reported that the costs mainly revolve around the infrastructure needed especially for storage and preservation of data sets. “Staff training is another significant cost that we can consider given that RDM is an emerging field that has not been incorporated into the current curricula for library and information science.” A question was also posed to faculty librarians on the anticipated costs for RDM. The results showed that the costs which are likely to be incurred are start-up costs that include costs to acquire ICTs resources for RDM, staff training costs, costs for preparing data and metadata, and advocacy

costs. Three faculty librarians reported that there are also costs that could be incurred after the research project for data preservation and archiving. The ICT manager was also asked about the anticipated costs, he highlighted that, from an ICT perspective, if the library is going to use existing repository systems, the costs will be incurred in acquiring additional servers for data storage and preservation and training on RDM. Other costs highlighted by the ICT manager include the costs of maintaining and upgrading the data repository once it is set. The ICT manager pointed that there is need to acquire portable storage devices such as hard drives to enable the movement of data sets from researchers to the library for processing.

A follow up question was asked to the University Librarian on whether there were mechanisms in place to cater for the RDM costs. It was reported that there were plans to lobby for funds for RDM activities from central UZ administration as well as engage in strategic partnerships with other organisations/corporates with a keen interest in RDM.

#### **4.5.3 Benefits of Research Data Management**

Researchers were asked to rate their level of agreement to the benefits of RDM. Figure 4.5 below illustrates that 74% of the researchers completely agreed that data can be re-used to answer new questions and opens up new interpretations and discoveries, 8% moderately agreed, 12% were neutral, 1% moderately disagreed and 6% completely disagreed. On whether RDM helps to protect and preserve valuable data, 74% strongly agreed, 12% moderately agreed, 4% were neutral another 4% moderately disagreed and 6% strongly disagreed. Researchers were asked if RDM enhances the visibility of their research, 37% strongly agreed, 54% moderately agreed, 7% were neutral, 1% moderately disagreed and 6% strongly disagreed. On whether RDM makes the comparison and co-analysis of data from multiple sources easier, 36% strongly agreed, 52% moderately agreed to the notion, 7% were neutral, 1% moderately disagreed and 5% strongly disagreed. Researchers were asked if RDM provides an opportunity to check or verify findings, 71% strongly agreed, 14% moderately agreed, 6% were neutral, 5% moderately disagree and 4% strongly disagreed. Lastly researchers were asked if RDM makes their articles and papers more useful and citable, 31% strongly agreed, 19% moderately agreed, 10% were neutral, 36% moderately disagreed and 5% strongly disagreed.



**Figure 4.5 Benefits of RDM**

- Key:** 1. Data can be re-used to answer new questions and opens up new interpretations and discoveries  
 2. RDM helps to protect and preserve valuable data  
 3. RDM enhances the visibility of research  
 4. RDM makes the comparison and co-analysis of data from multiple sources easier  
 5. RDM provides an opportunity to check or verify findings  
 6. RDM makes your articles and papers more useful and citable

In the “Additional Information” section of the questionnaire, one researcher reported that, “Data management plans are very essential for researchers. It looks like a very innovative idea that can help researchers use and re-use their data.” Another researcher reported that, “Data management is a noble idea which should be promoted across university disciplines.”

A similar question was posed to faculty librarians about what they foresee as some of the benefits of RDM. One faculty librarian said that offering RDM enhances the relevance of academic libraries and it provides value addition to the services in the research continuum and

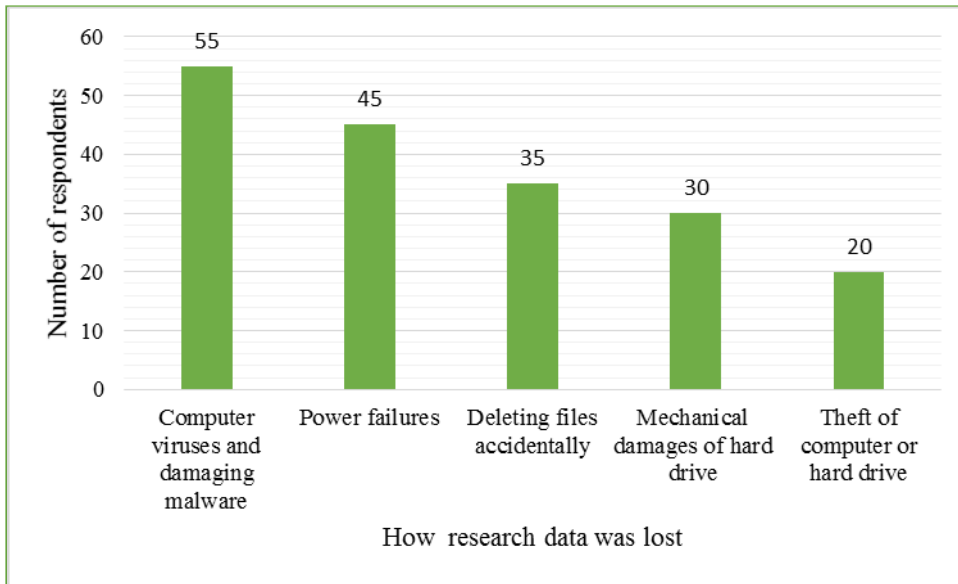
it is a best practice for the UZ as a university which aims to employ world class standards in research. All faculty librarians said that RDM is an emerging trend in academic libraries therefore, offering RDM services helps the UZ libraries to keep up with trends and offer all-encompassing research support services. Most faculty librarians said that RDM necessitates structured and easily accessible, legible, discoverable, visible and re-usable research data. Another response that emerged from eight faculty librarians was that RDM improves the quality of research by reducing data loss in research and makes it easy to verify and compare research findings. Most interview respondents said that RDM leads to long-term preservation and secure storage of research data which is critical in identifying trends in various academic disciplines.

The same question was posed to the University Librarian. The following response was given on the benefits of RDM,

*Research data management offer distinct benefits to academic staff and researchers, students, administrative services staff, and external collaborators and partners. For researchers, good data management enhances the visibility of research data and increases the number of citations, reduces the risk of data loss by keeping your research data safe and secure, facilitates the sharing and reuse of research data for future research and demonstrates research integrity and validation of research results. For institutions like the University of Zimbabwe, research data management can help to showcase the university's research data outputs to a global audience, attract new research collaborators and partners, strengthen the research environment and ensure that the university's researchers comply with the research data expectation of research funders and publishers.*

#### **4.5.5 Risks in RDM**

Researchers were asked if they have experienced data loss. The study revealed that 85 researchers have experienced research data loss and 19 said they did not experience any research data loss. Those who have experienced research data loss were further asked how they have experienced research data loss. Figure 4.6 below highlights the responses which were given on how researchers lost their data.



**Figure 4.6 Research data loss**

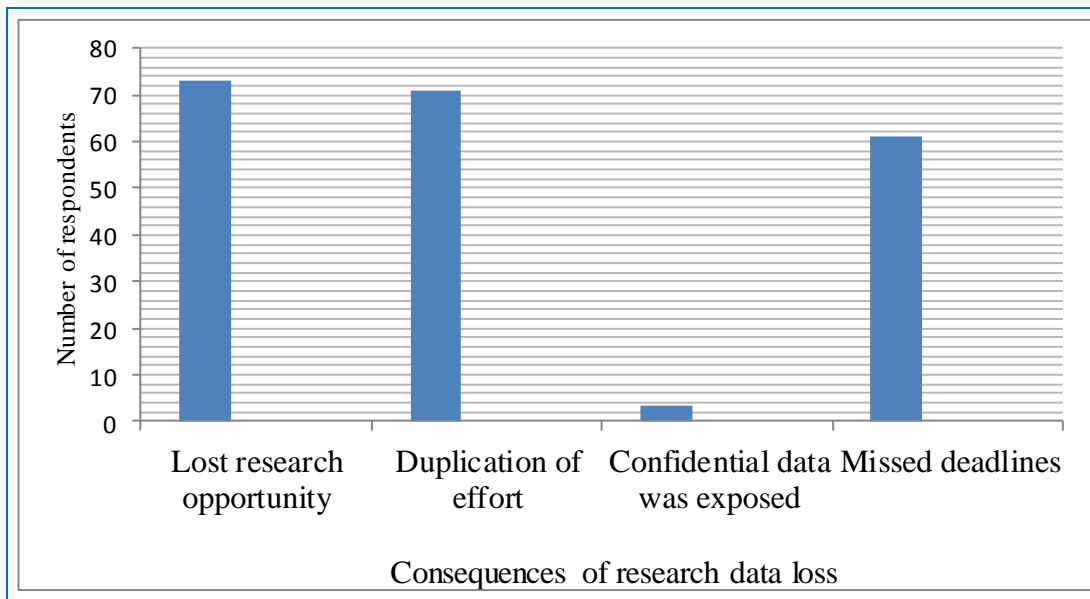
The results in figure 4.6 show that the highest number (55) of researchers lost their data through computer viruses and damaging malware, 45 lost through power failures, another 35 lost data by deleting files accidentally, 30 lost through mechanical damages of hard drive and 20 researchers lost through theft of computer or hard drive.

A similar question was posed to the library ICT manager on risks that are associated with digital data. He cited viruses and malware, hard drive failures, power outages and software failures. The ICT manager added that another risk associated with digital object is unauthorised access. A followed up question was asked on mechanisms that can be employed to mitigate the risks. He started by saying that losing data is costly and detrimental to the operations of the library as it affects service delivery. He said that there are periodic backups, the library has antivirus software and servers are protected with strong passwords and are in lockable rooms to avoid unauthorised access.

Researchers who reported that they have lost research data were further asked about the consequences of losing their research data. The study findings showed that the consequences of research data loss were; lost research opportunities (73), duplication of effort (71),



confidential data was exposed (11) and missed deadlines (61). Figure 4.7 depicts the responses which were given.



**Figure 4.7 Consequences of research data loss**

Researchers were asked if they have any data backup mechanisms for their research data, 45% said they have and 55% said they do not have. Those who said they have data backup mechanisms, 55% said they back up using external hard drives, 35% said they use cloud services and 10% said they use UZ servers.

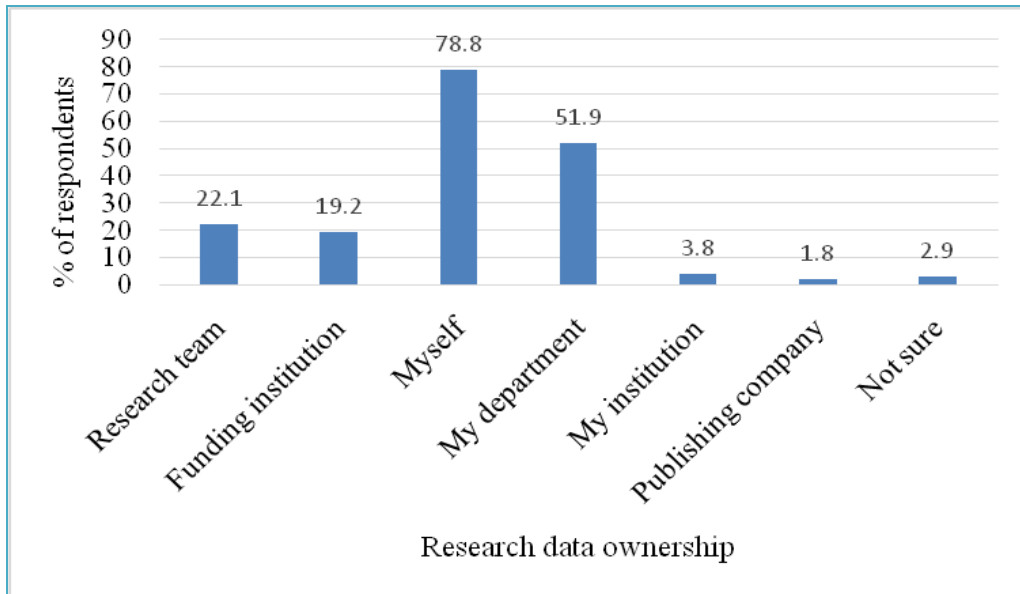
#### **4.6 Legal Obligations for RDM**

Legal obligations for setting up RDM services were addressed from two distinct angles namely research data ownership and policies for RDM.

##### **4.6.1 Research Data Ownership**

Researchers were asked about ownership of their research data. The study findings revealed that the highest number of respondents 79% (82) own data their data, 52% (54) have their research data owned by their departments. The results further shows that 22% (23) of the respondents have their data owned by the research team, 19.2% (20) said the funding institution, 4% (4) said my institution, 2% (2) said the publishing company and 2% (2) of the respondents were not sure about who owns their data. As additional information to the study,

one researcher mentioned that, “Some funding institutions monopolise the data especially where they provided the technical back up for the storage.” Figure 4.8 below shows the responses provided by researchers.



**Figure 4.8 Ownership of researchers’ research data at the UZ**

#### **4.6.2 Research Data Management Policies**

Researchers were asked if they had any legal concerns regarding the use of their research data. The results revealed that 51% had legal concerns while 49% did not have any concerns. For those who said they had legal concerns, most researchers (89%) cited intellectual property rights issues particularly copyright and 11% mentioned data protection issues.

On the same note, researchers were asked if they had any policies that govern research data management services. In response, the highest number of respondents 75% said they do not have any policies, 15% said they had policies and 10% did not respond to the question.

The university librarian was asked if there were enabling policies that were enacted for research data collection, storage, preservation and access. The librarian reported that, “The issue of copyright will be guided by the intellectual property policy of the University of Zimbabwe which I believe is already in place.” She further reported that other policies to

support the RDM activities will be enacted at the start of the implementation phase of RDM. We will of course be guided by international best practice when coming up with the policies, added the librarian.

Similarly, faculty librarians were asked about the perceived legal issues around the collection, storage, preservation and access to research data. One respondent mentioned that given the different types of research that are done at the UZ, there ought to be different policies that govern RDM. It was noted that there is consultancy research, research for academic purposes and research for publication. These types of research have different legal issues. It was also raised that some research is cumulative and the sharing of research data is difficult. All the faculty librarians mentioned that the most common legal issues are copyright for research data and data protection. It was also mentioned that in some cases there could be access restrictions for sensitive and commercially beneficial data. Two librarians pointed that the policies need to address issues data disposal and period of preserving the data. Other issues which were raised include plagiarism, ethical considerations regarding the use of and access to research data.

#### 4.7 Organisational Capabilities for RDM

Organisational capabilities for setting up RDM services were addressed from five distinct angles namely knowledge of and skills for RDM, researchers' attitudes towards RDM, management support, responsibility for RDM and stakeholders for RDM.

##### 4.7.1 Knowledge of Research Data Management

Researchers were asked to rate their knowledge of research data management. Table 4.2 below shows the responses which were given.

**Table 4.2 Knowledge of RDM by researchers at the UZ**

		Highest qualification				Total
		Professor	Dr	Phd/DPhil candidate	Masters	
Knowledge of RDM	Excellent	0	2	2	2	6
	Good	2	3	4	3	12
	Average	3	11	9	8	31
	Poor	3	19	16	10	48
	Very Poor	1	0	2	4	7
Total		9	37	31	27	104

The data shows that the highest number of respondents (48) had poor knowledge about RDM, followed by 31 who rated their RDM knowledge as average. Twelve (12) researchers rated their knowledge of RDM as good. Seven rated their knowledge of RDM as very poor and the remaining 6 rated their knowledge as excellent.

Similarly, faculty librarians were asked on whether they understood the concept of RDM. Six (6) faculty librarians reported that they have a general understanding of RDM while the remaining three (3) reported that they had no clue on what RDM refers to. The library ICT manager said that he had average knowledge about RDM. To enable the continuity of the interviews, the researcher explained in detail to the interviewees what RDM entails.

#### 4.7.2 Researchers' Attitudes Towards RDM

Researchers were asked if they need RDM services. Table 4.3 shows the responses which were given and this was moderated by qualifications

**Table 4.3 Researchers who require RDM Services**

		Designation				Total
		Professor	Dr	DPhil candidate	Masters	
Require RDM services	Yes	7	33	25	24	89
	No	2	2	1	1	6
	Not sure	0	2	4	2	9
Total		9	37	31	27	104

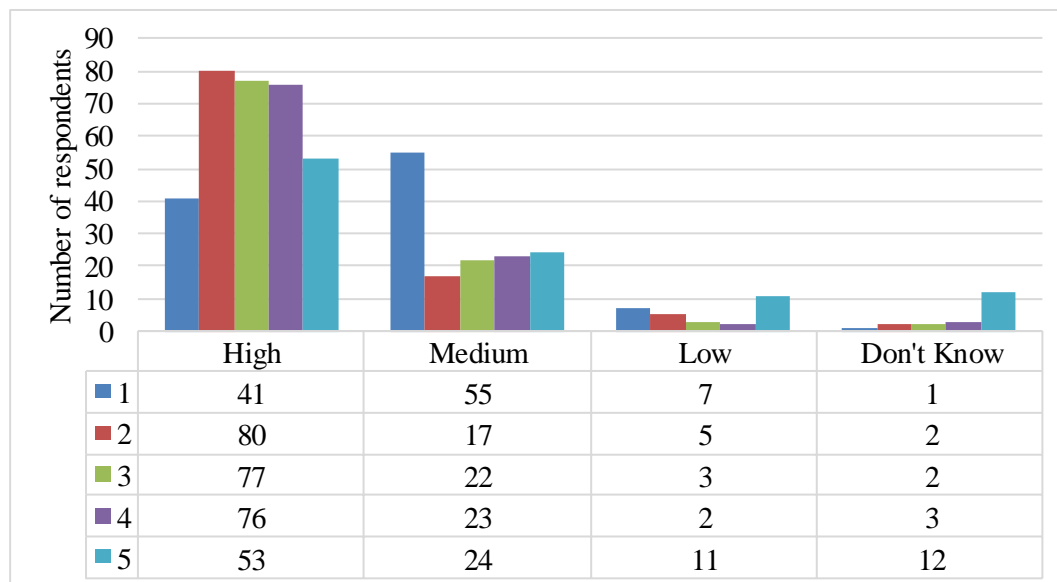
The highest number (89) of researchers said they require RDM services, followed by 9 who said they were not sure and 6 said they not require RDM services. The data shows that the highest number of those who require RDM services are DRs (33), followed by DPhil candidates (25), Masters holders followed with (24) responses and 7 professors require RDM services.

Faculty librarians were asked if they have received any requests for research data from library patrons. Only two (2) said they have received requests for research data. The rest reported that they only receive requests for final research outputs. One faculty librarian explained that the

reason why researchers do not request research data is not necessarily that they do not need research data but they are aware that the library currently does not provide RDM services.

A question was posed to researchers on whether they have ever been requested to submit research data management plans, 55% said yes while 45% said no.

Researchers were asked to rate their interest in RDM services. Figure 4.9 below shows that 41 researchers have a high interest in support in preparing data management plans for grant proposals, 55 have a moderate interest in the service, seven (7) showed a low interest in the service and one (1) respondent did not know about the service. On whether researchers have an interest in training on RDM, 80 showed a high interest, 17 medium interest, 5 showed low interest and 2 said they do not know about the service. Researchers' interests on data storage showed that 77 had a high interest, 22 moderate interest, 3 low interest and 2 said they did not know about the service. On data preservation, 76 researchers showed a high interest, 23 medium interest, 2 low interest and 3 did not know. On data access control, 53 researchers showed a high interest, 24 showed a medium interest, 11 showed a low interest and 12 said they did not know.



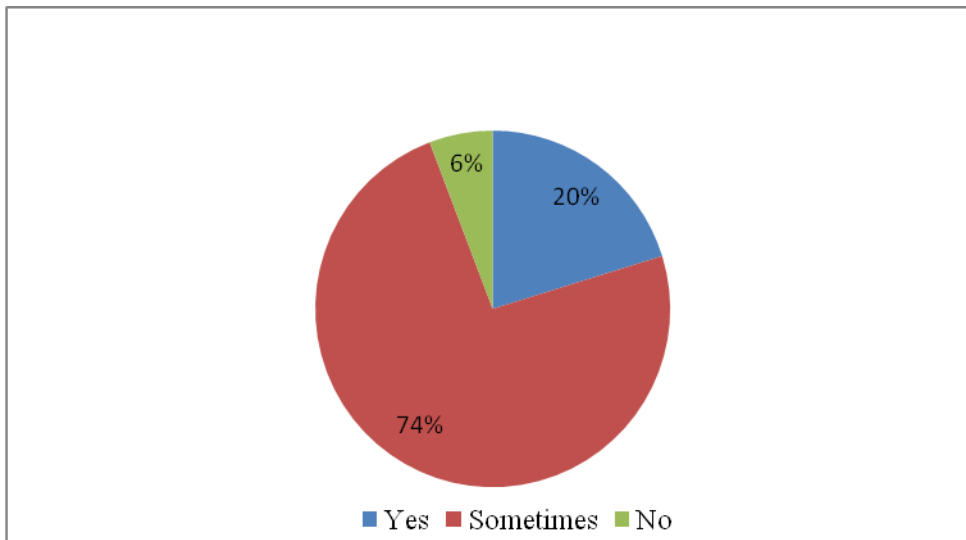
**Figure 4.9 Researchers' interests in RDM services**

**KEY:** 1. Support in preparing data management plans for grant proposals 2. Training on RDM. 3. Data storage. 4. Data preservation. 5. Data access control

Faculty librarians were asked what RDM services the UZ Library can currently provide. Faculty librarians reported that the library can start by offering informational services such as providing advice on writing data management plans, training on basic RDM. They argued that informational services are linked to other research support services that the librarians are providing such as training on literature searches and referencing and citation. Half of the faculty librarians were of the view that the library should also provide technical services such as data storage, data back-up, data preservation and giving access to research data. They argued that they can build on existing technical services such as content uploading on the institutional repository.

Similarly, the university librarian was asked about the RDM services that the library can begin with. She reported that, “The library has to be involved in the entire data cycle from planning, collecting, organising, managing, storage, security, backing up, preserving, and sharing data.” The ICT manager was asked the similar question and he said from a technical point of view, the library can start with data collection, storage and giving access to research data.

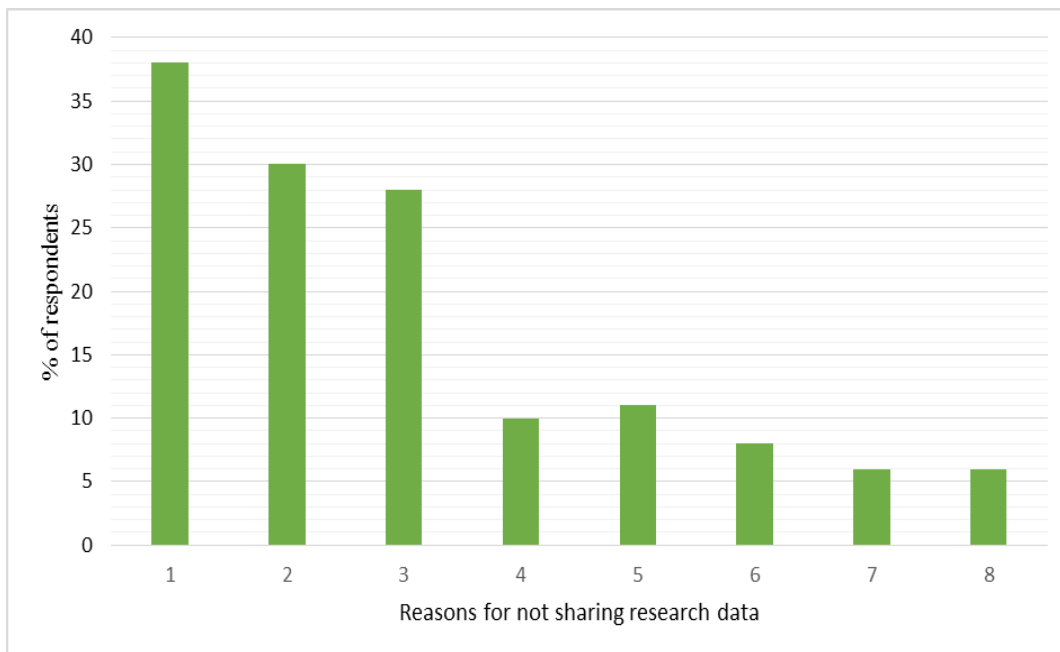
Researchers were asked if they share their research data with others. Figure 4.10 shows the responses which were given



**Figure 4.10 Research data sharing**

The results of the study showed that 74% of the researchers sometimes share their data, 20% said yes while 6% said no. Those who said they share and sometimes share their research data were further asked how they share their research data. On data sharing, 69 researchers said they share their data privately by email upon request, 18 said they submit data along with the article to a journal publisher, 12 respondents indicated that they share through a project's or intuition's website and 9 researchers reported that they share through an appropriate disciplinary data repository. One (1) researcher said that data is submitted together with final thesis for PhD studies as institutional requirements at their institution.

Researchers who reported not sharing and who sometimes share their research data were asked to provide their reasons for not sharing. Results showed that most researchers 40% (42) do not share their research data due to legal restrictions. Other reasons cited for not sharing research data were danger of misuse 30% (31), risk of misinterpretation and/or falsification of data 28% (29), increased competition in the 'publish or perish' game 7.7% (8), potentially undesired commercial use 9.6% (10), increased effort of time and/or cost 5.8% (6) and privacy violation 10.6% (11). Figure 4.11 demonstrates the reasons which were given by researchers for not sharing research data.



**Figure 4.11 Reasons for not sharing research data**

**Key:** 1. Legal restrictions (copyright, data protection e.t.c) 2. Danger of misuse 3. Risk of misinterpretation and/falsification of data 4. Potentially undesired commercial use 5. Privacy violation 6. Increased competition in the 'publish or perish' game 7. Increased effort in time and/or cost 8. Missing data standards

A question was asked to find out if researchers re-use their research data. This was controlled by qualification levels. The results showed that most researchers, (50) frequently re-use their research data. This is followed by 34 who occasionally re-use their research data, 9 researchers rarely re-use their data, 8 respondents always re-use their data and one (1) researcher never re-use the research data. The results also showed that the highest number of researchers who re-use (those who always, frequently, occasionally combined) is under DRs (36), followed by DPhil candidates (27), followed by those with Masters (22) and lastly Professors (7). Table 4.4 below shows the responses which were given:

**Table 4.4 Data re-use by researchers**

		Qualifications				Total
		Professor	Dr	DPhil/Phd candidate	Masters	
Re-use of research data	Always	0	4	2	2	8
	Frequently	7	22	10	11	50
	Occasionally	0	10	15	9	34
	Rarely	2	1	4	2	9
	Never	0	0	0	1	1
	Don't know	0	0	0	2	2
Total		9	37	31	27	104

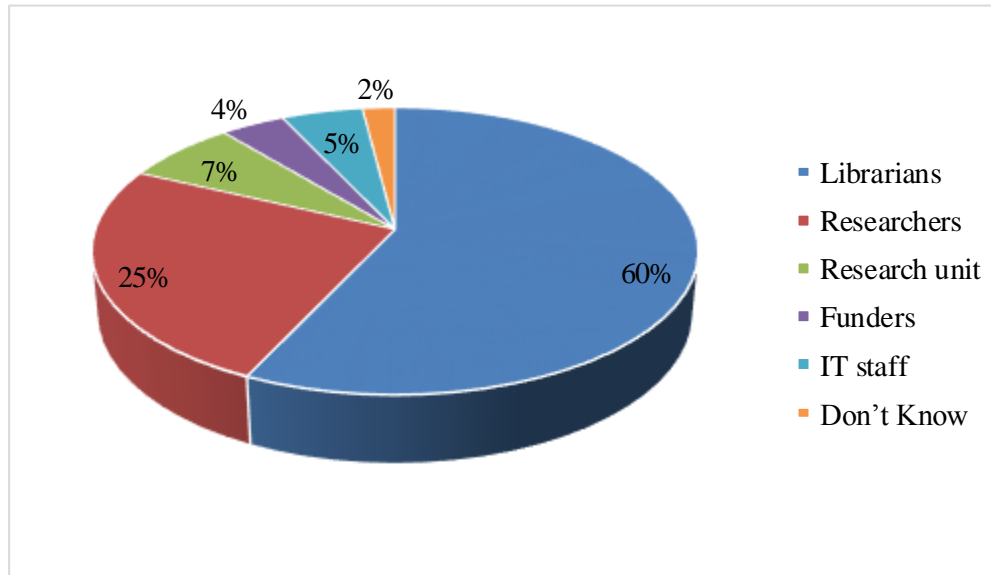
### 4.7.3 Responsibility for RDM

Researchers were asked about who should be primarily responsible for RDM. The results are presented in figure 4.12.

The results showed that more than half of the researchers (60%) said that the primary responsibility for RDM should be taken by librarians, 27% said researchers 7% said research unit, 5% said IT staff, 2% said they did not know. One researcher cited that, “Secure data



management services to one key functional research unit so that they provide insurance over data loss.”



**Figure 4.12 Primary responsibility for RDM at the UZ**

Furthermore, researchers were asked about who should be in the RDM team. All the respondents said that librarians, the research unit, funders, researchers and IT unit should be involved in RDM.

Similarly, faculty librarians were asked about who should be the main stakeholders for RDM. Faculty librarians said that librarians, the research unit, funders, researchers, IT department, university executive. Two faculty librarians said that external stakeholders such as the Research Council of Zimbabwe should also be involved so that they can advocate for the development of RDM at national level. The university librarian reported that the library would need to work with the producers of the data (academics and researchers). “This is important because academic and researchers would be required to come up with data management plans for their research as well as acquire the necessary skills for them to be able to produce meaningful and usable data”, explained the librarian.

### **4.7.3 Skills and Competences Needed for RDM**

After a brief description of what RDM entails, faculty librarians were asked about the kind of skills that are needed for RDM. Most faculty librarians reported that there was need for data curation skills (7), ICT skills (9), marketing and advocacy skills (6), scholarly communication (3), knowledge of the research process (4), metadata description (9), RDM workflows and systems (5), data recruitment (2) and understanding of the different types of research (3). Furthermore, the faculty librarians were asked if they possess the skills for data collection, storage, preservation and giving access to research data. All of them said they do not possess the skills. One faculty librarian reported that they might have ICT skills but it is not in the context of RDM. All faculty librarians agreed that they require training on the aforementioned skills in order to effectively execute RDM services. Half of the faculty librarians suggested that instead of leaving the responsibility of RDM with all the faculty librarians, there overall management of data should be done centrally. The University Librarian was asked whether the current set of librarians had the skills and knowledge on RDM in which she reported that they do not possess the skills. Asked on what can be done to enhance the skills and knowledge of librarians to take up RDM roles, the University Librarian said that some training on RDM is needed. The library ICT manager was asked if he had the necessary technical skills to set up the RDM platform, he reported that he can build up on existing skills, however, he pointed that since RDM is a new concept there was need for training.

### **4.7.4 Roles that Librarians can Take Up in RDM**

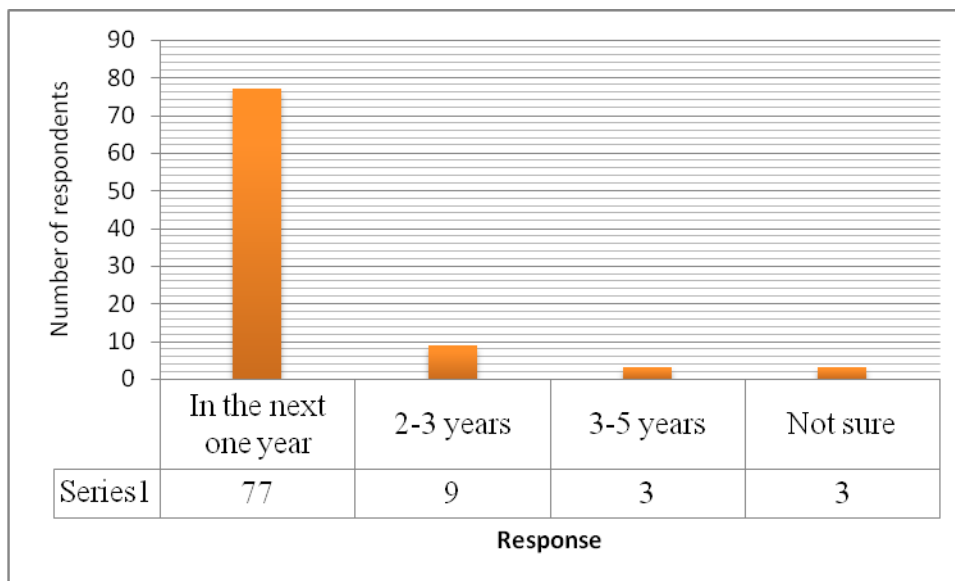
Asked on what RDM roles librarians can up, seven (7) said that librarians can take up advocacy roles. All faculty librarians identified training researchers as another role. Other roles which were identified include data storage, data preservation, data curation, RDM advisory roles, research data collection development and assist in the design of the RDM system. The University Librarian said that the library has to be involved in the entire data cycle from planning, collecting, organising, managing, storage, security, backing up, preserving, and sharing data.

#### 4.7.5 Management Support to RDM Services

Interview respondents were asked if there is adequate support from UZ top management on library projects. All the respondents agreed that top management renders adequate support to library projects in terms of infrastructure, financial and human resources. Regarding the proposed RDM programme, some librarians said that management support is need in securing funds for staff training, acquisition of ICT resources for data storage and preservation and in enacting and ratifying enabling policies. However, three (3) faculty librarians reported that when it comes to library policies, there is need for improvement. They cited an example of the Open Access Policy which was submitted in 2012 for adoption but until now it was not approved and there are no mandates despite concerted effort from the library. One faculty librarian emphasised that management support should be all encompassing if RDM is to succeed at the UZ.

#### 4.8 Scheduling RDM services at the UZ

Researchers were asked about how soon they require RDM services. The study revealed that the majority of the researchers (77) require RDM services in the next one year, 9 said they require RDM services in 2-3 years' time, 3 said in 3-5 years and 3 were not sure. Figure 4.13 shows the responses which were given.



**Figure 4.13** When researchers require RDM services

Similarly, the University Librarian was asked when the library intends to roll-out RDM services, she said in 2017. Asked about how long it takes to set up an RDM platform, the library ICT manager reported that if all the required resources are in place, it take a few days to set up the platform.

Researchers were asked on how RDM services can be implemented at the UZ. The data shows that over half of the respondents 70% (73) preferred a centralised multidisciplinary data repository. Ten percent (10%) (10) preferred a centralised discipline specific data repository, 7% (7) preferred a decentralised discipline specific data repository while 5% (5) preferred a decentralised multidisciplinary data repository. The remaining 9% (9) said they did not know how RDM services could be implemented. Table 4.5 illustrates the responses which were given.

**Table 4.5 How RDM services can be implemented**

<b>Type of data repository</b>	<b>Count</b>	<b>Percentage (%)</b>
Centralised discipline specific data repository	10	10
Centralised multidisciplinary data repository	73	70
Decentralised discipline specific data repository	7	7
Decentralised multidisciplinary data repository	5	5
I don't know	9	9
Total	104	100

A similar question was posed to the university librarian; she reported that the library exists to serve the entire university community so any data repository would have to cater for all disciplines. A similar question was posed to the ICT manager, he said it is better to start with a centralised multidisciplinary repository that encompasses all RDM services that include data storage, preservation and giving access to research data.

## **4.9 Data Interpretation and Discussion**

This section interprets the major research findings and shows how the results fit into previous studies.

The study sought to explore the feasibility of offering RDM services at the University of Zimbabwe library and it was guided by the following research questions:

1. What are the technological requirements for establishing research data management services at the University of Zimbabwe library?
2. What are the economic needs for setting up research data management services at the UZ library??
3. What are the legal obligations for the establishment of research data management services at the University of Zimbabwe library?
4. What are the organisational capabilities of the UZ library in providing research data management services?
5. How can research data management services be scheduled at the UZ library?

A general profile of researchers was important for this study as any assessment requires a clear picture of the characteristics and context of the beneficiaries. This is because some attitudes towards RDM may be influenced by educational qualifications, gender and period service in the department.

### **4.9.1 Information Technology Requirements for RDM**

The study findings showed that the technological requirements for data collection, storage, preservation and giving access include hardware and software that supports the storage, preservation and provision of access to research data. The findings are similar to those of Parsons (2013) who said that research data collection, storage, preservation and access is highly dependent on hardware and software used.

The study findings revealed that researchers at the UZ produce different types of research data. The findings were expected given the different academic disciplines that are at UZ. Most researchers reported that they generate data in text documents 87% followed by spreadsheets 62%. The findings are consistent with those of Tuyl and Michalek (2015) who found that the most common data types at the Carnegie Mellon University were text documents and spreadsheets. The data types may be considered low-difficulty formats for preservation purposes since there are clear transformations or migration pathways for these formats and the UZ libraries have extensive experience working with these types of documents. It was found that a few researchers 4% and 3% generate software applications source code and configuration data respectively. These two categories might pose more difficult to preserve, especially when considering the potential for obsolescence of formats and software applications for using these files and the need for an emulation strategy to preserve a functional environment to run them in the future (Van Tuyl and Michalek, 2015).

In terms of data storage devices, the study revealed that researchers stored their data in privately owned external hard drives and computers. Similarly, Schumacher and Vandecreek (2014) conducted a study for faculty members at five American universities to find out how they were managing their research data. Faculty members' responses revealed that most managed their data themselves, relying on some combination of individual devices and storage services. However, personal data storage devices maybe regarded as insecure given the vulnerabilities that face ICT devices. Assessing whether the UZ library had enough storage devices for research data, the library ICT manager said they have distributed servers that can be used for research data storage.

The study revealed that researchers produce different quantities of research data. Forty percent (40%) said they produce large quantities of data that range from 101GB-1TB. The findings are similar to those of Kennan and Markauskaite (2015) who conducted a study which sought to provide an insight into the research data management and sharing practices of academics at ten universities in New South Wales, Australia. The study revealed that most participants (40.2%) indicated that they generate large quantities of data that ranges from hundreds of Megabytes to Terabytes. In light of the research question on technological requirements, this

means that if the UZ library wants to successfully introduce RDM services, they need to invest in servers with a high storage capacity servers accommodate the huge quantities of research data that are generated by the researchers.

Researchers rated UZ computing facilities (hardware) as good (55%) and on storage media the highest number of researchers were not sure (38%), software for RDM (43%) and data back-up tools (43%) were rated as poor by the highest number of researchers. The library ICT manager reported that the existing ICT infrastructure can only support some RDM activities such as data storage. The university librarian reported that the library will find out if the existing technology can accommodate RDM and if not it was highlighted that mechanisms will be put in place to develop infrastructure for RDM. The librarian's sentiments were in line with those of Jones, Pryor and Whyte (2013) who highlighted that the technical infrastructure of institutional repositories can be extended to enable the curation of data without the development or purchase of an entirely new software platform. The results showed that the UZ library can extend existing repositories to incorporate RDM services. This is in accordance with a previous study by Jones, Pryor and Whyte (2013) who highlight that institutional repositories have typically been created to store research publications rather than data, but their technical infrastructure can be extended to enable the curation of data without the development or purchase of an entirely new software platform.

#### **4.9.2 Economic Needs for RDM**

The results of the study showed that economic needs for research data collection, storage, preservation and access include costs for acquiring, maintaining and upgrading ICT resources and staff training. This study produced results which corroborate previous findings by Hole et al. (2010) who assert that the costs involved in setting up RDM include costs for staff recruitment and training, consumables, ICT equipment and maintenance. It is interesting to note the UZ library has mechanisms in place to cater for the anticipated costs for setting up RDM. The finding is in agreement with Berman's (2008) assertion that during the onset of the RDM project, it is imperative to ascertain who will bear the costs of setting up and maintaining the service. In terms of the costs that are incurred in RDM, 69% of the researchers reported that they incur costs for data management activities. The results showed that researchers incur costs in data collection, acquiring data storage devices and acquiring software for data analysis.

The study revealed that most researchers (79%) do not receive funding for their research projects (79%). The study also showed that 68% of the researchers do not budget for RDM in their initial research funding proposals. The findings differ from recommendations given by Davidson et al. (2014) who advised that with increasing mandates to retain research data for the mid to long term period, researchers must now make informed decisions at the grant application stage about both the in-project data management costs and their potential to retain the data for as long as their funder requires, in most cases a period of at least ten years. These findings show that if the UZ library introduces RDM services which are perceived to be free to the UZ community, there is likely to be buy-in from the researchers.

It is interesting to note that researchers at UZ viewed RDM as beneficial in their research endeavours. The results showed that the majority of researchers agreed (as shown in figure 4.5) that data can be re-used to answer new questions and opens up new interpretations and discoveries, RDM helps to protect and preserve valuable data, RDM enhances the visibility of their research, RDM makes the comparison and co-analysis of data from multiple sources easier, RDM provides an opportunity to check or verify findings and RDM makes their articles and papers more useful and citable. The faculty librarians and the university librarians also showed that RDM is beneficial to both the researchers and the university at large. These results match those observed in earlier studies by Jones, Pryor and Whyte, (2013); Pryor (2012) and Heidorn, (2011) who pointed that properly managed and shared data have the potential to yield manifold benefits throughout and beyond the life of a research project when reused in primary research, follow-up, and synthesis studies, as well as in interdisciplinary and data-intensive research.

The study findings showed that most researchers (82%) have experienced research data loss. It was revealed that research data was lost mainly due to computer viruses and damaging malware (55), through power failures (45), deleting files accidentally (45), through mechanical damages of hard drive (34) and theft of computer or hard drive (20). The study findings also revealed that digital data faces the risks of hacking, viruses and malware, hardware and software failures and power cuts. The study findings showed that the consequences of research data loss were; lost research opportunities (73), duplication of effort



(71), confidential data was exposed (11) and missed deadlines (61). A possible explanation to this result could be that researchers lose their data because there are no data storage and backup systems that are provided by the university and the personal data storage devices used are vulnerable to attacks. In accordance with the present results, a previous study by Schumacher and VandeCreek (2015) showed that 31 (55.3%) of participants indicated that they have lost research data and been unable to replace them with backup files in the course of their professional career. This highlights that providing RDM services will be a noble idea at UZ given the rate at which researchers lose their data and the undesirable consequences. The study revealed that there are mechanisms in place at the UZ library to mitigate the risks.

The results of this study showed that most researchers at UZ do not have data backup mechanisms in place. This means that in any case of research data loss, researchers will face negative results in terms of financial resources, opportunities and time. The findings are similar to those of Baccarini and Melville (2011) who highlighted that losing research data in academia may lead to dire effects which are detrimental to the success of a research project.

#### **4.9.3 Legal Obligations for RDM**

The study findings showed that the legal requirements for the collection, storage, preservation and access to research data include copyright issues, institutional RDM policy, access control, and data protection. These findings are similar to those of Fitzgerald and Pappalardo (2007) who assert that the legal framework for a RDM service should cater for copyright, privacy issues and data protection and access. One of the issues that emerge from these findings is that just like the final intellectual products such as journal articles and textbooks, research data must be protected by policies. The study findings show that most researchers do not have any policies that govern their RDM activities. It was revealed that there were intentions to develop RDM policies by the UZ library at the start of the implementation of RDM services. The results show that the UZ Library is in the direct path of planning to develop an RDM policy before implementation. This is in accordance with the recommendations given by Patel (2016) who recommends that before embarking in the RDM journey there has to be an institutional RDM policy that clearly spell out the following in the context of data management: 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, terms and conditions of the use of data, protection of confidentiality of sensitive data, protection of data against security breaches and intellectual property concerns. Interestingly, it was found that the UZ has an Intellectual Property Rights policy which caters for copyright issues.

The results from the study showed that research data is owned by different players that include research team, researchers themselves, funding agencies, the departments and the UZ. This can be attributed to the fact that there are different types of research that take place at UZ that include consultancy research, research for tenure and research for adding knowledge. These results match those observed in an earlier study by Corall (2015) who asserted that the question of who should own the data raises so many questions given the number of players involved in data generation.

#### **4.9.4 Organisational Capabilities for RDM**

The results of the study showed that librarians need to possess the following skills in order to take up RDM roles: ICT skills, digital curation, advocacy and training skills. Similarly, Cox and Pinfield (2014) found that the specific skills needed for RDM include data curation skills, technical IT skills and knowledge of research methods and the need for disciplinary knowledge. The findings of the study revealed that currently the UZ library staff do not possess the requisite skills and knowledge for research data collection, storage, preservation and access. The findings fit the researcher's anecdotal understanding that none of the faculty librarians had any course in RDM during their training as the area was not covered in the past LIS curriculum. The result shows that librarians need to be trained on all the activities that take place throughout the research data lifecycle. The findings reinforce previous results of a baseline survey of academic librarians by Tenopir et al. (2011) which showed that many academic librarians do not feel prepared to take on RDM new roles in spite of plans by their libraries to offer RDM services. Similarly, Cox and Pinfield (2014) conducted a study which revealed that libraries had patches of the relevant skills but that they were not seen as widely enough spread.

The results of this study showed heterogeneity of the participants' knowledge of RDM especially among researchers and library staff. This result can be attributed to the different types of research that is conducted by researchers at UZ. For instance some researchers are conducting research which is funded by organisations that require data management plans thus they know what RDM entails. The study findings also showed that faculty librarians had partial knowledge about RDM. The responses given were qualified acknowledging that faculty librarians needed RDM skills. Similarly, Cox and Pinfield (2014) found that there was an RDM skills and knowledge deficit in most academic institutions.

The results of the study revealed that researchers at UZ were interested in RDM and the associated services. This can be attributed to the fact that most researchers were at some point requested to submit data management plans, they incur costs in RDM and they have previously lost research data. This seemingly obvious result however, may help to better understand the services that researchers need in the design of RDM programmes. Asked about what RDM services the library can currently provide, there was an agreement from interview participants that the library should provide both informational services and technical services that include data storage, preservation, back up and providing access. However, the findings from this study are in contradiction with previous studies by Tenopir et al. (2017) on RDM services in European libraries which revealed that academic research libraries are more likely to offer consultative-type RDM services than hands-on/technological services. Consultative services frequently involve a personal client-librarian relationship and inform the client about such things as how to find information on data management plans, metadata standards, or data citation practices. From previous studies, it can be noted that the lower and slower up-take of technical services compared to consultative services may reflect the fact that these services require a substantial investment in time, resources, and new technical knowledge.

The study finding showed mixed views from researchers regarding their willingness to share their research data. This result is expected given that some researchers at UZ conduct research which is of commercial value and some data is sensitive to share in data repositories. The main reason why researchers are reluctant to share their data could be the fact that most researchers occasionally and frequently re-use their research data. The present findings seem

to be consistent with previous findings by Keil (2014) in a study on research data needs from academic libraries taking a the perspective of a faculty researcher which revealed that sharing data can be particularly unnerving to researchers who may perceive a loss of a competitive edge for their next follow-up manuscript or grant proposal.

The results of this study revealed alternative opinions on who should take up the primary responsibility for research data management at the UZ. However, most researchers pointed that the primary responsibility for research data management should be bestowed upon librarians. The results fits the researcher's understanding that librarians are traditionally known as the key custodians of information resources that is why researchers commonly agreed that the primary responsibility for RDM should be with librarians. The results resonate the findings of previous studies which acknowledged that librarians are well positioned to play an important role in RDM (Kennan and Markauskaite, 2015; Tenopir et al., 2015; Erway, 2013; Jones, Pryor and Whyte, 2013; Lyon, 2012; Corral, 2012; Alvaro et al., 2011.).

The findings from this study showed that stakeholders for RDM should include librarians, the research unit, funders, researchers and IT department. The explanation to this result could be that the management of research data is a joint effort and brings together different services, roles and responsibilities. Given the complexity and multi-disciplinary nature of RDM, there ought to be various players to contribute to the success of the programme. The findings of this study are in accordance with previous findings by Chiware and Mathe (2015) and Jones et al. (2013) who pointed that RDM partners at the institutional level commonly include the library, information technology (IT) services, and the office of research, records and archives services and research units and centres.

The study findings revealed that there is management support for library projects especially those that support research. These findings indicate that if the UZ library is to start RDM services, they will receive support from top university management. The findings are in accordance with those of Tenopir et al. (2015) who found that management support is important for securing funding for RDM since libraries do not always have the resources to implement such a programme on their own.

#### **4.9.5 Scheduling RDM Services**

The study findings revealed that most researchers prefer a centralised multidisciplinary data repository. Likewise, the interview respondents all agreed that there is need to provide a service that caters for all academic disciplines so that other faculties will not feel left out. On the contrary, previous studies provided that there are factors such as the availability resources, skills and technology that have a bearing on whether an institution should take a campus wide approach or departmental approach to RDM (Shen and Varvel, 2013). There are mixed views from literature on how RDM services can be rolled out. One of the main themes that arose from previous studies is the need to start with a small number of RDM services, and then look to expand those services (Toups and Hughes, 2013; Wright et al., 2014). However, research has shown that this approach may prove to be problematic if the demand for the service is high. The findings also tally with recommendations by Jones (2012) who said that libraries can begin by offering or mediate secure storage for dynamic and static research data in co-operation with institutional IT units and/or seek exploitation of appropriate cloud services.

The results of the study showed that most researchers require RDM in the next year. It is interesting to note that from the university librarian's view the UZ library had plans to offer RDM services in 2017. Asked about how long it takes to set up the RDM platform, the library ICT manager said it takes a few days if all the necessary resources are in place. On the contrary, Steeleworthy (2014) argues that developing an IT and staffing strategy to provide the technology and support that maintain an RDM programme is not always feasible in the short-to-medium term.

#### **4.10 Summary**

This chapter showed the presentation, analysis and interpretation of results. The first section showed the presentation of results followed by the interpretation of results. Results from the questionnaires were presented first followed by interview responses. Research findings were presented in tables, graphs, descriptive statistics and narratives. The results were organised into themes around the research objectives. Findings showed that in terms of technology, there is necessary hardware and software that can accommodate textual research data for a start since it does not require too much storage space. The results of the study showed that most

researchers at UZ generate large quantities of data that range from 101GB-1TB and it is mostly in textual format. It was also found that existing ICT infrastructure can be extended to accommodate RDM and if it is not enough, there was commitment from the Library administrator that they will source resources to finance RDM activities. It was found from the study that making use of existing repositories is cheaper but having a licence for a data repository is desirable though it is costly. The study revealed that the costs involved in setting up RDM services include staff training costs, costs of acquiring ICT resources and advocacy costs. Copyright, institutional RDM policy and data protection were the main legal obligations which were revealed. The study findings showed that there was an RDM skill deficit among librarians, researchers agreed that RDM should be primarily bestowed among librarians. It was also revealed that there was UZ management support on library projects. Librarians, IT staff, researchers and funders were highlighted as the major stakeholders in setting up RDM services. The results of the study showed that the preferred mode of implementing RDM services at the UZ was to start with a centralised multi-disciplinary data repository.

## CHAPTER 5

### SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

#### 5.1 Summary

The study was conducted under the background that the UZ Library was planning to offer research data management services to keep up with emerging trends in academic librarianship and extent its research support service portfolio. The library's preparedness was questioned thus the study aimed at exploring the feasibility of offering research data management services at UZ. The study objectives were guided by the TELOS feasibility model and the UK Data Archive Research Data Lifecycle Model which helped to bring out the key activities involved in RDM. The objectives of the study sought to identify the technological requirements for establishing RDM services at the UZ library. Additionally, the study sought to identify the economic needs for setting up RDM services at the UZ library. Moreover, this study sought to review the organisational capabilities of the UZ library in offering RDM services. Another objective was to establish the legal obligations for providing RDM services at the UZ library. Lastly, the study sought to find out how RDM services can be scheduled at the UZ library. Ideally it would have been better to target all the ten faculties at UZ but due to constraints of time and financial resources, the researcher focused on a selected sample from nine faculties. It is hoped that the findings of the study might be beneficial to the UZ Library, LIS profession and it might add to the body of existing literature. Literature was reviewed to set the tone for this study and orient the researcher on previous studies on RDM.

The study adopted a pragmatic worldview, mixed methods research design and a case study research strategy to answer the research questions. The case study research strategy was used taking the UZ as the unit of analysis and research data management as the subject of inquiry. Questionnaires and semi-structured interviews were used as data collection methods. The study population was drawn from 138 UZ researchers (permanent fulltime lecturers) and participants were selected using proportionate stratified random sampling from nine (9) faculties. Participants from the library were selected using purposive sampling and nine faculty librarians, the University Librarian and the library ICT manager were selected.

The study findings revealed that the UZ library had the necessary technological resources to accommodate some RDM services. The study findings also revealed that there are costs that are involved in setting up RDM services and there was willingness from the University Librarian to secure funds for the new services. Upon inquiry it was found that the legal environment for RDM needs to be strengthened. On organisational capabilities, the study showed that researchers were willing to submit some of their research data and librarians did not possess the necessary skills and knowledge to take up RDM roles. The study showed that the library can start with a centralised multidisciplinary data repository which can be implemented in the next one year.

## **5.2 Conclusions**

This section presents conclusions for each research question. The conclusions were drawn from the findings of the study and the discussion from Chapter 4.

### **5.2.1 What are the technological requirements for establishing research data management services at the University of Zimbabwe library?**

The results of the study showed that technological requirements for setting up RDM services at the UZ include hardware and software for research data storage and preservation that include servers. However, the study findings showed that the available technological infrastructure was able to accommodate research data collection, storage and access. It can be concluded that given the technological resources available in the UZ library, the library was ready to offer research data collection, storage and access. The library was not ready to offer data preservation services. The results of the study revealed that researchers at UZ produce different types and quantities of research data with textual data being the most prominent. The results from the library showed that the technological infrastructure can accommodate textual and structured data and it had a storage capacity of 5-10 TB. It can also be concluded that the UZ library can start by storing textual data as it does not require a lot of storage space. Another conclusion drawn from the study is that technology is vital when considering offering RDM services.



### **5.2.2 What are the economic needs for setting up research data management services at the UZ library?**

From the study findings, it can be concluded that the economic needs for setting up RDM services at UZ include staff training costs and costs of acquiring ICT resources. The study showed that there was willingness from the library management to secure funding for RDM services. It can therefore, be concluded that the UZ library had adequate economic resources to support RDM activities. From the findings, it can also be concluded that ascertaining the costs and how they will be met is pertinent in the development of successful RDM services.

### **5.2.3 What are the legal obligations for the establishment of research data management services at the University of Zimbabwe library?**

The legal requirements for setting up RDM services include copyright, data protection and institutional RDM policy. The study revealed that the UZ library had not enacted any of the policies for research data collection, storage, preservation and access. The researcher concluded that the legal framework on RDM was still very much in a flux at the UZ library therefore, it was not ready in as far as RDM legal issues are concerned.

### **5.2.4 What are the organisational capabilities of the UZ library in providing research data management services?**

The findings of the study revealed that the basic skills that are needed to take up RDM roles include digital curation, data storage, preservation, advocacy and policy development. The results of the study revealed that librarians do not possess the requisite skills to take up RDM roles. It can therefore, be concluded that the UZ library was not ready in terms of skills to offer research data collection, storage, preservation and access. The study also showed that researchers need RDM services. Thus, it can be concluded that there is likely to be buy in from researchers if the UZ library implements RDM services. The study findings show that the establishment of RDM services involves librarians, IT staff, funders, and researchers. It can be concluded that the responsibility of offering RDM does not solely lie with the library. Most researchers highlighted that the primary responsibility for RDM should be with the library. It can be concluded that the UZ library should champion the introduction of RDM activities.

### **5.2.5 How can research data management services be scheduled at the UZ library?**

The study findings showed that the UZ library exists to serve the entire university community. It was also revealed from the study findings that the library has to be involved in the entire data cycle from collection, storage, preservation and access to research data. It can be concluded that when implementing RDM services, the library need to provide support throughout the research data lifecycle. The findings of the study also showed that the preferred type of the data repository was a centralised multidisciplinary data repository. It can therefore, be concluded that when implementing RDM services, the library need to consider all the academic disciplines and the data repository should be centralised. The study findings also showed that RDM services are needed as soon as possible. The researcher concludes that the UZ library was in the right track in considering to offer RDM services as it is demand from researchers.

### **5.2.6 Relevance of the conceptual framework**

The findings of the study showed that the TELOS feasibility framework and the UK Data Archive Research Data Lifecycle Model are essential in assessing institutional readiness to offer RDM services. The models managed to bring out the requirements for setting up RDM services as well as the activities involved in RDM.

### **5.2.7 Overall Conclusion**

The UZ library was not fully ready to take up RDM services. In terms of the technological, scheduling and economic requirements for setting up RDM services the UZ library was prepared. However, the UZ library was not prepared in terms of the RDM legal obligations and skills that are required to take up RDM activities. Given the current technological, economic and schedule environment, the library is prepared to offer research data collection and storage services. Researchers at UZ require RDM services so the library was in the right path in planning to offer RDM services.

## **5.3 Recommendations**

The following recommendations which have both policy and practical implications were drawn based on study findings:

## **Recommendations for the University of Zimbabwe and the UZ Library**

1. The UZ library needs to upgrade its ICT infrastructure in order to incorporate all the RDM activities.
2. The UZ library needs to make use of existing repositories in order to cut on RDM start-up costs.
3. The UZ library needs to form strategic partnerships with organisations that are already offering RDM services so that they can learn from the best practices.
4. Setting up RDM services is costly therefore, there UZ library needs to have fundraising activities for the RDM programme.
5. The UZ library needs to actively participate in institutional research data policy development, including resource plans. Encourage and adopt open data policies where appropriate in the research data life cycle.
6. RDM is a new concept at the UZ so there is need to organise library staff training from RDM basics throughout the activities of the research data lifecycle. Where possible, the library should send staff for exchange programmes to institutions that have established RDM programmes.
7. There is need to raise awareness and educate researchers and the UZ community by offering training on research data management and provide adequate tools and structural support for research practice.
8. RDM is an activity that involves an assortment of players. The UZ library needs to engage all the key stakeholders to foster an interoperable infrastructure for data access, discovery and data sharing.
9. The UZ library should set up an RDM working group rather than assigning responsibility to one person.
10. The UZ library should develop procedures and related activities, clarify front and back office roles throughout the research data lifecycle.
11. The UZ library should support the lifecycle for research data by providing services for storage, discovery and permanent access.
12. An RDM position paper should be prepared and presented to the UZ executive so that when the library requests for funding from central administration, they will all be in a clear position.

### **Recommendation for the LIS Profession**

The researcher recommend that before embarking on the RDM journey, there is need to first conduct a thorough feasibility analysis in order to bring out the requirements given that RDM is a complex issue involving multiple activities carried out by various actors addressing a range of drivers and influenced by a large set of factors.

### **5.4 Recommendation for further research**

This study took a general approach to the TELOS framework in exploring the feasibility of offering RDM services at the UZ library. There is need for further studies that provide an in-depth analysis of the technological, economic, legal, organisational and schedule feasibility of offering RDM services. Additionally, the study focused on digital research data management only; it is therefore, recommended that further studies be conducted covering the management of non-digital research data at UZ.

## REFERENCES

- Addis, M. and Arkivum Ltd. 2016. RDM: The costs of long-term retention and access to research data. Available at: <http://www.arkivum.com/wp-content/uploads/2014/04/ResearchDataWebinar3.pdf> (Accessed 2017 January 23)
- Adidu, F.A. and Olannye, P.A. 2006. Basic Small business entrepreneurship: A modern approach. Agbor: Royal Pace Publishers.
- Alvaro, E., Brooks, H., Ham, M., Poegel, S. and Rosencrans, S. 2011. E-Science Librarianship: Field Undefined. Issues in Science and Technology Librarianship (66). Available at: <http://www.istl.org/11-summer/index.html> (Accessed 2017 February 10)
- Association of Research Libraries. 2010. E-science and data support services: A study of ARL member institutions. Association of Research Libraries, Washington, DC. Available at: <http://www.arl.org/storage/documents/publications/escience-report-2010.pdf> (Accessed 2017 February 23)
- Auckland, M. 2012. Re-skilling for research: An investigation into the role and skills of subject and liaison librarians required to effectively support the evolving information needs of researchers. London: Research Libraries UK. Available at: <http://www.rluk.ac.uk/wp-content/uploads/2014/02/RLUK-Re-skilling.pdf> (Accessed 2017 February 10)
- Babbie, E., 2001. The practice of social research. Cape Town: Oxford University Press.
- Baccarini, D. and Melville, T. 2011. Risk management of research projects in a University context-An exploratory study. Available at: [http://epublications.bond.edu.au/cgi/viewcontent.cgi?article=1016&context=aubea\\_2011](http://epublications.bond.edu.au/cgi/viewcontent.cgi?article=1016&context=aubea_2011) (Accessed 2017 January 25)
- Ball, A. 2012. Review of Data Management Lifecycle Models. Available at: <http://opus.bath.ac.uk/28587/1/redm1rep120110ab10.pdf> (Accessed 2016 December 12)
- Bause, K., Radimersky, A., Iwanicki, M. and Albers, A. 2014. Feasibility studies in product development. Procedia CIRP. [Volume 21](#), Pages 473-478. Available at: <http://www.sciencedirect.com/science/article/pii/S2212827114006684> (Accessed 2016 December 10)
- Beagrie, N., Chruszcz, C. and Lavoie, B. 2008. Keeping Research Data Safe: a cost model and guidance for UK universities. Available at: <https://www.webarchive.org.uk/wayback/archive/20140615221657/http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pdf> (Accessed 2017 January 23)
- Beagrie, N. 2011. A guide to the KRDS benefits framework. Available at: [https://www.beagrie.com/static/resource/KRDS\\_BenefitsFramework\\_Guidev3\\_July%202011.pdf](https://www.beagrie.com/static/resource/KRDS_BenefitsFramework_Guidev3_July%202011.pdf) (Accessed 2017 January 23)

Beagrie, N., Eakin-Richards, L. and Vision, T. 2010. Business models and cost estimation: Dryad Repository case study. In *iPRRES2010 Vienna Conference Paper, September. Österreichische Computer Gesellschaft, Vienna* (pp. 365-371). Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.384.7168&rep=rep1&type=pdf#page=365> (Accessed 2016 December 16)

Berman, F. 2008. Got data?: a guide to data preservation in the information age. *Communications of the ACM* 51, 50.

Boriwondo, M. 2016. *Research Data Management Services at the Women's University in Africa Libraries*. Bulawayo: National University of Science and Technology.

Bowen, D.J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Bakken, S., Kaplan, C.P., Squiers, L., Fabrizio, C., et al. 2009. How We Design Feasibility Studies. *Am J Prev Med* 36, 452–457

Brandt, D. S. 2007. Librarians as partners in e-research: Purdue university libraries promote collaboration. *College & Research Libraries News*, 68(6). Available at: <http://crln.acrl.org/content/68/6/365.full.pdf> (Accessed 2017 January 10)

Brown, M.L. and White, W. 2014. Case Study 2: University of Southampton - a partnership approach to research data management, in Pryor, G., Jones, S. and Whyte, A. (Eds.), *Delivering Research Data Management Services*, Facet Publishing, London, pp. 135–162.

Bryman, A. 2004. *Social research methods*. 2nd ed. Oxford: Oxford University Press

Buchhorn, M., and McNamara, P. 2006. Sustainability Issues for Australian Research Data. The Report of the Australian e-Research Sustainability Survey Project. Available at: <http://www.apsr.edu.au> (Accessed 03 December, 2016)

Burnhum, A. 2012. Research data definitions. Accessed at: [http://www2.le.ac.uk/services/research-data/documents/UoL\\_ResearchDataDefinitions\\_20120904.pdf](http://www2.le.ac.uk/services/research-data/documents/UoL_ResearchDataDefinitions_20120904.pdf) (Accessed 2016 December, 3)

Burns, N. and Grove, S. K. (2001). *The Practice of Nursing Research: Conduct, Critique, and Utilization*. 4th ed. Philadelphia: W. B. Saunders.

Carlson, J. 2012. Demystifying the data interview: Developing a foundation for reference librarians to talk with researchers about their data. *Reference Services Review*, 40(1), 7–23. doi:10.1108/00907321211203603

Cambridge English Dictionary. 2016. Researcher. Available at: <http://dictionary.cambridge.org/dictionary/english/researcher> (Accessed 2016 August 28)

Cheek, F.M. and Bradigan P.S. 2010. Academic health sciences library research support. *Journal of the Medical Library Association*, 98 (2), pp. 167–171 Available at: <http://dx.doi.org/10.3163/1536-5050.98.2.011> (Accessed 2017 February 23).

Chioma, E.F., Chioma, O. and Augustus-Daddie, J. 2016. Effect of feasibility study on the growth of SMEs in Imo-state: a study of selected SMEs in Owerri zone. *International Journal for Research in Business, Management and Accounting*. Vol. 2(2). Available at: <http://internationaljournalsforresearch.com/Pdf/International%20Journal%20for%20Research%20in%20Business,%20Management%20and%20Accounting/BMA-FEB-2016/BMA-FEB-10-2016.pdf> (Accessed 2017 January 23)

Chiwere, E., and Mathe, Z. 2015. Academic libraries' role in Research Data Management Services: a South African perspective. *South Afr. J. Libr. Inf. Sci.* 81, 1.

Choudhury, S. 2012. ASERL webinar: Data conservancy and data management services at Johns Hopkins University Libraries. Available at: <https://vimeo.com/36050328> (Accessed 2017 February 28)

Corrall, S. 2012 Roles and responsibilities: Libraries, librarians and data. In: Pryor G, editor. *Managing research data*. London: Facet. pp. 105–133.

Corrall, S., Keenan, M.A. and Afzal, W. 2013. Bibliometrics and research data management services: Emerging trends in library support for research. *Library Trends*, 61 (3) (2013), pp. 636–674 Available at: <http://dx.doi.org/10.1353/lib.2013.0005> (Accessed 2017 January 6)

Cole, G.J. 2016. Establishing a research data management service at Loughborough University.

Collis, J. & Hussey, R. 2009 *Business Research: A practical guide for undergraduate and postgraduate students*. 3rd edition. New York: Palgrave Macmillan

Cox, A.M., Kennan, M.A., Lyon, L., and Pinfield, S. 2017. Developments in research data management in academic libraries: Towards an understanding of research data service maturity. *J. Assoc. Inf. Sci. Technol.* Available at: <http://eprints.whiterose.ac.uk/101389> (Accessed 2017 February 23)

Cox, A. D. and Knight, G. 2015. Building a research data management service for the London school of hygiene & tropical medicine. *Program: Electronic Library and Information Systems* 49, 424–439.

Cox, A.M. and Pinfield, S. 2014. Research data management and libraries: Current activities and future priorities. *J. Librarians. Inf. Sci.* 46, 299–316. Available at <http://journals.sagepub.com/doi/abs/10.1177/0961000613492542> (Accessed 23 October, 2016)

Cox, A.M. Pinfield, S. and Smith, J. 2016. Moving a brick building: UK libraries coping with research data management as a 'wicked' problem. *Journal of Librarianship and Information Science*. Vol. 48(1) 3–17. Available at: <http://eprints.whiterose.ac.uk/80934/11/RDM%20wicked%20problem%20JOLIS%20published%20version%201603%20VoR.pdf> (Accessed 2017 February 23)

Cox, A.M., Verbaan E, Sen, B.A. 2013. Upskilling liaison librarians for research data management. *Adriadne* (70). Available at: <https://www.ariadne.ac.uk/issue70/cox-et-al> (Accessed 2017 October 23)

Cox, A. D., Searle, S., Wolski, M., Simons, N., and Richardson, J. 2015. Librarians as partners in research data service development at Griffith University. *Program: Electronic Library and Information Systems* 49, 440–460.

Creamer, A. 2015. Current issues and approaches to curating student research data. *Bulletin of the Association for Information Science and Technology*, 41(6), 22–25. doi: 10.1002/bult.2015.1720410610

Creswell, J.W. 2009. *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks: SAGE Publications.

DataOne. Best practices. <https://www.dataone.org/best-practicesv> (Accessed 5 October, 2016)

Davidson, J., Jones, S., Molloy, L. and Kejser, U.B. 2014. Emerging good practice in managing research data and research information within UK Universities. *Procedia Computer Science*, 33, pp.215-222. Available at: <http://www.sciencedirect.com/science/article/pii/S1877050914008254> (Accessed 2016 December 3)

Davies, R., Ayris, P., McLeod, R., Shenton, H. and Wheatley, P. 2007. How much does it cost? The LIFE Project-costing models for digital curation and preservation. *Liber Quarterly*, 17(3-4). Available at: <https://www.liberquarterly.eu/articles/10.18352/lq.7895> (Accessed 23 January, 2017)

De Rosa, M., Giovannini, R., Giuliani, S., Lunati, G. and Taddeo, M. (N.D). Feasibility study for the creation of the digital library. Available at: [http://www.iccu.sbn.it/upload/documenti/Esdf\\_Summary.pdf](http://www.iccu.sbn.it/upload/documenti/Esdf_Summary.pdf) (Accessed 2017 January 25)

Delserone, L.M. 2008. At the watershed: Preparing for research data management and stewardship at the University of Minnesota libraries. *Libr Trends* 57: 202–210. Available at: [http://muse.jhu.edu/journals/library\\_trends/v057/57.2.delserone.html](http://muse.jhu.edu/journals/library_trends/v057/57.2.delserone.html). (Accessed 2016 September 23)

Der, A. 2015. Exploring the academic libraries' readiness for research data management: cases from Hungary and Estonia.



- Digital Curation Centre (DCC). 2016. Data lifecycle model. Available at: <http://www.dcc.ac.uk/resources/curation-lifecycle-model> (Accessed 2017 February 17)
- Doty, J., Kowalski, M. T., Nash, B. C. and O’Riordan, S. F. 2015. Making student research data discoverable: A pilot program using Dataverse. *Journal of Librarianship and Scholarly Communication*, 3(2), 1–25. doi: 10.7710/2162-3309.1234
- E-infrastructures Austria. Questionnaire national research data survey. Available at: [https://e-infrastructures.at/fileadmin/user\\_upload/p\\_e\\_infrastructures/PDFs/Questionnaire\\_e-Infra\\_Research\\_Data\\_Survey\\_Aug15.pdf](https://e-infrastructures.at/fileadmin/user_upload/p_e_infrastructures/PDFs/Questionnaire_e-Infra_Research_Data_Survey_Aug15.pdf) (Accessed 2017 March 2)
- Erway, R. 2013. Starting the Conversation: University-wide Research Data Management Policy. Dublin, Ohio: OCLC Research. Available at: <http://www.oclc.org/content/dam/research/publications/library/2013/2013-08.pdf>. (Accessed 2017 February 20)
- Eynden, V., Corti, L., Woollard, M., Bishop, L. and Horton, L. 2010. Managing and sharing data: best practices for researchers. Essex: UK Data Archive. Available at: <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf> (Accessed 2017 January 6)
- Federer, L. 2016. Research data management in the age of big data: Roles and opportunities for librarians. *Information Services and Use* 36 35–43 35 DOI 10.3233/ISU-160797.
- Fitzgerald, A. and Pappalardo, K. 2007. Building the infrastructure for data access and reuse in collaborative research: An Analysis of the Legal Context. Available at: <http://eprints.qut.edu.au/8865/1/8865.pdf> (Accessed 2017 February 11)
- Flores, J.R., Brodeur, J.J., Daniels, M.G., Nicholls, N, and Turnator, E. 2010. Libraries and the Research Data Management Landscape. Available at: <https://www.clir.org/pubs/reports/pub167/RDM.pdf> (Accessed 2017 February 24)
- Gabridge, T. 2009. The last mile: Liaison roles in curating science and engineering research data. *Research Library Issues*, 265, 15–21. Available at: <http://www.arl.org/bm~doc/rli-265-gabridge.pdf> (Accessed 2017 February 23)
- Gold, A.K. 2007. Cyber infrastructure, data, and libraries, Part 1: A cyber infrastructure primer for librarians. *D-Lib Magazine*. Volume 13 Number 9/10. Available at: [http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1015&context=lib\\_dean](http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1015&context=lib_dean) (Accessed 2016 October 23)
- Grace, S., Whyte, A. and Rans, J. 2015. ‘Building UEL’s Research Data Repository Capabilities with EPrints’. *DCC RDM Services case studies*. Edinburgh: Digital Curation Centre. Available at: <http://www.dcc.ac.uk/resources/case-studies> (Accessed 2016 December 3)
- Gray, D.E. 2014. *Doing research in the real world*. London: Sage Publications.

Greenbaum, D. and Gerstein, M. 2003. A universal legal framework as a prerequisite for database interoperability. *Nature Biotechnology* (21) 979.

Habert, B. and Huc, C. 2010. Building together digital archives for research in social sciences and humanities. *Social science information*, 49(3), pp.415-443. Available at: <http://journals.sagepub.com/doi/abs/10.1177/0539018410371570> (Accessed 2016 December 3)

Hall, J.A. 2007. *Accounting Information Systems*. 8th Edition. South-Western Cengage Learning

Hammond, M and Wellington, J. 2013. *Research methods: key concepts*. Oxon: Routledge.

Heidorn, P. B. 2011. The Emerging Role of Libraries in Data Curation and E-science. *Journal of Library Administration* 51 (7–8): 662–672. Available at: <http://www.tandfonline.com/doi/abs/10.1080/01930826.2011.601269#.VbpjALcwdaQ> (Accessed 2017 February 6)

Henderson, M.E., and Knott, T.L. 2015. Starting a Research Data Management Program Based in a University Library. *Med. Ref. Serv. Q.* 34, 47–59.

Henshaw, C., M. Savage-Jones, M. and Thompson, D. 2012. A Digital Library Feasibility Study. *LIBER Quarterly*. 20(1), pp.53–65. Available at: <https://www.liberquarterly.eu/articles/10.18352/lq.7975/> (Accessed 2017 January 29)

Henty, M. (2008) *Dreaming of data: The library's role in supporting e-research and data management*.

Henty, M. and Kingsley, D. 2007. Readiness and responsibility for managing research data: institutional perspectives. Available at: <https://openresearch-repository.anu.edu.au/handle/1885/46754> (Accessed 2017 January 23)

Hey, T. and Trefethen, A. 2003. The Data Deluge: An e-Science Perspective. *Grid Computing - Making the Global Infrastructure a Reality*, 809–824.

Higman, R. and Pinfield, S. 2015. Research data management and openness: the role of data sharing in developing institutional policies and practices. *Program: electronic library and information systems*, 49(4), pp.364-381. doi: 10.1108/PROG-01-2015-0005

Hitchcock, S. 2007. Digital preservation service provider models for institutional repositories. *D-Lib Magazine*. 13(S/6). Available at: [www.dlib.org/dlib/may07/hitchcockIOshitchcock.html](http://www.dlib.org/dlib/may07/hitchcockIOshitchcock.html) (Accessed 2017 January 13)

- Hofstrand, D. and Holf-Clause, M. 2009. What is feasibility study? Accessed at: <http://www.exension.iastate.edu/agdm> (Accessed 2017 January 23)
- Hole, B., Lin, L., McCann, P. and Wheatley, P. 2010. LIFE3: A predictive costing tool for digital collections. In iPres. 7th International Conference on Preservation of Digital Objects
- Johnson, R.B. and Onwuegbuzie, A.J. 2004. A Research Paradigm Whose Time Has Come. *Educational Researcher*, (33)7: 14-26. [Online] Available at: <http://edr.sagepub.com/content/33/7/14.full.pdf+html> (Accessed 2017 Mar 16)
- Johnston, L. R. 2014. Developing a data curation service: Step 1: Work with what you've got. *Bulletin of the Association for Information Science and Technology*, 40(4), 45–47. Available at: <http://onlinelibrary.wiley.com/journal/10.1002> (Accessed 2017 February 27)
- Jones, S. 2012. 10 recommendations to get started with research data management. Available from: <http://www.dcc.ac.uk/blog/10-recommendations-get-started-research-data-management> (Accessed 2017 January 29)
- Jones, S., Pryor G., and Whyte, A. 2013. How to Develop Research Data Management Services—A Guide for HEIs. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available from: <http://www.dcc.ac.uk/resources/how-guides> (Accessed 2016 October 6)
- Kahn, M., Higgs, R., Davidson, J., and Jones, S. 2014. Research Data Management in South Africa: How We Shape Up. *Aust. Acad. Res. Libr.* 45, 296–308.
- Keil, D.E. (2014). Research Data Needs from Academic Libraries: The Perspective of a Faculty Researcher. *Journal of Library Administration*, 54(3), pp.233-240. Available at: <http://dx.doi.org/10.1080/01930826.2014.915168> (Accessed 2017 February 13)
- Kennan, M.A., and Markauskaite, L. 2015. Research data management practices: A snapshot in time. *Int. J. Digit. Curation* 10, 69–95. Available at: <http://www.ijdc.net/index.php/ijdc/article/view/329> (Accessed 2017 January 20)
- Knight, G. (2015). Research data management at LSHTM: web survey report. London School of Hygiene and Tropical Medicine: London. Available at: <http://blogs.lshtm.ac.uk/rdmss/files/2013/04/LSHTM-RDM-Web-Survey-Report.pdf> (Accessed 27 February, 2017).
- Koltay, T. 2016. Are you ready?: Tasks and roles for academic libraries in supporting Research 2.0. *New Library World*, 117, 94-104. doi:10.1108/NLW-09-2015-0062
- Koopman M.M and De Jager K. 2016. Archiving South African digital research data: How ready are we? *South African Journal of Science* Volume 112, Number 7/8. <http://www.sajs.co.za> (Accessed 2017 January 3)

Kothari, M. 2004. Research Methodology: Methods and Techniques. 2nd ed. New Delhi:New Age international Ltd

Kumar, R. 2011. Research methodology: A step-by-step guide for beginners. 2nd ed. New Delhi: Sage.

Kurtz, D.T. and Boone, M.T.C. 2001. Business research methods. New Delhi: EssEss.  
Lee, Y., Chung, D. and Kim, Y. 2007. Public R&D Risk and Risk Management Policy. World Academy of Science, Engineering and Technology 34: 205-208

Lewis M.J. 2010. Libraries and the management of research data. In: McKnight S, editor. Envisioning future academic library services. London: Facet. Available at: [http://eprints.whiterose.ac.uk/11171/1/LEWIS\\_Chapter\\_v10.pdf](http://eprints.whiterose.ac.uk/11171/1/LEWIS_Chapter_v10.pdf). (Accessed 2016 October 5)

Linde, P., Wessels, B.A., Sveinsdottir, T. and Noorman, M. (2015). How Can Libraries and Other Academic Institutions Engage in Making Data Open?. Available at: <https://www.diva-portal.org/smash/get/diva2:834214/FULLTEXT01.pdf> (Accessed 2017 March 10)

Lindstrom, J. and Shonrock D. D. 2006. Faculty-Librarian Collaboration to Achieve Integration of Information Literacy. “*Reference & User Services Quarterly*” Vol. 46, No. 1 (Fall 2006) , pp. 18-23 [Online]. Available at: <http://www.jstor.org/stable/20864595> . (Accessed 2016 August 28)

Lotter, Lucia. 2014. “Reflections on the RDM Position in South Africa.” Paper presented at the LIASA Research Data Management Workshop, Cape Town, South Africa. Available from [http://www.dcc.ac.uk/webfm\\_send/1631](http://www.dcc.ac.uk/webfm_send/1631) (Accessed 2016 December 13)

Lyon, L. 2012. The informatics transform: Re-engineering libraries for the data decade. *Int J Digital Curation*. 7:126–138. Available from: <http://ijdc.net/index.php/ijdc/article/view/210>. (Accessed 2016 October 5)

Macanda, M., Rammutloa, M., and Bezuidenhout, R. 2014. Research data management at UNISA. Available at: <http://uir.unisa.ac.za/handle/10500/13907> (Accessed 2017 February 13)

MacColl, J. 2010. Library roles in university research assessment. *LIBER Quarterly*, 20 (2), pp. 152–168. Available at: <http://research-repository.st-andrews.ac.uk/bitstream/10023/1677/1/MacColl2010LIBERQuarterly20LibraryRoles.pdf>(Accessed 2017 February 27)

Marakas, G.M and O’Brien, J.A. 2014. Introduction to information systems. 6<sup>th</sup> Edition. New York: McGraw-Hill Education.

Matingwina, T. 2015. Health information dissemination among undergraduate students in Zimbabwe with particular reference to the National University of Science and Technology: a study in developing an integrated framework for health information dissemination. Cape

Town: University of Cape Town. Available at: [https://open.uct.ac.za/bitstream/handle/11427/15574/thesis\\_hum\\_2015\\_matingwina\\_thomas.pdf?sequence=1](https://open.uct.ac.za/bitstream/handle/11427/15574/thesis_hum_2015_matingwina_thomas.pdf?sequence=1) (Accessed 2017 January 4)

Mclure, M., Level, A. V., Cranston, C. L., Oehlerts, B., and Culbertson, M. 2014. Data curation: A study of researcher practices and needs. *Portal: Libraries and the Academy*, 14(2), 139–164.

Muir, B.J. 2004. *The Exeter anthology of Old English poetry: an edition of Exeter Dean and Chapter MS 3501. Rev. 2nd [CD-ROM] Edition.* Exeter: Exeter University Press.

Mugo, F.W. 2010. Sampling in research. Available at: <http://www.answers.com/topic/sampling.html> (Accessed 2017 March 15).

National Science Board. 2005. Long-lived digital data collections: enabling research and education in the 21<sup>st</sup> century. Available at: <https://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf> (Accessed 2017 January 20)

Naum, A. 2014. Research data storage and management: library staff participation in showcasing research data at the University of Adelaide. *Aust. Libr. J.* 63, 35–44.

Ndhlovu, P. 2016. The state of preparedness for digital curation and preservation: a case study of the National University of Science and Technology Library. Bulawayo: National University of Science and Technology

Neale, P., Thapa., S. And Boyce, C. 2006. Preparing a case study: a guide for designing and conducting a case study for evaluation input. Available at: [http://www2.pathfinder.org/site/DocServer/m\\_e\\_tool\\_series\\_case\\_study.pdf](http://www2.pathfinder.org/site/DocServer/m_e_tool_series_case_study.pdf) (Accessed 2017 March 17).

Nera, G. 2015. *The management of research records at Africa University, Zimbabwe.* Bulawayo: National University of Science and Technology.

Newton, M. P., Miller, C. C., and Bracke, M. S. 2011. Librarian roles in institutional repository data set collecting: Outcomes of a research library task force. *Collection Management*, 36(1), 53–67. doi: 10.1080/01462679.2011.530546

Nielsen, H.J and Hjørland, B. 2014. Curating research data: The potential roles of libraries and information professionals. *Journal of Documentation*, 70(2): 221–240. Available at: <http://www.emeraldinsight.com/doi/full/10.1108/JD-03-2013-0034> (Accessed 2017 January 20)

Nugroho, R.P., Zuiderwijk, A., Janssen, M. and de Jong, M. 2015. A comparison of national open data policies: lessons learned. *Transforming Government: People, Process and Policy*, 9(3), pp.286-308. Available at: <http://dx.doi.org/10.1108/TG-03-2014-0008>

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

Olive, P. 2012. The SAGE dictionary of social research methods: purposive sampling. Available at: <http://www.srmo.sagepub.com/view/the-sage-dictionary-of-social-researchmethods/n162.xml> (Accessed 2017 March 15)

Oxford University. 2014. Data. UK: Oxford University Press. Available at: [http://www.oxforddictionaries.com/us/definition/american\\_english/data](http://www.oxforddictionaries.com/us/definition/american_english/data) (Accessed 2017 January 6)

Palumbo, L.B., Jantz, R., Lin, Y.H., Morgan, A., Wang, M., White, K., Womack, R., Zhang, Y. and Zhu, Y. 2015. Preparing to Accept Research Data: Creating Guidelines for Librarians. *Journal of eScience Librarianship*, 4(2), p.1. Available at: <http://escholarship.umassmed.edu/jeslib/vol4/iss2/1/> (Accessed 2017 January 29)

Parahoo, K. 1997. Nursing research: principles, process and issues. Basingstoke: Macmillan.

Parsons, T. 2013. Creating a research data management service. *International Journal of Digital Curation*, 8(2), pp.146-156. Available at: <http://dx.doi.org/10.2218/ijdc.v8i2.279> (Accessed 2016 December 6)

Patel, D. 2016. "Research data management: a conceptual framework ", *Library Review*, Vol. 65 Iss 4/5 pp. 226 - 241 Available at: <http://dx.doi.org/10.1108/LR-01-2016-0001> (Accessed 2017 January 25)

Pinfield, S., Cox, A.M., and Smith, J. 2014. Research Data Management and Libraries: Relationships, Activities, Drivers and Influences. *PLOS ONE* 9, e114734. Available at: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114734> (Accessed 2017 January 6)

Potter, W.G. C. Cook, M. Kyriallidou. 2011. ARL profiles: Research libraries. Association of Research Libraries, Washington, D.C. Available at: <http://www.arl.org/storage/documents/publications/arl-profiles-report-2010.pdf> (Accessed 2017 February 23)

Pryor G. 2012. Managing research data. London: Facet.

Pryor, G., Jones, S. and Whyte, A. (Eds.). 2014. Delivering Research Data Management Services. London: Facet Publishing.

Rajic, C. 2016. Advanced computer architecture : a practical approach, parallel algorithms, parallel programming, super computers. New Delhi : S. Chand and Company.

Ray, J.M. 2014. Research Data Management: Practical Strategies for Information Professionals (West Lafayette, Indiana: Purdue University Press).

Research Data Management Forum. 2008. RDMF2: Core Skills Diagram. Available at:

<http://data-forum.blogspot.com/2008/12/rdmf2-core-skills-diagram.html> (Accessed 2017 February 10)

Remenyi, D., Williams, B., Money, A. and Swartz, E. 2003. *Doing research in business and management: An introduction to process and method*, London: Sage Publications.

Rice, R., Ekmekcioglu, Ç., Haywood, J., Jones, S., Lewis, S., Macdonald, S., and Weir, T. 2013. Implementing the Research Data Management Policy: University of Edinburgh Roadmap. *International Journal of Digital Curation* 8, 194–04. Available at: <http://www.ijdc.net/index.php/ijdc/article/view/8.2.194/327> (Accessed 2016 December 20)

Richardson, J., Nolan-Brown, T., Loria, P., and Bradbury, S. 2012. Library research support in Queensland: a survey. *Aust. Acad. Res. Libr.* 43, 258–277. Available at: <http://dx.doi.org/10.1080/00048623.2012.10722287> (Accessed 2017 February 23)

Saunders, M., Lewis, P. and Thornhill, A. 2009. *Research methods for business students*. 5th ed. Harlow: Pearson Education Limited

Schopf, H., Violaine, P and Rebouillat, V. 2016. *Research Data in Current Research Information Systems*. Available at: [https://halshs.archives-ouvertes.fr/sic\\_01331537/document](https://halshs.archives-ouvertes.fr/sic_01331537/document) (Accessed 2017 January 23)

Schumacher, J. and VandeCreek, D. 2015. Intellectual capital at risk: Data management practices and data loss by faculty members at five American universities. *International Journal of Digital Curation*, 10(2), pp.96-109. Available at: <http://dx.doi.org/10.2218/ijdc.v10i2.321> (Accessed 2016 December 3)

Science Europe. 2016. *Research Data Management*. Available at: [http://www.scienceeurope.org/wp-content/uploads/2016/05/SE-KE\\_Briefing\\_Paper\\_Funding\\_RDM.pdf](http://www.scienceeurope.org/wp-content/uploads/2016/05/SE-KE_Briefing_Paper_Funding_RDM.pdf) (Accessed 2017 January 23)

Steeleworthy, M. (2014). Research data management and the Canadian academic library: An organizational consideration of data management and data stewardship. *Partnership: The Canadian Journal of Library and Information Practice and Research*, 9(1), p.1. Available at: <http://search.proquest.com/openview/900957496adc3f619059b0fc664950d7/1?pq-origsite=gscholar&cbl=38957> (Accessed 2017 January 20)

Steinhart, G., Saylor, J., Albert, P., Alpi, K., Baxter, P. and Brown, E. 2010. *Digital research data curation: Overview of issues, current activities, and opportunities for the Cornell University Library*. A report of the Cornell University Library Data Working Group. Available at: [http://ecommons.library.cornell.edu/bitstream/1813/10903/1/DaWG\\_WP\\_final.pdf](http://ecommons.library.cornell.edu/bitstream/1813/10903/1/DaWG_WP_final.pdf) (Accessed 2017 February 23)

Strasser, CA. 2014. *Data Management for Libraries: A LITA Guide*, ALA TechSource, New York, US. Available from: Proquestebary. (Accessed 15 December 2016).

- Stuart, D. 2014. Libraries could play key role in managing research data. Available at: <https://www.researchinformation.info/feature/libraries-could-play-key-role-managing-research-data> (Accessed 2017 February 03)
- Surkis, A. 2015. Research data management. *Journal of American Librarianship*. 103(3). DOI: <http://dx.doi.org/10.3163/1536-5050.103.3.011> (Accessed 2017 January 6)
- Soehner, C., Steeves, C., and Ward, J. 2010. E-Science and Data Support Services: A Study of ARL Member Institutions. *Assoc. Res. Libr.*
- Shen, Y., and Varvel, V.E.J. 2013. Developing Data Management Services at the Johns Hopkins University. *Journal of Academic Librarianship* 39, no. 6 Pages Doi: 552–557. doi:10.1016/j.acalib.2013.06.002
- Sompel, H.V.D., Payette, S., Erickson, J., Lagoze, C. and Warner, S. 2004. Rethinking scholarly communication: building the system that scholars deserve. *D-Lib Magazine*; 2004 [10] 9. Available at: <http://dlib.org/dlib/september04/vandesompel/09vandesompel.html> (Accessed 2017 February 10)
- St Andrews University. 2014. Costing Research Data Management. Available at: <https://www.st-andrews.ac.uk/media/research-data-management/documents/generic/How-to%20Guide%20-%20Costing%20RDM.pdf> (Accessed 2017 January 23)
- Stanford University Libraries. 2016. Share and preserve research data. Available at: <https://library.stanford.edu/research/data-management-services/share-and-preserve-research-data> (Accessed 2016 December 20)
- Swan, A. and Brown, S. 2008. Skills, role and career structure of data scientists and curators: an assessment of current practice and future needs. JISC, Key Perspectives. Available at: <http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareersfinalreport.pdf> (Accessed 2017 January 10)
- Sykes, J. 2009. Managing the UK's Research Data: Towards a UK Research Data Service. *New Review of Information Networking*. 14(1), 21-36. Available at: <http://www.tandfonline.com/doi/full/10.1080/13614570902953556> (Accessed 2016 November 18)
- Tenopir, C., Sandusky, R.J., Allard, S., and Birch, B. 2013. Academic librarians and research data services: preparation and attitudes. *IFLA J.* 39, 70–78.
- Tenopir, C., Sandusky, R.J., Allard, S., and Birch, B. 2014. Research data management services in academic research libraries and perceptions of librarians. *Libr. Inf. Sci. Res.* 36, 84–90. Available at: <http://www.sciencedirect.com/science/article/pii/S0740818814000255> (Accessed 2017 January 16)
- Tenopir, C., Talja, S., Horstmann, W., Late, E., Hughes, D., Pollock, D., Schmidt, B., Baird, L., Sandusky, R., and Allard, S. (2017). Research Data Services in European Academic



Research Libraries. *Liber Quarterly*. 27. Available at: <https://www.liberquarterly.eu/articles/10.18352/lq.10180> (Accessed 2017 February 23)

Thurmond, A. V. 2001. The point of triangulation. *Journal of Nursing Scholarship*.33 (3): 253-258.

Tickle-Degnen, L. 2013. Nuts and Bolts of Conducting Feasibility Studies. *Am J Occup Ther*67, 171–176.

Toups, M. and Hughes, M. 2013. When data curation isn't: A redefinition for liberal arts universities. *Journal of Library Administration*, 53(4), 223–233. doi: 10.1080/01930826.2013.865386

UK Data Archive. 2015. Create and manage data: Research Data Lifecycle. Available at: <http://www.data-archive.ac.uk/create-manage/life-cycle> (Accessed 2016 October 5)

UK Data Archive. 2011. Managing and sharing data. Available at: <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf> (Accessed 2016 December 13)

UK Data Service. (N.D). Data management costing toolkit and checklist. Available at: <http://www.data-archive.ac.uk/media/247429/costingtool.pdf> (Accessed 2017 January 25)

University of Cape Town. 2014. Research data management policy. Available at: <http://www.digitalservices.lib.uct.ac.za/dls/rdm-policy> (Accessed 2017 February 3)

University College Dublin Library. 2015. Research Data Management: Why manage research data? Available at: [http://libguides.ucd.ie/data/why\\_manage](http://libguides.ucd.ie/data/why_manage) (Accessed 2016 December 6)

University of Pretoria. 2015. Research data management: home. Available at: <http://up-za.libguides.com/c.php?g=356288&p=2404132> (Accessed 2017 February 3)

University of Virginia Library 2016. Research Data Services, Steps in the data life cycle. Available at: <http://data.library.virginia.edu/data-management/lifecycle/>. (Accessed 2017 February 3)

University of Queensland. 2017. What is research data management?. Available at: <http://guides.library.uq.edu.au/research-data-management> (Accessed 2017 February 23)

University of Zimbabwe (UZ). 2016. University of Zimbabwe Five-year strategic plan: 2016-2020. Harare: University of Zimbabwe.

University of Zimbabwe. Vice Chancellor's. 2015 Graduation Speech. Available at: <http://www.uz.ac.zw/index.php/vice-chancellors-2015-graduation-speech> (Accessed 2016 October 23)

University of Zimbabwe Library. 2016. UZ Library 2017-2020 Strategic thrusts. Harare: University of Zimbabwe.

Van Deventer, M. and Pienaar, H. 2015. Research data management in a developing country: a personal journey. *International Journal of Digital Curation*, 10(2), pp.33-47 Available at: <http://www.ijdc.net/index.php/ijdc/article/view/380> (Accessed 2017 January 23)

Van Tuyl, S. and Michalek, G. 2015. Assessing Research Data Management Practices of Faculty at Carnegie Mellon University. *Journal of Librarianship and Scholarly Communication*, 3(3), eP1258. Available at: <http://dx.doi.org/10.7710/2162-3309.1258>

Van Wyk, J. and Van der Walt, I. 2014. Going Full Circle: Research Data Management at the University of Pretoria. Available at: <https://www.slideshare.net/johannvanwyk/goingfullcirclerdmupvanwykvanderwalt26nov2014> (Accessed 2017 February 23)

Wessels, B., Finn, R. L., Linde, P., Mazzetti, P., Nativi, S., Riley, S. 2014. Issues in the development of open access to research data. *Prometheus*, 32(1), 49–66.

Wilhem, E.E., Olster, E. and Shoulson, I. 2014. Approaches and Costs for Sharing Clinical Research Data. Available at: [http://www.kusp.ualberta.ca/en/KUSPNews/2014/February/~/\\_media/Knowledge%20Utilization%20Studies%20Program/Documents/Articles/Wilhelm\\_et\\_al\\_Data\\_Sharing\\_in\\_Health\\_Research\\_JAMA\\_2014\\_\(3\).pdf](http://www.kusp.ualberta.ca/en/KUSPNews/2014/February/~/_media/Knowledge%20Utilization%20Studies%20Program/Documents/Articles/Wilhelm_et_al_Data_Sharing_in_Health_Research_JAMA_2014_(3).pdf) (Accessed 2017 January 23)

Wilkinson, M. Amos, H. Flaherty, B. Hearne, S. Lynch, H. Lamond, H. Dewson, N. Kmiec, M. Nicolle, J. Talia, E. Ellio, G. 2016. Research data management feasibility report. Available at: <http://www.universitiesnz.ac.nz/files/CONZUL%20Project%20Report%20NRDR%20Feasibility%2020160617%20FINAL.pdf> (Accessed 2017 January 3)

Wilson, J.A., Fraser, M.A., Martinez-Uribe, L., Jeffrey, P., Patrick, M., Akram, A., and Mansoori, T. 2010. Developing Infrastructure for Research Data Management at the University of Oxford. Ariadne.

Witt, M., and Horstmann, W. 2016. International approaches to research data services in libraries. *IFLA Journal*. 42, 251–252. Available at: <http://journals.sagepub.com/doi/pdf/10.1177/0340035216678726> (Accessed 2017 January 20)

Whyte A, and Tedds J. 2011. Making the case for research data management. Edinburgh: Digital Curation Centre. Available at: [http://www.dcc.ac.uk/webfm\\_send/487](http://www.dcc.ac.uk/webfm_send/487). (Accessed 2017 January 6)

Woolfrey, L. 2010. Skills development to assist data usage for policymaking in Africa. Available at: <https://www.ifla.org/past-wlic/2010/86-woolfrey-en.pdf> (Accessed 2017 February 10)

Wright, S., Whitmire, A., Zilinski, L. and Minor, D. 2014. Collaboration and tension between institutions and units providing data management support. *Bulletin of the Association for Information Science & Technology*, 40(6), 18–21. Available at:

[http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223) (Accessed 2017 February 27)

Yin, R.K. 2014. Case study research: design and methods. Los Angeles: SAGE.

Yoon, A. and Schultz, T. 2016. Research Data Management Services in Academic Libraries in the US: A Content Analysis of Libraries' Websites. *College & Research Libraries*, pp.crl16-948. Available at: <http://crl.acrl.org/content/early/2016/11/16/crl16-948.full.pdf+html> (Accessed 2017 February 23)

## Appendix A

### Questionnaire for Researchers

My name is Nancy Nhendodzashe, a Master of Science student in Library and Information Science at the National University of Science and Technology (NUST). I am conducting a study aimed at exploring the feasibility of offering research data management services at the University of Zimbabwe (UZ). In the context of this study, research data refers to recorded factual material commonly retained by and accepted in the scientific community which is necessary to validate research findings. I am kindly asking for your participation in responding to the questions asked herein. Your participation in this study will help assess the viability of offering research data management services at the UZ.

The findings of the study will be used strictly for academic purposes and respondents will be kept anonymous. Kindly spare approximately 10 minutes of your time to respond to this questionnaire. Should you require any clarification or more information, do not hesitate to contact me on the following email addresses: [nhendodzashe@uzlib.uz.ac.zw](mailto:nhendodzashe@uzlib.uz.ac.zw) , [nancie1987@gmail.com](mailto:nancie1987@gmail.com).

#### SECTION A: PRELIMINARY INFORMATION

1. Select your faculty

Faculty	Tick
Agriculture	<input type="checkbox"/>
Arts	<input type="checkbox"/>
Commerce	<input type="checkbox"/>
Education	<input type="checkbox"/>
Engineering	<input type="checkbox"/>
Law	<input type="checkbox"/>
Science	<input type="checkbox"/>
Social Studies	<input type="checkbox"/>
Veterinary Sciences	<input type="checkbox"/>

2. How long have you been in the faculty?

- 0-5 Years   
6-10 Years   
11-15 years   
Over 15 Years

3. What is your highest qualification?

- Professor   
Dr   
Phd/DPhil Candidate   
Masters   
Other (Specify) \_\_\_\_\_

4. Gender  
Male  Female

## SECTION B. TECHNOLOGICAL REQUIREMENTS FOR RESEARCH DATA MANAGEMENT

5. *Research data management is the collection and organisation of data produced in particular investigations or research projects from its entry to the research cycle to the dissemination of research findings.*

How do you rate your knowledge of Research Data Management?

Excellent  Good  Average  Poor  Very Poor

6. Which types of data sets do you commonly generate?

### Data Types Generated

Tick all that apply

- Text documents (DOC, ODF, PDF, TXT, etc.)
- Structured text (HTML, JSON, TEX, XML etc.)
- Spreadsheets (XLS, ODS, CSV, SAS, Stata, SPSS, etc.)
- Databases (MS Access, MySQL, Oracle etc.)
- Geospatial Data (ESRI, Georeferenced TIFF, CAD data etc)
- Graphics/Images (JPEG, SVG, PNG, GIF, TIFF etc.)
- Audio (MP3, WAV, AIFF, OGG etc.)
- Video/Film (MPEG, AVI, WMV, MP4 etc.)
- Software applications Source code (CSS, JavaScript, Java etc.)
- Configuration data (INI, CONF etc.)
- Software Applications
- Other (Specify) \_\_\_\_\_

7. Where do you regularly store your research data?

### Data storage devices

Tick all that apply

- Locally on my work computer
- Locally on my private computer
- Centrally on departmental servers
- Centrally on a server of the university (institutional repository)
- On an external hard drive (including USB drive)
- On CDs/DVDs
- At an external data centre
- In a cloud service (eg Drop box, Google drive)
- Other (Specify) \_\_\_\_\_

8. Please estimate the total volume of your research data based on the storage space you require (your estimated average per year)

Small (< 50 GB)  Medium (50 – 100 GB)  Large (101 GB – 1 TB)   
 Very large (1 TB – 1PB)  I am not sure

9. Rate the following technology resources provided by the UZ

	Excellent 1	Good 2	Average 3	Poor 4	Not sure 5
Computing facilities (Hardware)					
Storage media					
Software for Research data management					
Data Backup tools					

**SECTION C. ECONOMIC NEEDS FOR RESEARCH DATA MANAGEMENT**

10. Do you have funders for your research?

Yes  No

If YES, who are they (e.g. My institution, research grants from donors etc) \_\_\_\_\_

11. Do you budget for research data management in your initial research funding proposals?

Yes  No

12. Do you incur any costs in managing your research data?

Yes  No

If YES, specify the costs  
 \_\_\_\_\_

14. Have you ever experienced research data loss?

Yes  No

If YES proceed to question 15 and 16. If NO Proceed to question 17

15. How did you lose the data?(Tick all that apply)

- Deleting files accidentally
- Computer viruses and damaging malware
- Mechanical damages of hard drive
- Power failures
- Theft of computer or hard drive
- Other (Specify) \_\_\_\_\_

16. What were the consequences of losing your research data?(Tick all that apply)

Lost research opportunity

- Duplication of effort
- Confidential data was exposed
- Missed deadlines
- Other (Specify) \_\_\_\_\_

17. Do you have any data backup mechanisms for your research data?

Yes  No

If YES, how do you backup your research data?

\_\_\_\_\_

### 18. Benefits of Research Data Management

*Please tick the appropriate number to indicate the level of your agreement or disagreement with the following statements on a scale of 1 to 5, where 1 = completely disagree, 2 = moderately disagree, 3 = neutral (neither disagree nor agree), 4 = moderately agree, and 5 = completely agree.*

	5	4	3	2	1
Data can be reused to answer new questions, opens up new interpretations and discoveries					
Research data management helps to protect and preserve valuable data					
Research data management improves the visibility of my research					
Research data management makes the comparison and co-analysis of data from multiple sources easier					
Research data Management provides an opportunity to check or repeat research and verify findings					
Research data management makes your articles and papers more useful and citable by others					

## SECTION D. LEGAL OBLIGATIONS FOR RESEARCH DATA MANAGEMENT

19. Who owns data from your research projects? Please choose all that apply:

Research team  Funding institution  Myself  The Department  UZ   
 Not Sure   
 Other (Specify) \_\_\_\_\_

20. Do you have any legal concerns (e.g Copyright, data protection, patents etc) regarding the use of your research data?

Yes  No

If YES, Explain the legal concerns \_\_\_\_\_

21. Do you have any policies that govern your research data management activities?

Yes  No  Don't know

If YES, provide the policies \_\_\_\_\_

**SECTION E. ORGANISATIONAL CAPABILITIES FOR RESEARCH DATA MANAGEMENT AND SCHEDULING OF RDM SERVICES**

22. Do you require Research Data Management services?

Yes  No  Not sure

If YES, proceed to Questions 23, 24 and 25. If NO/ Not Sure Proceed to question 26

23. How soon do you require Research Data Management services?

In the next one year  2-3 years  3-5 year  Not sure

Other (specify) \_\_\_\_\_

24. How can research data management services be implemented at the University of Zimbabwe? *Choose one response*

Start with a centralised discipline specific data repository  Start with a centralised multidisciplinary data repository

Start with a decentralised discipline specific data repository  Start with a decentralised multidisciplinary data repository

I don't know

Other \_\_\_\_\_

25. Rate your interest in the following research data management services

Research data management services	High	Medium	Low	Don't know
Support in preparing data management plans for grant proposals				
Training on research data management				
Storage of research data				
Preservation of research data				
Research data access control				

26. Have you ever been requested to submit Data Management Plans?

Yes  No

27. Do you share your research data with others?

Yes  Sometimes  No

If YES/Sometimes, Proceed to question 28. If NO/Sometimes, proceed to question 29



28. How are you currently sharing your research data? *(Tick all that apply)*

- Through an appropriate disciplinary data repository
- Privately by email upon request
- Through the project's or institution's web site
- Submit data, along with the article, to a journal publisher
- Other (Specify) \_\_\_\_\_

29. What are your reasons for not sharing your research data? *Please choose all that apply*

- Privacy violation  Danger of misuse
- Increased effort of time and/or cost  Risk of misinterpretation and/or falsification of data
- Potentially undesired commercial use  Missing data standards
- Increased competition in the “publish or perish” game  Other legal restrictions (e.g. copyright, patent law, trademark protection, use protection, etc.)
- Other (Specify) \_\_\_\_\_

30. Do you or your research unit re-use your research data? *Choose one response*

Always	<input type="checkbox"/>
Frequently	<input type="checkbox"/>
Occasionally	<input type="checkbox"/>
Rarely	<input type="checkbox"/>
Never	<input type="checkbox"/>
Don't know	<input type="checkbox"/>

31. Who do you think should be primarily responsible for research data management at the University of Zimbabwe? *Choose one response*

- Librarians  Researchers
- Research Unit  Information Technology
- Funders  I don't know
- Other (Specify) \_\_\_\_\_

32. Who do you think should be in the research data management team? *Choose all that apply*

- Librarians  Researchers
- Research Unit  Information Technology
- Funders  I don't know
- Other \_\_\_\_\_

**33. Please provide any other information that might be useful to this study-**

---



---



---

**THANK YOU FOR YOUR PARTICIPATION AND CONTRIBUTION**

## **Appendix B**

### **Interview Guide for Faculty Librarians**

My name is Nancy Nhendodzashe, a Master of Science student in Library and Information Science at the National Science and Technology. I am doing a study aimed at exploring the feasibility of offering research data management services at the University of Zimbabwe (UZ). Your participation in this study will help to assess the viability of offering research data management services at the UZ. The findings of the study will be strictly for academic purposes and respondents will be kept anonymous. Kindly spare your time to respond to the questions.

1. Do you understand the concept of research data management?
2. What kind of skills and competences do you think are needed for Research Data Management (RDM)?
3. Do you possess these skills (mentioned in question 2) needed for research data management?
4. If NO, what training needs to you require in order to execute RDM services?
5. What roles do you think librarians can take up in RDM?
6. Are you aware of the research data produced in your faculty? If YES, what kinds of research data are produced in your faculty?
7. Have you ever received any requests for research data from library patrons?
8. What do you think are the anticipated costs for setting up RDM services?
9. What do you foresee as some of the benefits of offering RDM services?
10. Who do you think are the main stakeholders in RDM?
11. Is there adequate support from top university management on library projects?
12. How do you think RDM services can be implemented at the UZ library?
13. Which RDM services do you think the library can currently provide?
14. What do you perceive as the legal issues surrounding the collection, storage, preservation and access to research data?
15. Do you have any other information that you think is useful to this study?

**Thank You**

## **Appendix C**

### **Interview Guide for University Librarian**

My name is Nancy Nhendodzashe, a Master of Science student in Library and Information Science at the National Science and Technology. I am doing a study aimed at exploring the feasibility of offering research data management services at the University of Zimbabwe (UZ). Your participation in this study will help to assess the viability of offering research data management services at the UZ. The findings of the study will be strictly for academic purposes and respondents will be kept anonymous. Kindly spare your time to respond to the questions.

1. I understand that you are planning to offer research data management services as a Library, does the UZ Library have the necessary ICT infrastructure for Research Data Management (RDM)? If NO, are there mechanisms to acquire additional ICT resources that support RDM?
2. Have you enacted enabling policies for RDM (e.g copyright, data access, storage and preservation policies) yet?
3. What do you perceive as some of the benefits of offering RDM services to the UZ community?
4. What do you foresee as some of the costs of managing research data?
5. If there are any costs, are there mechanisms in place to cater for the costs?
6. Do you think the current set of librarians have the skills and knowledge on RDM?
7. If NO, what can be done to enhance their skills and knowledge to take up RDM roles?
8. Which RDM roles do you think the library can begin with (e.g consultative/informational roles on RDM or technical roles etc)?
9. How can RDM services be implemented at the UZ Library (e.g start with a discipline specific data repository or a multidisciplinary data repository etc)?
10. When do you intend to roll out RDM services?
11. Other than the Library, who else do you think should be involved in RDM activities?
12. Do you have any other information that you think is useful to this study?

**Thank You**

## **Appendix D**

### **Interview Guide for the Library ICT Manager**

My name is Nancy Nhendodzashe, a Master of Science student in Library and Information Science at the National Science and Technology. I am doing a study aimed at exploring the feasibility of offering research data management services at the University of Zimbabwe (UZ). Your participation in this study will help to assess the viability of offering research data management services at the UZ. The findings of the study will be strictly for academic purposes and respondents will be kept anonymous. Kindly spare your time to respond to the questions.

1. Do you understand the concept of research data management?
2. What are the technological requirements for setting up RDM services?
3. What kind of data types are supported by the current technological infrastructure?
4. To what extent is the current IT infrastructure adequate to cater for research data collection, storage, preservation and access?
5. What is the better option for implementing research data management services- using existing institutional repositories or licences for data repositories?
6. What are the risks associated with digital data?
7. What are the mechanisms that can be employed to mitigate the risks?
8. Do you possess the technical skills that are needed for setting up RDM services?
9. How long does it take to set up the RDM platform?
10. Which RDM services do you think the library can currently provide?
11. How can RDM services be implemented at the UZ library? (e.g. Centralised or decentralised discipline specific or multidisciplinary data repository)
12. Do you have any other information that you think is useful to this study?

**Thank You**

Appendix E

Research Authorisation Letter



NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

P.O. Box AC-035 - Anson - Bulawayo, Zimbabwe  
Off. Gweru Road/Cecil Ave. Jct

Telephone: 253-9-262842/283413  
Fax: 253-9-266803

DEPARTMENT OF LIBRARY AND INFORMATION SCIENCE

22 February 2017

The Librarian  
University of Zimbabwe Library  
P.O Box MP45  
Mt Pleasant  
HARARE

Dear Sir/Madam

Re: **MS. NANCY NHENDODZASHE (N01521888L): SUPPORTING LETTER**

Ms. Nancy Nhendodzashie (N01521888L) who is currently doing the MSc Degree in Library and Information Science has asked me to write her a letter supporting her request to do a research project at your organisation.

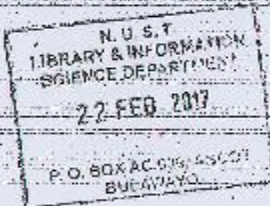
Ms. Nhendodzashie's research topic is "Exploring the feasibility of offering research data management services at the University of Zimbabwe Library."

It will be very much appreciated if you could assist by giving her permission to research.

If you are willing to assist, please indicate your willingness by signing, dating and stamping the second copy of this letter and returning it to our office.

Thank you.

  
Mr. L.R. Fuzire  
Chairman



SIGNED

DATE

23/02/2017

OFFICIAL STAMP