The feasibility of offering standardised data literacy services at selected private university libraries in Kenya

by

Stanislaus Litsalia Agava

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Information Science

in the

DEPARTMENT OF INFORMATION SCIENCE,

FACULTY OF ENGINEERING,

THE BUILT ENVIRONMENT AND

INFORMATION TECHNOLOGY

at the

UNIVERSITY OF PRETORIA

Supervisors:

Dr MJ van Deventer and Prof. TJD Bothma

FEBRUARY 2024

DECLARATION

I, Stanislaus Litsalia Agava (16383053), affirm that this is my sole original creation. When referencing external works, I have duly acknowledged them. This thesis has (or parts thereof have) not been presented by me to any other higher education institution for a degree or any academic distinction.

Name: Stanislaus Litsalia Agava

Signature:

ACKNOWLEDGEMENTS

I would like to extend my heartfelt gratitude and appreciation to the following individuals and entities:

- I am immensely thankful to the Almighty God for the gift of life and good health that sustained me throughout the entire journey of writing my thesis.

- To my dedicated mentors and supervisors, Dr M.J. van Deventer and Prof. T.J.D. Bothma. I owe you a debt of gratitude for journeying with me, for your unwavering academic guidance and support since the inception of this process

- My deepest appreciation goes to my wife, who served as an unwavering pillar of strength during moments of stress and doubt. Your encouragement and understanding were invaluable.

- To my son, your serene presence created an ideal working environment, and your understanding of my busy schedule has been a source of motivation since 2018.

- My daughter, with your engaging father-daughter games, you provided a sanctuary of relaxation and peace during the challenging times.

- To my brothers and sisters, your consistent encouragement bolstered my spirits throughout this journey.

- I extend my sincere thanks to Tangaza University College, my first employer, for their partial funding support for my studies.

- My gratitude also goes to my current employers, who not only provided the necessary workspace but also generously funded various aspects of my study.

Your collective support and understanding have been instrumental in making this thesis a reality, and I am deeply appreciative of your contributions to this academic endeavour. May God bless ALL of you immensely.

DEDICATION

In loving memory of my mother Agnes, your spirit forever inspires me.

ABSTRACT

In the era of rapid data generation and digital transformation, the ability to effectively navigate, interpret, and utilize data is fundamental for researchers in Kenyan universities. This study delves into the critical areas of data literacy to address the challenges researchers encounter while managing research data. Defined as the capacity to read, understand, create, and communicate data as information, data literacy, is necessary for researchers in diverse fields. It provides researchers with skills to access, analyse, interpret, evaluate, and employ data to enhance problem-solving and informed decision-making. Despite the ubiquity of data, researchers in Kenyan universities grapple with data management challenges. This study's primary objective was to assess the feasibility of implementing standardized data literacy services within private university libraries in Kenya. The study was aimed at achieving the following four key objectives: evaluating the data literacy needs of faculty members and postgraduate students in selected private universities in Kenya; assessing the organizational infrastructure within these private university libraries to determine their readiness for delivering data literacy services; investigating the technical infrastructure readiness of these libraries for implementing data literacy training and developing a comprehensive framework that encapsulates the requisites for successful data literacy initiatives in Kenyan universities. The study employed both quantitative and qualitative approaches, involving surveys, interviews, and document analysis. For data analysis the study used statistical tools and thematic analysis to extract meaningful insights from the collected data. Findings reveal a pronounced gap between existing library services and researchers' data literacy needs. Majority of the participants expressed dissatisfaction with the current level of data-related support, highlighting the necessity for education and training on data management, metadata creation, and research data services. Furthermore, challenges associated with data management, including data publishing and open data practices, were evident. The study identified key stakeholders in the implementation of data literacy programs in universities, including libraries, ICT departments, research directories, government agencies and regulators as well as research funders. A key role to be played by the library was well pronounced emphasizing the integral role of academic libraries could play in promoting data literacy. Libraries were pointed out having shown to possess the necessary capacity, infrastructure, and expertise to facilitate data literacy programs effectively. This study aligns with the Radical Change theory, Stakeholder theory, Intellectual Capital model, and the Bielefeld University RDM training model to provide a comprehensive framework for data literacy initiatives. It underscores the urgency for academic libraries to adapt to evolving research needs and embrace their role in fostering data literacy. The study lays the foundation for enhancing data literacy among researchers in Kenyan universities and offers practical recommendations to bridge existing gaps, emphasizing collaboration among stakeholders and the adoption of policies and infrastructure supporting open science and research data. The developed data literacy framework holds the potential of improving research quality, innovation, and societal impact within Kenya's academic community.

Keywords: Data Literacy, Research Data Management, Academic Libraries, Data literacy framework, Feasibility Study, Standardized Services, Kenya Universities

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS

ACRL	Association of College and Research Libraries
ALA	American Library Association
ASA	American Statistical Association
CPD	Continuing Professional Development
CSTB	Computer Science and Telecommunications Board
CUE	Commission for University Education
DIRISA	Data-Intensive Research Infrastructure for South Africa
DMP	Data Management Plan
DMPs	Data Management Plans
EBIT	Engineering and Built Environment
EC	European Commission
HEI	Higher Education Institutions
IASSIST	International Association for Social Science Information Services and Technology
ICT	Information Communication Technology
ISI	International Statistical Institute
MIT	Massachusetts Institute of Technology
NACOSTI	National Commission for Science, Technology & Innovation
NeDICC	Network of Data and Information and Curation Communities in South Africa
NHGIS	National Historic Geographic Information System
NRC	National Research Council
OECD	Organisation for Economic Co-operation and Development
RDA	Research Data Alliance
RDAP	Research Data Access and Preservation
RDM	Research Data Management
RDS	Research Data Services
SADA	South Africa Data Archive
SANParks	South African National Parks
SPSS	Statistical Package for the Social Sciences
USA	United States of America
UCT	University of Cape Town

UG	University of Uganda
UK	United Kingdom
UL	University Librarian
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNISA	University of South Africa
UP	University of Pretoria
USFSP	University Library of South Florida St. Petersburg

CHAPTER 1

1.   INTRODUCTION

The first chapter of this study highlights the fundamental issue of data literacy's importance in the twenty-first century, notably in research and academia. It emphasizes the necessity of data literacy and the value of data in decision-making. The chapter explores several data literacy projects as well as libraries' increasing role in tackling this topic. It describes the research problem as the need for improved data literacy, identifies gaps in data literacy training and policies, and emphasizes libraries' role in closing these gaps. The study's direction is guided by the research objectives and questions. The chapter concludes with a summary of the research methods, ethical considerations, limits, and a brief description of the following chapters.

1.1   Background/Context

In the 21st century, possessing data literacy is imperative for research and for adapting to an economy driven by information (Gibson & Mourad, 2018; Koltay, 2017). The value of data has been likened to currency, given its economic value (Eggers, Hamill & Ali, 2013). The United Nations has stressed the importance of data in the decision-making process, referring to it as the essential foundation of decision-making and the primary resource for ensuring accountability (Morales et al., 2014). Without high-quality data providing the right information at the right time, designing, monitoring, and evaluating effective policies becomes almost impossible.

The emergence of e-research or e-science has widened the need for training researchers and data handlers. E-research is the development and support of advanced information and computational technologies to enhance all phases of research processes (Luce, 2008). Access to high-bandwidth networks, massive data storage capacity, and advanced data analysis and visualization tools have significantly changed research and scholarship. These changes call for a new approach to handling data, placing more responsibilities in the hands of critical stakeholders in institutions of higher learning such as the libraries and librarians, who are playing a significant role in the lifecycle of research.

Koltay (2017b) notes that data literacy skills are necessary worldwide for students and researchers, given that research is one of the key pillars of academia. Researchers in a university community, mainly postgraduate students and faculty members, play a crucial role

in channelling out research in line with many universities' research agenda. They collect, process, analyse, and disseminate data, carrying out research data management activities. The academia is beginning to recognize the need to empower its community of researchers with more data-based skills to meet data management-related requirements and channel out reliable data.

The library plays a fundamental role in supporting teaching, learning, and research in higher learning institutions. Information literacy services offered by the library have raised information literacy levels among library users and for that reason, there is a growing debate on the library's role in data related services (Tammaro et al., 2016; Heidorn, 2011; Corrall, 2012) and implementing data literacy programs as an extension of existing services to researchers. It is evident that offering advanced data literacy training can enable academic libraries and librarians to enhance researchers' data management skills. (Dai, 2020; Perrier & Barnes, 2018).

It is upon academic libraries in the 21st-century to deploy a four-fold response to the data challenge: 1) hire specialized staff (data librarians) or provide data management and analysis training for librarians; 2) intensify the collection or compilation of data sources and providing access to such sources; 3) participate in the development of institutional data repositories to preserve and share original research data, and 4) incorporate data literacy in their instructional programs and services (ACRL, 2015). Without a guiding framework, the implementation of these initiatives is inefficient.

Although data literacy is a new area of study and practice, studies have shown the advent of data support services offered by various libraries. Some of the studies include Patel (2016), Liu and Ding (2016), and Chigwada, Chiparausha, and Kasiroori (2017). Findings reveal either poor policing or lack of policy in relation to data management. Other findings indicate rigid institutional adaptability coupled with ignorance about data literacy, particularly among top management.

1.2   What is data literacy

Data literacy is an emerging concept that has become relevant in contemporary times, especially in research. It refers to the ability of individuals to access, interpret, evaluate critically, manage, handle, and ethically use data (Palsdottir, 2021). There are varying definitions of data literacy, with different scholars using alternative terms such as data information literacy, research data literacy, and science data literacy. However, the essence of

data literacy lies in the ability of individuals to transform data into information and actionable knowledge, which requires specific skill sets and knowledge bases. Carlson, Fosmire, Miller and Nelson (2011) describe data literacy as the ability to understand what data mean, including how to read graphs and charts appropriately, draw correct conclusions from data, and recognise when data are being used in misleading or inappropriate ways. Similarly, Mandinach and Gummer (2013) define data literacy as the ability to understand and use data effectively to inform decisions. Wanner (2015) describes data literacy as a sought of competence that provides individuals with the skills to find and evaluate data, as well as cite and ethically use data. Wolff et al. (2016) argue that current definitions of data literacy do not account for changes in the nature of data sets, which are becoming larger and more complex. Therefore, a clear definition of data literacy is necessary for it to be effectively introduced in the society through education.

1.3   Data literacy initiatives

Why data literacy? It is presumed that effective use of data empowers people to make objective, evidence-based inferences and fundamental decisions affecting their lives, both as individuals and societies. As data continues to claim a rightful space in the context of other literacies, data literacy becomes not just an important academic exercise but also an important civic skill that needs to be developed. As an academic exercise, it makes students and researchers data literate. As a civic skill it is meant for lifelong learning (Nayek & Sen, 2015). In a bid to confirm the vital role of data and the need for data literacy, Kenneth Prewitt wrote in his introduction to the *Association of Public Data Users* publication on statistical literacy indicating that, "statistical literacy can contribute to more responsible data presentation and greater understanding of data, [and] inform public policy debates about statistical programs" (Prewitt, 1999:1).

Though data literacy is still a new area for exploration, some initiatives have emerged in recent years revealing efforts towards data literacy-related programmes or activities. At the University of California, Los Angeles, for instance, Elizabeth Stephenson and Patti Schifter Caravello co-taught a one-credit data literacy and statistical literacy sociology course. The course was designed in order to develop skills in critical thinking and information evaluation among undergraduate students (Stephenson & Caravello, 2007).

After surveying data management practices at the faculty of science, Qin and D'ignazio (2010a) developed a science data literacy course. While utilising a variety of pedagogical strategies to engage students, the course was broadly based around three modules; fundamentals of science

3

data and data management, managing data sets in aggregation and broader issues in science data management (ibid).

Carlson et al. (2011) conducted an extensive study involving students and staff in the science and engineering departments at Purdue and the University of Illinois, Urbana-Champaign. The study aimed to assess the data literacy needs of both faculty and students in a Geoinformatics course. Using ACRL's information literacy competency standards to filter the results, Carlson et al (2011) developed learning goals for data literacy instruction. Their proposed core competencies for data information literacy included; introduction to databases and data formats, discovery and acquisition of data, data management and organisation, data conversion and interoperability, quality assurance, metadata, data curation and reuse, cultures of practice, data preservation, data analysis, data visualisation and ethics, including citation of data (ibid).

On this front, data literacy is considered to be as important as reading and writing. Unfortunately, the world of research is facing a major skill gap. There are no unified standards for data literacy (Mandinach & Gummer, 2012), although many researchers have proposed core competencies and curriculum guidelines based on evidence-based practices. The concept is quickly emerging as a key priority in today's world economy where development is increasingly defined by data collection and the knowledge and skills required to use the data. We are in perpetual production of streams of data from all sectors of life, hence the need for skills to work with and interpret the data.

1.4   Problem statement

The need for data literacy skills is increasingly important as the 21st century progresses. This is due to the growing reliance on data-driven decision-making across various fields. In academia, researchers (faculty and postgraduate students) play a critical role in collecting, analysing, and disseminating data to advance knowledge and address pressing societal issues. However, there is a lack of adequate data literacy training for this community or a standardized data literacy framework, hence affecting the quality of research outputs and undermining the credibility of research (Adika & Kwanya, 2020). Libraries have traditionally played a vital role in supporting teaching, learning, and research in universities, but according to some studies there is a growing debate about their role in data curation and providing data literacy services (Xu, 2023; Koltay, 2019; Cox, Kennan, Lyon & Pinfield, 2017; Heidorn, 2011). The Association of College and Research Libraries (ACRL) has called for libraries to incorporate

data literacy into their instructional programs and services to help researchers develop advanced data management skills.

However, there are gaps in policy and practice regarding data management, and the implementation of data literacy programs, is often inefficient due to a lack of guiding frameworks. Additionally, studies have shown that some academic institutions have poor data governance, they lack policies in relation to data management, while others have either rigid institutional adaptability or are ignorant about data literacy, especially among top management. While the imperative for data literacy grows in academia, the apparent absence of a standardized data literacy framework hampers private university libraries in Kenya from effectively addressing the increasing demand for these skills. The misalignments observed in policy, human resourcing, and infrastructure development, further exacerbate the challenges faced by libraries in delivering quality data literacy services, highlighting the critical need for a comprehensive and evidence-based intervention framework.

1.5   Research objectives

It is in view of the prevailing context that this study embarked on identifying the most important gaps that academic institutional libraries (also known as academic libraries) must be able to address. The study assessed the feasibility of offering standardized data literacy services at selected private university libraries in Kenya. By evaluating the data literacy needs of faculty and postgraduate students and the organizational and technical infrastructure readiness of the libraries, the main objective of the study was to develop a framework to guide the implementation of a relevant data literacy program, to improve data literacy practices and enhance the quality of research outputs in Kenyan academic institutions.

The research used the following sub-objectives. The research set out to:
1. Critically review existing literature on data literacy, in order to establish a comprehensive framework encompassing essential services and components that would facilitate an essential discussion about the successful implementation of a data literacy initiative in selected private universities in Kenya
2. Asses the data literacy needs of faculty and postgraduate students in selected private universities in Kenya.
3. Asses the organizational infrastructure of selected private university libraries in Kenya to determine the feasibility of offering data literacy services.

4. Assess the technical infrastructure readiness in the selected private university libraries in Kenya to facilitate the implementation of data literacy training.

1.6 Research questions

The central research question that guided this study was: *What would a generally accepted, evidence-based, broad framework of data literacy interventions that includes policy, human resourcing and infrastructure development (to be used by university libraries in Kenya), look like?*

In order to respond to the main question, the following sub-questions were used to guide the course of the study.

1. How does the existing literature on data literacy inform the development of a comprehensive framework with essential services and components for the successful implementation of a data literacy initiative in selected private universities in Kenya?
2. What are the data literacy needs of faculty members and postgraduate students as researchers in selected private universities in Kenya?
3. What is the existing organisational infrastructure within the selected private university libraries in Kenya that supports data literacy initiatives for researchers?
4. To what extent are the selected private university libraries in Kenya technically equipped to implement data literacy programs?

Having established the research objectives and questions, in the following section, the researcher provided a concise overview of the research methodology employed in this study. A more detailed and comprehensive explanation of the research process can be found in Chapter 4.

1.7 Research methodology

The researcher provides a brief overview of the research process used in this study in this part, with a more detailed explanation available in Chapter 4.

1.7.1 Research paradigm, research approach, and research design

Research paradigm: The pragmatic paradigm was selected as the research paradigm for this study due to its fundamental principles and practical advantages. Pragmatism challenges the notion that a single scientific method can comprehensively address all aspects of social reality (Kivunja & Kuyini, 2017). Instead, it advocates for a pluralistic approach, one that seamlessly combines various research strategies to gain a more practical understanding of human

6

behaviours. At the core of pragmatism lies an emphasis on the significance of experience and the utilization of mixed research methods.

Pragmatism not only promotes methodological diversity but also encourages the integration of methods and the application of knowledge gained from one study to other situations, thus enhancing the transferability of research findings (Hesse-Biber & Johnson, 2015). This choice was grounded in a careful consideration of the study's context, the specific phenomenon under investigation, and the practical constraints, especially in smaller and time-constrained investigations. The integration of paradigms within the pragmatic framework ensures that research maintains high standards in terms of validity, relevance, reliability, and the generation of actionable outcomes.

Research approach: Research approaches are defined as the methodologies that outline the researcher's techniques and plans to achieve the study's objectives (Creswell & Creswell, 2018). The selection of a research approach is influenced by various factors, including the research problem, the specific concerns being addressed, the researcher's experience, and the study's population. For the purposing of achieving this study, the researcher chose the mixed methods approach, which combines quantitative and qualitative methodologies. The mixed methods approach, which incorporates elements from both quantitative and qualitative research, has gained popularity among researchers. It involves the gathering, analysis, and integration of quantitative and qualitative data in a single study to enhance the understanding of a research problem. This approach allowed for a more comprehensive and complementary use of data, acknowledging trends, generalizations, and participants' perspectives.

This choice was relevant because it was necessary to apply various procedures, to strengthen specific aspects of the study, and to allow for data triangulation/convergence. By combining quantitative and qualitative approaches, the researcher aim was to gain a broader and deeper understanding of the research problem, resulting in richer and more diverse data. This approach is well aligned with the pragmatist paradigm which was selected as the study's worldview.

Research design: Research designs are essential components that guide data collection, analysis, interpretation, and presentation methods in a study. They are chosen based on factors such as the study's intent, procedures, and strengths and challenges associated with each design. For this study, the researcher chose to use the convergence design, which is a specific form of triangulation that focuses on bringing together different types of data or evidence to reach a unified or cohesive understanding of a research issue. Convergence design is a well-known and

7

widely used research design when using a mixed research methods approach. It involves collecting both quantitative and qualitative data separately but on the same topic simultaneously. This design aims to leverage the strengths of both quantitative and qualitative methods while mitigating their individual weaknesses.

The application of the convergence mixed methods design involved collecting and analysing quantitative and qualitative data concurrently during the same stage of the study. This allowed for the comparison and contrast of results, making it suitable for researchers aiming to validate or expand quantitative findings with qualitative insights.

The chosen design facilitated a comprehensive understanding of the research problem by integrating multiple data sources and methods. It aligned with the pragmatist paradigm selected for the study, emphasizing the importance of using diverse approaches to gain insights into the research phenomenon.

1.7.2   The study population

The population of the study in this research consisted of private chartered universities located within the Nairobi metropolitan area, encompassing Nairobi and areas of proximity to the city from the surrounding counties. This region is known for its rapid growth, attracting private investors and universities. In total, there were 12 private universities in this area, but the study focused on those that offered PhD programs. The selected universities included the Catholic University of Eastern Africa, United States International University, African International University, Adventist University of Africa, and Daystar University. In summary, the decision to focus on specific private universities was a pragmatic one, driven by the limited number of public universities in the area, practical constraints. Some institutions also declined to participate. This approach allowed for a manageable and feasible study within the context of defined resources and time constraints.

The study's population included various groups of participants directly involved in research activities and data management within these universities. These groups comprised postgraduate students (specifically PhD students), full-time faculty members (ranging from lecturers to full professors), heads of university libraries, and research/reference librarians. These participants were chosen based on their active engagement in research and data-related activities, making them valuable sources of information for the study. Part-time faculty members were excluded due to their contractual differences and mobility between universities, which limited their contribution to university research requirements.

University librarians were considered eligible for inclusion in the study, given their roles in overseeing library services and influencing policy changes, such as data literacy initiatives. Full-time faculty members and PhD students were included because of their active involvement in research, including data collection, handling, and publication requirements. Research/reference librarians, who provide research-related services, including data management training, were also part of the study's population.

1.7.3 Sampling

The study employed two main sampling procedures: purposive sampling and full/total population sampling.

Purposive Sampling: This technique involved selectively identifying specific participants from the target population based on their knowledge, experience, availability, willingness to participate, and ability to provide essential information related to the research. It was primarily used to choose private universities offering PhD programs within the Nairobi Metropolitan area, faculty and doctoral students actively involved in research, and university librarians who play a crucial role in library programs and services, as well as research/reference librarians who provide data-related services to researchers. The researcher aimed to overcome potential bias by including multiple and diverse participants selected randomly.

Full/Total Population Sampling: This method, a type of purposive sampling, involved examining the entire population when specific characteristics or attributes were common across it. In this study, full/total population sampling was used only for selecting researchers, including full-time faculty members and PhD students in the six universities. The researcher relied on gatekeepers within the universities to distribute the questionnaire since the researcher was not granted direct access to participants' email addresses. The entire population of full-time faculty members and PhD students in the five universities received the questionnaire link.

Sample Size: The sample size for the study was determined using the SurveyMonkey sample size calculator, considering a confidence level of 80%, a margin of error of 5%, and a z-score of 1.28. The calculated sample sizes for each group were as follows:

- Fulltime Lecturers: 291

- Postgraduate Students (PhD): 323

- Research/Reference Librarians: 5

- University Librarians: 5

These sample sizes were used to ensure that the study's findings were statistically reliable and representative of the selected population groups.

1.7.4 Data collection

The study employed both quantitative and qualitative data collection tools within a mixed methods research design.

Questionnaires: In this study, questionnaires were used to collect quantitative data from researchers, including PhD students and faculty members (see Appendix I). The questionnaires included both closed-ended and open-ended questions, aiming to obtain factual information as well as opinions, attitudes, and perceptions. Questionnaires were chosen for their cost-effectiveness, standardized answers, and the anonymity they provide to respondents.

Interview schedules: The study employed semi-structured interviews to collect qualitative data from University Librarians and Research/Reference Librarians (see Appendices II and III). Semi-structured interviews offer flexibility while ensuring that relevant topics are covered. Open-ended questions allowed for probing and clarification, especially regarding complex data literacy concepts. Interviews were conducted individually at the participants' workplaces, with mutually agreed-upon schedules.

The two distinct sets of data, one quantitative and the other qualitative, allowed for a comprehensive exploration of the research questions and objectives. The combination of data collection tools aimed to collect persuasive and dependable findings, enriching the study's overall quality and reliability.

1.7.5 Data analysis

The study employed a mixed methods research approach and utilized a convergence design for data collection and analysis (refer to Section 1.7.1).

The study employed a mixed methods research approach and utilized a convergence design for data collection and analysis (refer to Section 1.7.1).

This method facilitated the synthesis of both quantitative and qualitative data analysis techniques, ensuring a well-rounded examination of the research questions and objectives.

Quantitative Data Analysis: Quantitative data, generated from closed-ended questions in the questionnaire, were imported into the IBM SPSS statistical package version 21. The process

involved data file creation and data validation. The data processing, interpretation, and analysis followed. The results were presented using charts and tables, complemented by explanatory text to convey findings, categorizations, and correlations among different data sets.

Qualitative Data Analysis: Qualitative data, stemming from both open-ended questions in the questionnaires and interview responses, were analysed using content analysis. Content analysis is a method for interpreting the content of text data by systematically classifying, coding, and identifying themes or patterns.

The study's data analysis process aimed to uncover underlying patterns, trends, and relationships in the data, as well as to create a comprehensive understanding from responses to research questions. The researcher acknowledged the complexities of analysing mixed methods data and employed a side-by-side technique, first analysing quantitative data and then qualitative data before merging the findings. This approach ensured a comprehensive exploration of the research objectives and a meaningful integration of data from different sources.

1.7.6 Validity and reliability

Validity and reliability are crucial aspects of research methodology that help ensure the credibility and accuracy of study findings.

Validity: Validity refers to the extent to which a research instrument accurately measures what it is intended to measure. In this study, the researcher employed multiple strategies to enhance validity. Open-ended questions in the data collection tools were generated from the study's objectives, ensuring relevance. The use of convergence, involving different data collection techniques (questionnaires and interviews), contributed to internal validity by confirming the findings. External validity, which assesses the applicability of findings beyond the study context, was addressed by selecting a representative sample to make it possible to extrapolate the results to other Kenyan private chartered universities.

Reliability: Reliability focuses on the consistency and repeatability of research results. To increase the reliability of the study, the researcher explicitly described each phase of the research process, including its purpose, design, and participants. This detailed documentation enhances the consistency of the study.

Both validity and reliability are essential to ensure that the research produces meaningful and dependable results, enhancing the overall quality and acceptance of the study

11

1.7.7 Ethical Clearance

Ethical clearance is a crucial aspect of research, ensuring the integrity of the study and the protection of participants' rights. The research was considered for ethical soundness by both the University of Pretoria (EBIT 170/2022, see Appendix IV) as well as the National Commission for Science, Technology and Innovation (NACOSTI) (see Appendix V). The researcher took several steps to adhere to ethical requirements:

Protection from Harm: The researcher ensured that participants did not face any physical or psychological harm during the study. This included considering potential harm to participants and taking steps to minimize it. Sensitive and potentially traumatizing questions were avoided in the data collection instruments.

Informed Consent: Participants were provided with detailed information about the study, including its purpose and what participation entailed. They were required to provide explicit, active, and signed consent, acknowledging their rights to access their information and the right to withdraw from the study at any point.

Anonymity and Confidentiality: Anonymity was maintained by disassociating names from responses during data coding and using aliases or pseudonyms for individuals and places in qualitative research. Confidentiality was upheld by ensuring that although the researcher was aware of participants' identities, he took definite measures to hide their identities. Aggregated findings, rather than personal data, were released to the public.

Honesty and Compliance: The researcher presented their work with honesty, disclosing any potential conflicts of interest and accurately reporting data, results, methods, and procedures. There was no falsification or misrepresentation of data.

By adhering to these ethical considerations, the researcher ensured the integrity of the study and the protection of participants' rights throughout the research process.

1.8 Limitations to the study

The researcher identified the following limitations in relation to the study:

Sampling bias: The study's sample primarily comprised faculty members, postgraduate students, university librarians, and research/reference librarians, conducted exclusively within private universities. This could potentially introduce a sampling bias, as the study's outcomes may be influenced by the selection of participants or the sampling methodology employed. In order to mitigate this bias, the researcher applied methodological transparency. In this case the

researcher provided a detailed account of the study's sampling methodology, which included the criteria used for participant selection and the rationale behind focusing exclusively on private universities. This transparency allows readers and to assess the potential impact of the sampling strategy on the study's outcomes.

Limited applicability: Given that the ultimate goal of this study was to develop a data literacy framework tailored for researchers, there is a limitation in terms of its broader applicability. The framework may have limited relevance outside of academic contexts, as the study participants were exclusively drawn from universities.

1.9   Outline of chapters

This study has a total of eight chapters. This section, therefore provides a brief outline of each of the chapters.

1.9.1   Chapter 1

This chapter provides a comprehensive introduction of the research context. The chapter delves into various definitions of data literacy and advocates for a precise and inclusive interpretation. It also highlights the emergence of data literacy initiatives and their dual importance in academia and the broader society. The problem statement emphasizes gaps in data literacy training and proposes standardized solutions. Research objectives, a central question, and sub-questions are detailed, along with the chosen research methodology, population, and ethical considerations. Lastly, the chapter acknowledges study limitations and a brief outline of the chapters.

1.9.2   Chapter 2

The chapter conducted a comprehensive review of the existing related literature on data literacy. The review focused on data literacy in relation to data related services offered by various university libraries. The chapter provided an in-depth analysis of literature related to data literacy, Research Data Management (RDM), and Research Data Services (RDS). The purpose of the review was to enable the researcher to understand the specific area of study, data literacy, in relation to other relevant fields and to draw upon existing research as a point of reference for the present study. The researcher analysed the available literature based on the study's objectives and the broader context of data literacy. The review aimed to provide a solid foundation for the study's research questions and offered insights into the existing gaps and areas for further investigation in the field of data literacy.

1.9.3   Chapter 3

In Chapter 3, the researcher established the theoretical foundations for the study, laying the groundwork for the envisioned data literacy framework. The chapter introduces key concepts essential for comprehending the study's focus and presents a background on the frameworks and models to be employed. Two theories and two models were introduced, anchoring the study in established theoretical frameworks. Additionally, an initial data literacy framework emerged from the synthesized literature and elements drawn from the identified frameworks and models, providing a solid foundation for the study's development.

1.9.4   Chapter 4

The chapter provides an in-depth discussion of the research methodology employed in the study. It encompasses the research paradigm, research approach, research design, target population, sampling methods, and sample size. The chapter also delves into aspects such as pre-testing, validity, reliability, data analysis methods, and ethical considerations, all of which collectively form the methodological framework guiding the study's execution.

1.9.5   Chapter 5

Chapter 5 provides an overview of the research findings primarily obtained through a questionnaire. The data collected consisted mainly of quantitative information, although a small portion of the questionnaire contained open-ended questions that yielded qualitative data.

1.9.6   Chapter 6

Chapter 6 presents the interview findings, which primarily comprise qualitative data collected from university librarians and research/reference librarians at the participating institutions. The chapter offers a thorough analysis of this data, shedding light on the participants' experiences, perspectives, and opinions. The results are presented descriptively. The primary goal of this chapter is to provide a comprehensive and profound comprehension of the research subject matter.

1.9.7   Chapter 7

Initially, the research findings, were presented separately in quantitative and qualitative formats in Chapters 5 and 6. Chapter 7 was dedicated to converging and discussing the results. Convergence was used to make sense of certain outcomes derived from the study's findings. In this chapter, the findings are compared and contrasted to achieve a more holistic comprehension of the research outcomes.

1.9.8 Chapter 8

Chapter 8 offers a succinct summary of the study's objectives and their successful achievement. It outlines the core components of the research design and emphasizes key findings related to the primary research question and sub-questions. The chapter encompasses recommendations based on these findings and suggests potential directions for future research endeavours. Its primary focus is on providing an in-depth discussion of the newly developed data literacy framework.

1.10 Summary

Chapter 1 provides a brief discussion while situating the aim of the study. The chapter makes reference to the emergence of data literacy programs and underscores the growing involvement of libraries in addressing the challenges associated with data literacy. Furthermore, the chapter characterizes the research problem as the need for improved data literacy in academia, and it identifies gaps in data literacy training and policies. It provides an overview of the research objectives, questions, methods, ethical considerations, and limitations, as well as stressing the growing importance of data literacy and the role of libraries in bridging gaps in this area hence laying the groundwork for the subsequent chapters.

The chapter concludes by providing an outline of subsequent chapters in which Chapter 1 introduces the importance of data literacy and sets the stage for the study; Chapter 2 provides a comprehensive review of existing literature related to data literacy; Chapter 3 establishes the theoretical foundations for the study; Chapter 4 details the research methodology; Chapter 5 presents research findings from a questionnaire; and Chapter 6 focuses on interview findings from librarians. Chapter 7 converges and discusses the results; and Chapter 8 offers a summary of research objectives, findings, and recommendations, with a focus on the newly developed data literacy framework.

CHAPTER 2

2. LITERATURE REVIEW

Chapter 2 of the research study addresses the goal to conduct a comprehensive review of the existing literature on data literacy. The review focuses on data literacy as an essential training program for researchers and as a service offered by university libraries. The chapter provides an in-depth analysis of literature related to data literacy, research data management (RDM), and research data services (RDS). Data literacy is discussed in the context of other literacies, of the data lifecycle and of FAIR principles. The researcher also considered several case studies to consider the emerging practice at academic institutions. Finally, the possible challenges in providing data literacy training were considered.

2.1 Introduction

The purpose of the review is to enable the researcher to understand the specific area of study, data literacy, in relation to other relevant fields and to draw upon existing research as a point of reference for the present study. The researcher analysed the available literature based on the study's objectives and the broader context of data literacy. The review aims to provide a solid foundation for the study's research questions and to offer insights into the existing gaps and areas for further investigation in the field of data literacy.

Robert J. Moore, in his 2014 article titled "Why data literacy will Be the Most Important New Skill of the 21st Century," argues that data literacy is an essential skill that individuals must possess in the modern era of information abundance. Moore notes the evolution of the printing industry and the challenges of accessing information in the past, in contrast to the present era where we are surrounded by vast volumes of data thanks to the evolution of information and communication technologies (ICTs). He contends that data literacy is a fundamental change in human behaviour required to cope with the broad availability of data and analytical tools (Moore, 2014).

Moore's assertion is supported by Bryla (2018), who notes that data literacy is becoming more applicable in real-life situations than previously thought. Bryla emphasizes the importance of data literacy skills for both researchers and citizens, as they help in critical thinking, problem-solving, and decision-making in various fields. In summary, the articles by Moore and Bryla emphasize the significance of data literacy as a 21st-century skill necessary for effective engagement with the modern world's data-driven environment.

Data literacy is a relatively new concept that is gaining attention in both academic research and practical applications. Several studies have shown that various institutions and organizations, including academic libraries, are offering data support services. For instance, Chigwada, Chiparausha, and Kasiroori (2017), Patel (2016), and Liu and Ding (2016) have conducted studies that demonstrate the emergence of data support services. However, some of these studies have revealed significant challenges related to data literacy. For instance, some institutions have either poor policies or no policies regarding Research Data Management (RDM). Furthermore, some findings indicate that there is a lack of institutional adaptability and awareness among top management regarding data literacy. These challenges may hinder the effective implementation of data literacy programs and limit their potential impact. Therefore, there is a need for further research to identify and address these challenges to enhance the effectiveness of data literacy programs.

Several studies have raised concerns about the role of libraries in the research lifecycle of scientific studies, especially regarding the need for rigorous processes for research acceptability, reliability, and conformability (Chigwada, 2020; Nitecki & Davis, 2017; Cox & Pinfield, 2014). This study aims to address these concerns, specifically regarding the readiness of academic libraries in the 21st century to develop data literacy programs for researchers. In the modern era, data is considered the "new currency," and it is essential to handle and manage data created throughout the research cycle with utmost care. This is necessary to ensure data confirmation, scrutiny, as well as use, reuse, and sharing, among other things.

However, many institutions fall short in implementing effective strategies to oversee the entire process of Research Data Management (RDM), which is a crucial component of data literacy. Ineffectiveness has led to the partial implementation of RDM in the data lifecycle, thus lowering data quality in some cases. This issue is highlighted by Eggers, Hamill, and Ali (2013), who assert that data is the new currency in the digital economy, and the effective management of data is crucial for organizational success.

Though, as mentioned before, that data literacy is fairly a new concept, it is steadily gaining traction, being adapted and practised within research environments especially in the global north (Fontichiaro & Oehrli, 2016). Developments and practices around data literacy have therefore led to other concerns or opportunities hence resulting in some emerging themes.

2.2 Emerging themes from data literacy literature

Despite the novelty of data literacy as a 21st century concept and practice, various themes have emerged from different studies in the wake of describing data literacy and its relation to other disciplines or practices. This study will limit discussions or views about data literacy as captured or synthesised in three themes; data literacy as research data management service, as a set of practices for data lifecycle and data literacy for decision making. The first two will predominantly influence discussions in this chapter. These themes are discussed below.

2.2.1 Data literacy with an emphasis on research data management

The advent of e-science and cyberinfrastructure has led to a situation whereby researchers are dealing with vast amounts of data which has led to a growing demand for RDM services (Khan, Kim & Chang, 2018). RDM activities are therefore supposed to be structured and anchored in a data literacy programme. Data literacy needs to be recognised as a critical skill in the knowledge base of not only researchers but also those involved in providing RDM services too (Koltay, 2016). Data literacy, therefore, entails the provision of research data management services. Different authors including Koltay (2016), Searle, Wolski, Simons and Richardson (2015), Prado and Marzal (2013) and Qin and D'Ignazio (2010a) have developed this line of thought as expressed in some of their studies as summarised in Table 2-1.

Table 2- 1: Data literacy to perform data management

Koltay (2016:303)	Data literacy is "closely related to **research data services** that include research **data management**."
Searle, Wolski, Simons and Richardson (2015:11)	Data literacy "should include activities that support researchers in building the skills and knowledge required to **manage data** well".
Prado and Marzal (2013:126)	Data literacy "…enables individuals to **access, interpret, critically assess, manage, handle and ethically use** data."
Qin and D'Ignazio (2010:3a)	Data literacy entails "…knowledge and **skills in collecting, processing, managing, evaluating, and using** data for scientific inquiry."

The subject on data literacy, with a focus on research data management, highlights the growing demand for organized RDM services in the context of increasing data and e-science.

2.2.2 Data literacy with an emphasis on the data lifecycle

Data literacy, according to this study reflects on one's ability to apply a "set of activities and practices undertaken in the production, consumption and management of data" as captured in various stages of a data lifecycle model. The "activities and practices include data creation, data acquisition, data normalisation, data analytics, data storage, data disposal, and data use/re-use" (Khan, Kim & Chang, 2018:3). An elaborate discussion on data lifecycles in relation to data literacy was done later in the chapter. Some of the studies that have captured this include studies by Okamoto (2017:120), Maybee and Zilinski (2015:2) and Carlson et al. (2011:24-25), as presented in Table 2-2.

Table 2- 2 : Data literacy as a set of practices

Okamoto (2017:120)	"…The ability to **access, critically assess, interpret, manipulate, manage, summarize, handle, present, and ethically use** data."
Maybee & Zilinski (2015:2)	Data literacy encompasses "skills related to **accessing, managing, communicating, preserving and ethically using** data."
Carlson et al. (2011:24-25)	Data literacy **skills** include: "discovery and acquisition of data; databases and data formats; data conversion and interoperability; data processing and analysis; data visualization and representation; data management & organization; data quality and documentation; metadata and description; cultures of practice; ethics & attribution; data curation and re-use; and data preservation."

This study defines data literacy as the ability to apply a set of actions and practices throughout the data lifecycle, which includes data generation, acquisition, standardization, analytics, storage, disposal, and use/reuse. This complex idea of data literacy highlights the significance of skills related to data access, management, communication, preservation, and ethical use at many phases of the data lifecycle.

2.2.3 Data literacy for sound decision making

Data literacy is evolving as a 21st-century skill that each citizen needs to acquire. According to Khan, Kim and Chang, (2018:2), data literacy is emerging as a "term of the data revolution discussion, mainly from the perspective of data analytics and data science, which supports data-driven decision making" hence viewed as a skill that each world citizen needs to acquire. Table 2-3 presents a summary of various studies that highlight the need to focus on the ability of the recipient of a data literacy programme.

Table 2- 3 : Data literacy for decision-making

Wolff et al. (2016:16)	"…Data literate reader has the ability to properly **evaluate** the evidence that is presented in these scenarios so that they can make critical judgements on the **reliability of the information** presented and can better understanding how their own contributed data is being utilized and make more informed decisions."
Johnson (2012:79)	Data literacy is "…the ability to process, sort, and filter vast quantities of information, which requires knowing how to search, how to filter and process, to produce and synthesise."
Carlson et al. (2011:5)	Data literacy is all about "…understanding what data **mean**, including how to read graphs and charts appropriately, **draw correct conclusions** from data, and recognize when data are being used in misleading or inappropriate ways."
Love (2004:22)	Data literacy is "…the ability to examine multiple measures and multiple levels of data, to consider the research, and to **draw sound inferences**."

Data literacy is emerging as a vital skill enabling data-driven decision making. It involves the capacity to evaluate evidence critically, process vast amounts of information, understand the meaning of data, and draw sound inferences, ultimately contributing to more informed decision-making and a better understanding of data's role in one's life.

These three themes continue to dominate and shape data literacy discussion as shown in a number of studies that are available today, thanks to scholars who continued to probe into its relevance and need as a 21st century skill (Goldman, Chen & Palau, 2023; Hamad, Al-Fadel & Al-Soub, 2021; Huang, Cox & Sbaffi, 2021; Goben & Griffin, 2019; Majid, Foo & Zhang, 2018). Having outlined the parameters upon which this study will be referring to data literacy, it is good to distinguish it from or show how related it is from other 21st century skills. This is in consideration that data literacy is not the only skill that is relevant to the researcher today.

2.3 Data literacy in the context of the data revolution

The current era is considered by many to be a new economic revolution, often referred to as the data era or the new industrial revolution. This revolution is characterized by the increasing use of data to solve complex social issues and develop new ideas. According to Jim Gray (2007) based on the transcript of a talk given to the NRC-CSTB1 in Mountain View, CA, as quoted by Tansley & Tolle (2009), the current era can be classified as the fourth paradigm in the

evolution of science. Gray's classification identifies four paradigms: the empirical stage, which focused on experimental sciences, the theoretical stage, which concentrated on scientific laws and theories, the computational stage, which was primarily concerned with simulating complex phenomena, and the current eScience or data exploration stage, where data is the central component. Gray's classification was presented in a talk to the National Research Council (NRC) and the Computer Science and Telecommunications Board (CSTB) in California, where he argued for the existence of the fourth paradigm. The data era, therefore, is the fourth paradigm that characterizes the current era (Tansley & Tolle, 2009).

The United Nations recognises the data revolution and its widespread impact on society. According to a report by the Independent Expert Advisory Group on Data Revolution for Sustainable Development, the data revolution is characterised by the explosive growth in the volume, speed, sources, dissemination and range of data available. This is driven by new technologies such as mobile phones and the Internet of Things, as well as non-traditional sources of data such as citizen-generated data, qualitative data and perceptions data (Morales et al., 2014). The UN acknowledges the importance of data literacy in the 21st century, which is crucial in enabling individuals and communities to leverage the potential of data towards creating a more sustainable future (United Nations Economic Commission for Europe, 2012). As such, it is essential for educational institutions to prioritise data literacy as a critical skill to develop in their students to equip them with the necessary knowledge and skills to navigate and thrive in a data-driven world. This includes not only data analytics but also data ethics and critical thinking about data (Morales et al., 2014). By fostering data literacy, educational institutions can contribute to the creation of a more informed and data-literate society, which is essential for achieving sustainable development goals.

Mandel (2017) notes in a report prepared for the United Nations that industries related to data have shown much faster productivity growth as compared to physical industries, with the digital sector in the United States experiencing productivity growth averaging 2.7% between 2000 and 2015, compared to 0.8% in the physical sector. Furthermore, the data industry is creating more jobs than the physical industry, with hours worked in the digital category rising to 9.6% compared to 5.6% on the physical side since 2007. The report underscores the importance of data as a key resource in the modern world (Kitchin, 2014), highlighting the need to incorporate data literacy as a mainstream training program in discussions about the "data revolution" to enable individuals to effectively navigate the impact and value of data in society.

The value of data is constantly rising and is becoming a focus of technological competition among international data companies such as Google, Facebook, Apple, Amazon, and Microsoft. These companies thrive and run data economies based on the amount of data they possess as it gives them leverage and advantage over their competitors (Press, 2018). The same applies to nations where developed countries and multinationals are running data economies. Gilligan (2016) confirms that the data economy comprises of organisations and governments providing personalised services and solutions to existing challenges. However, the effective use of voluminous data relies on the data literacy of these organisations and governments. Data literacy, therefore, plays a critical role in the 21st-century skill set and should emerge as a mainstream training programme in discussions about data revolution (Mandel, 2017; Kitchin, 2014).

2.3.1 Open data and big data initiatives

Data has been recognized as a crucial resource in today's economies due to the major data initiatives in the scientific sphere. Governments and organizations have spearheaded these initiatives, which have had either a direct or indirect impact on the focus on data as a key component of development (Hart, 2019). The Organisation for Economic Co-operation and Development (OECD) has launched Open Knowledge and Science Commons initiatives that emphasize the value and sharing of data in a global market. According to OECD, open access in research data can enhance the quality and productivity of science systems globally, and a declaration was made for open access to publicly funded research. However, security, property rights, and privacy measures must be considered and followed. (OECD, 2004).

In 2012, the Obama administration introduced the Big Data initiative to advance the means of managing and analysing large and diverse data sets. The initiative aimed to accelerate scientific discovery and promote new fields of inquiry by emphasizing research data as the key component in advancing science and research. (U.S. Office of Science and Technology Policy, 2012).

Similarly, the European Commission (EC) released a communication titled: *Towards better access to scientific information: Boosting the benefits of public investments in research*, in 2012, which called for improved access to publicly-funded scientific research through widely circulating publications and data with the help of digital media. The Commission strongly agreed with the OECD on the importance of disseminating scientific information as a way of encouraging scientific and technological progress (European Commission, 2014).

These initiatives aim to promote and encourage access to data from publicly-funded research. Governments have taken up the initiative of publishing all data meant for public consumption on open webs for easy access as one of the benefits (Sivarajah et al., 2017; Laboutkova, 2015). Governments' efforts and declarations can bolster data reuse, which is a crucial component of data literacy. Therefore, leading by example is essential to promoting access to research data. This means that the relationship between open data, big data, and data literacy is symbiotic, where open and accessible data, both in quantity and diversity, contribute to a more informed and data-literate society, ultimately fostering scientific and technological progress.

2.3.2   Data literacy for public data consumption

The value and impact of openly published data depend on the public's ability to interpret and comprehend it. However, the general public often finds data difficult to interpret, making it challenging to derive value from it (Gray & Darbishire, 2011). To address this issue, organisations such as the International Statistical Institute (ISI), the American Statistical Association (ASA), the International Association for Social Science Information Services and Technology (IASSIST), and the United Nations Economic Commission for Europe have championed the need to introduce data competency instruction at all educational levels. They argue that data literacy is essential for three reasons.

Firstly, data literacy empowers citizens to keep the government accountable and transparent. Citizens need to understand and interpret government data to effectively hold the government accountable. Secondly, data literacy enables people to understand and solve local problems by empowering them to work with current data and even generate their data. Thirdly, data literacy builds the financial effect of Big, Small, and Open Data. As organisations expect to gain business value from data, workers who are data competent have turned out to be valuable in the current economy (D'Ignazio & Bhargava, 2015). And finally, the link between research data and public data consumption lies in the transferable skills acquired through data literacy training, enabling individuals to navigate and leverage diverse data sources, contributing to evidence-based research, informed decision-making, and broader societal engagement with data-driven knowledge.

2.4 Data literacy in the context of other literacies

As we experience numerous technological evolutions and innovations impacting research processes in the 21st century, the quest for researchers to be well equipped with different literacies continues to grow. The gradual adoption of technologies in research today facilitates

23

quick access to information, connectivity with other researchers and self-publication. This raises the need for various literacies among researchers. There are different types of literacies that researchers in the 21st century should be yearning to acquire so as to remain relevant and competent today. Despite the many forms of literacies that exist today this section limited its discussion on information literacy, statistical literacy and digital literacy in relation to data literacy considering their close relationship. One of the studies that has tried to make a comparison between these literacies is Schield's (2005) study which delves into the discussion of data literacy, information literacy and statistical literacy. The study indicates that it is relatively difficult to distinguish between the three literacies or promote one of them without involving the rest as all of them are relevant in research today.

An elaborate discussion of what data literacy is, was done in the preceding sections of this study, including Chapter 1 and parts of this chapter. As defined in Chapter 1, data literacy is simply being able to understand, use and manage data. Diverse areas such as information literacy and statistical literacy influence key principles of data literacy (Prado & Marzal, 2013; Fry, 2004). A data literate researcher needs skills such as data retrieval and data evaluation skills which are drawn from information literacy skills. Furthermore, data literacy includes statistical analysis skills.

Schield (2005) argues that it is hard to separate statistical literacy, information literacy, and data literacy since they all deal with similar issues. Ridsdale et al. (2015) support this notion and confirm that data literacy is based on the same theoretical principles as statistical and information literacy. Therefore, these three literacies are interrelated and are essential in dealing with the challenges that students face in higher education.

2.4.1   Information literacy

Due to the rapid increase in information and information sources, the 21st century has been named as the information era. Information literacy has, therefore, become a necessary skill so much needed for one to sift through volumes of information. It empowers individuals with lifelong skills to become independent learners (Deja, Rak & Bell, 2021; Olakunle & Olanrewaju; 2019; Ranaweera, 2008).

Information literacy is all about evaluation of information depending on the need for information and the context. According to Horton (2013), information literacy entails possessing the essential skills, attitudes, and knowledge to determine when information is required to address a problem or make a decision. It involves the ability to express this need

for information effectively, conduct efficient searches, retrieve, interpret, and comprehend the information, organize it, evaluate its credibility and authenticity, assess its relevance, and, if necessary, communicate it to others, all with the ultimate aim of achieving specific objectives. For one to be considered information literate, the American Library Association (ALA) (2015) and the Association of College and Research Libraries ACRL (2000) have outlined a set of competency standards that must be achieved which include the ability to determine the need for information, ability to effectively and efficiently access information, be able to evaluate information sources, use the information and adhere to ethical requirements applied in the use of information. Furthermore, someone who is information literate should be in a position to make sense of the social, legal and even economic issues related to information. Generally, they should be able to understand the implications surrounding the use of information.

Information literacy skills could also be applied by a data literate researcher. Researchers will need to know when data is needed, where to access and evaluate data, how to retrieve data and use data. Olakunle and Olanrewaju (2019) conducted research indicating that possessing information literacy is essential for researchers as it has a positive correlation with research productivity. Incorporating data literacy into research also entails integrating certain information literacy skills when working with data. The establishment of data literacy draws heavily from the framework, training methods, and practices of information literacy (Gross, Latham & Julien, 2018).

2.4.2 Statistical literacy

Statistical literacy is a critical element of working with data considering the interpretation of data that a researcher is required to do. Gal (2002:2-3) describes it as the "ability to interpret and critically evaluate" statistical products, as well as their ability to "discuss or communicate their reactions" to statistical products.

There is a growing significance on the value of statistics in today's world of information (Yusof, 2021; Sharma, 2017). The need to apply it is steadily becoming paramount considering the data-driven context of research. Some Institutions of Higher Learning have therefore incorporated a component of statistics in research methodology courses with the understanding that researchers need it, especially during data analysis. In the context of this study, statistical literacy would, therefore, be considered to be one of the key components in a data literacy programme. Researchers cannot, therefore, claim to be data literate without some basic statistical skills which are essential for most research.

2.4.3 Digital literacy

Gradual development in Information and Communications Technology (ICT) has transformed the learning process hence acquiring a prominent role in many spheres today. With the use of the internet, there is an almost unlimited number of sources and resources of information. The use and application of ICTs in diverse spheres has therefore led to the need for a digitally literate society. Digital literacy comprises of a range of diverse digital skills. These skills are considered essential for one to succeed in an increasingly digital world. Digital literacy is described as one's ability to use ICTs to find, evaluate, utilize, share and create content.

Presently, digital technologies have impacted research processes at all levels. From the planning of research to the distribution and reuse of research data and findings, technology has come in handy either to expedite the process or make it efficient. This means that the synergy between research and the technological infrastructure requires some level of competence which every researcher must-have. This is where digital literacy comes in as it equips the researcher with skills on how to use digital technology, communication devices or networks to locate, analyse, use and produce data or information (Robelo & Bucheli, 2018).

The ever-evolving digital technologies have led to changes in technologies hence resulting in quick access to information. Furthermore, these changes have opened boundaries leading to wide collaborations in research. In research, digital evolutions have spurred a steady but growing impact since the 1990s when the concept of online research arose and was adopted in research as a "new and vibrant field of research methods" (Corti & Fielding, 2016:1). Data literacy and digital literacy are, therefore, complementing competencies that every researcher must be aware of and must strive to attain. A data literate researcher should equally be digitally literate, well equipped with skills to use digital technologies to manipulate data. Digital literacy would, therefore, be considered as a critical component of data literacy. What this study suggests is that a researcher should be able to apply emerging technologies to drive research.

To put it concisely, data literacy doesn't exist in isolation; rather, it flourishes and achieves success when integrated with other emerging literacies, as previously discussed. When designing and executing a data literacy program, it's crucial to incorporate elements from these other literacies into the process. As we delve into the conceptual framework for data literacy training (see figure 3.1), it's essential to acknowledge and account for the various literacies mentioned here.

26

2.5 Data literacy and data lifecycle models

Section 2.2.3 and Table 2-3 of the source material provide a definition of data literacy as the set of practices involved in producing, managing, and consuming data throughout the data lifecycle. Data literacy aims to enable researchers to produce high-quality data and manage it effectively. Data quality, management, and utilization can be achieved through a set of practices that span the various stages of a data lifecycle model. A data lifecycle model is a series of phases through which data passes during its lifetime (Cox & Tam, 2018). Several of these models exist (cite at least UK Data Archive, DCC, DataONE). Motivate why you chose UK Data Archive model.

Though data lifecycle starts with data planning and ends with data re-use, it is good to note that data is continually re-purposed, meaning new data products continue to be created, processed, distributed, discovered, analysed and archived. Data, therefore, could be said to have a longer lifespan considering all the stages that are involved. The lifecycle may extend beyond what is considered to be the active research project period. The researcher may continue to work on data even after funding has ceased. They may do so by carrying out follow-up projects to either analyse or add to the data. Furthermore, data may be re-used by other researchers, hence prolonging its intended lifespan.

Data lifecycle connotes the entire process of data management activities. Therefore, a data literacy programme needs to consider incorporating the training of researchers on how to adhere to the needs of all the stages of a data lifecycle. While offering an overview of the different stages involved in successful management and preservation of data for use and reuse, Plale and Kouper (2017) assert that data lifecycle remains to be key in the achievement of invaluable research.

The increase in the emphasis on data sharing today highlights the importance of data lifecycle as a means of achieving data quality (Higman & Pinfield, 2015). Different scientific disciplines, contexts and even research intentions have led to the development of various data lifecycle models to achieve data quality. Today, according to a study by Sinaeepourfard et al., (2015) there are different versions of data lifecycle models. They differ in terms of focus, perspective, or relevance of the discipline. According to Carlson's (2014) description, the classification of research data lifecycle models could be in various forms, that is linear, circular, non-linear or other models. The form depends on the context of the model which could be individual-based, organisation-based or community-based.

27

According to Wissik and Durco (2016), the different versions are attributed to variations in practices across various fields of study and disciplines. Sinaeepourfard et al., (2015:1) assert that "there is no global and comprehensive framework, from data creation to data consumption, to be widely utilized in different fields". However, despite the differences, there are key concepts that cut across which each model entails. Broadly, we can distinguish six stages of a data lifecycle which include: data planning, data collection, data processing, data publication, data preservation and data re-use.

As a demonstration of the data lifecycle, this study will adapt a UK Data Archive research data lifecycle model as shown in Figure 2.1. The UK Data Archive model emphasises acquisition, curation and archiving of digital data (Hart et al., 2016). The original phases of the model include data creation, data processing, data analysis, data preservation, data access and re-use. The model has been lauded as being "a good choice for archiving and discovering data across the digital data". However, due to lack of data planning and data collection as key stages in the model, the researcher finds the UK Data Archive model limiting without these two. This study will, therefore, adapt the UK Data Archive model and modify it by adding data planning and data collection stages in order to complete the lifecycle. The reason for the addition of these two stages is that the researcher considers data planning and data collection to be very important. Every researcher needs to be well versed in planning their research project as well as us in data collection. The two stages have a significant impact not only on the outcome of the research but also on the quality of data for posterity.

Hence, Figure 2.1 illustrates the seven steps to be executed at various stages of the research lifecycle for both the successful culmination of a research project and the attainment of high data quality, incorporating data planning and data collection as additional components. Comprehensive descriptions of each of these phases can be found in subsequent sections.
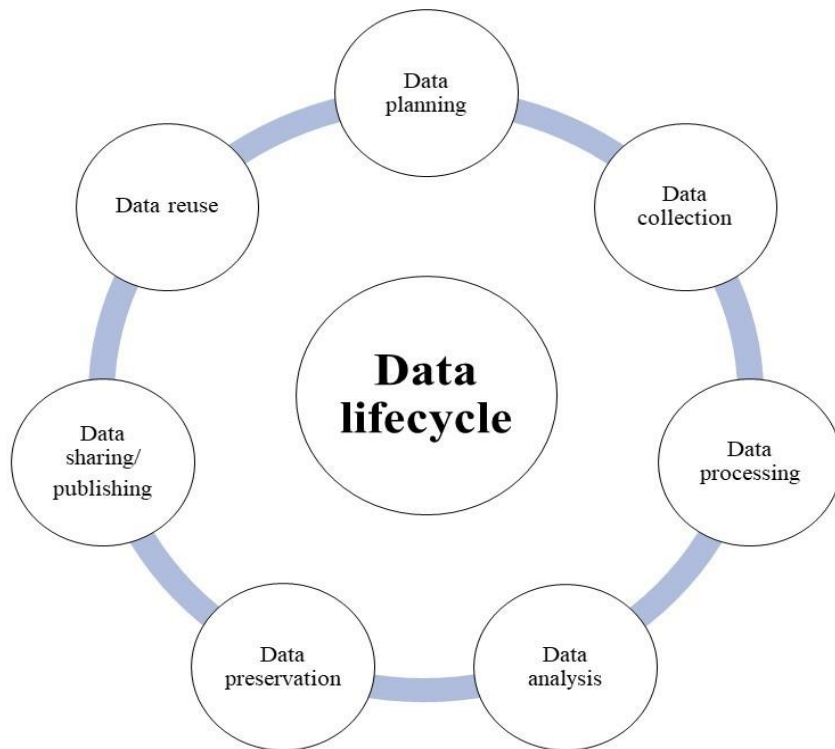
Figure 2- 1: Research data lifecycle

Source: Modified UK Data Archive research data lifecycle model (2019)

2.5.1   Data planning

Developing a DMP before commencing research is a good research practice as it guides the entire research project against duplication, data loss, theft, inappropriate use of data, and security breaches, ensuring the accuracy, reliability, integrity, quality, and security of data (Gajbe, Tiwari & Singh 2021; Deng & Hu, 2014: 4). It also helps the researcher in developing procedures in the initial stages of the project, maintaining consistency, and saving time and effort. Moreover, a DMP increases the researcher's profile through data dissemination and re-use.

Effective data management is critical to the research process as it enhances data quality, facilitates data sharing, promotes collaboration, and ensures research integrity (Janssen, et al., 2020; Tenopir et al., 2015). Data management plans have been identified as a useful tool for ensuring effective data management (Mannheimer, 2018; Jones et al., 2013). A DMP is a document that outlines the steps that researchers will take to manage the data generated, collected, or used during a research project. The document provides information on data organization, storage, backup, preservation, sharing, data privacy, ethical considerations, and legal compliance with relevant regulations (Borycz, 2021; Jones et al., 2013).

29

Despite the benefits of DMPs, their adoption among researchers is still low, particularly in developing countries, where researchers face several challenges that hinder their adoption, such as lack of awareness, limited resources, and inadequate institutional support (Buhomoli & Muneja, 2021; Williams, Bagwell & Zozus, 2017). Additionally, researchers may have limited knowledge or skills in data management practices (Birkbeck, Nagle & Sammon, 2022; Tenopir et al., 2015).

In the context of open data and research collaborations, data planning has become a crucial process, enabling researchers to identify potential areas of difficulty or conflict, which can be resolved smoothly by research colleagues and collaborators before they escalate (Gray & Darbishire, 2011:11). In some instances, a DMP is considered a requirement by research funders, Research Ethics Committees, and institutional RDM policies. Therefore, researchers are obliged to submit a DMP alongside their research proposals.

The choice of a DMP tool depends on the researcher, the funder, or the magnitude of the research project. Researchers can choose between paper and electronic DMP tools, and some funders have their own customized DMP tools, in which researchers must fill in the appropriate details. Today, most research projects use electronic DMPs, such as the DMPonline developed by the Digital Curation Centre (DCC). The DMPonline tool provides customized guidance to help researchers develop their data management plans and includes different templates for funders to customize DMPs as per their specific requirements.

Therefore, developing a DMP is a necessary skill for 21st-century researchers. Both postgraduate students and faculty engaged in research need to be aware of and trained in DMP development. Any data literacy training program aimed at empowering researchers with data skills would, therefore, need to incorporate DMP development.

2.5.2   Data collection

Data collection entails the gathering of data depending on the research's variables (Chigwada, Chiparausha & Kasiroori, 2017). Data collection is also described as the phase which entails carrying experiments or survey or making observations. Every research is fulfilled by the collection of data, be it primary or secondary data. The collected data enables the researcher to respond to the identified research objectives or research questions. The collected data would also help in testing the identified hypotheses or evaluate outcomes. The whole process of how to collect data and the methods to be used depends on the nature and magnitude of the study.

The output of research is heavily impacted by the quality of the data collected. The validity of the findings of a study depends heavily on the process of collecting data. And therefore, the researcher must ensure its accuracy. A data literate researcher should, therefore, be able to know where to collect data, how to collect data, methods for collecting data and the different instruments to be used in collecting data. Data collection would, therefore, form part of a data literacy training programme especially for postgraduate students.

2.5.3   Data processing

Data processing is basically starting to work with the collected research data in order to achieve the study's objective. The collected raw data makes no sense unless it is processed. Data processing involves converting unprocessed data into meaningful information (Rudo, 2013). The process would, therefore, include a sequence of steps carried out to extract useful information from raw data such as data entry, data digitisation (especially where data had been collected using non-electronic means), data transcription or coding and data translation. At this stage, the researcher is also supposed to do data cross-checking, data validation as well as data cleaning and data description. Data processing would also include anonymization of data to ensure no data reveals the identity of the respondents (Vickers, 2011).

All these activities are carried out in order to achieve data quality. One of the criticisms against the UK Data Archive research data lifecycle model by Sinaeepourfard et al. (2015) is that the model does not have measures in place to achieve data quality. However, the researcher holds a contrary opinion considering all the activities that take place within the third phase of the lifecycle which indicates that these series of steps are geared towards achieving good quality of data before its analysis.

Data processing involves data cleaning. Data cleaning in this context might involve calibrating data instrument, separating signal from noise, especially where data was collected using an audio recording. To achieve quality data to work with, researchers, therefore need to be well equipped with data processing skills. They shouldn't just be aware of the series of stages in the data process but also should be able to do it. And in that case, considering the huge amount of data a single study may produce, which researchers must work with, there has been technological enhancement that allows researchers to process large volumes of data with ease. From computer programming language to computer software, these technologies have become handy in the lifecycle of research data, including in the processing of data. Though data

processing can be outsourced by a researcher, general knowledge of the researcher especially on how to use some of the ICT tools for data processing is key.

2.5.4 Data analysis

Once quality data has been obtained through data processing and cleaning, the researcher must move on to analysis, which involves extracting meaningful information from the processed data (QuestionPro, 2019). While some research data lifecycle models combine data analysis with data processing, the UK Data Archive model treats them as separate stages due to the unique activities required at each stage. Data analysis involves examining the processed data to produce insights and information that form the basis of the research findings. This stage includes activities such as data interpretation, derivation, and production of research outputs from the data (Bhat, 2019).

The main objective of data analysis is to gain a comprehensive understanding of the content, context, and quality of research data. During the analysis phase, researchers generate new datasets, draw conclusions, and document the methods used. Statistical analysis is the most common activity in data analysis, which involves applying statistical methods to data to identify patterns, detect trends, make generalizations, and estimate the level of uncertainty associated with the data. The specific activities involved in data analysis may vary depending on whether the researcher is interested in obtaining quantitative or qualitative results (Islam, 2020). Consequently, there are various software tools available for carrying out either type of analysis.

Data analysis can be performed using either quantitative or qualitative methods, and the choice of method influences the selection of software or platform used for the analysis. Researchers require advanced data literacy skills to analyse complex data, particularly in the context of big data. For quantitative data analysis, researchers commonly use software tools such as Statistical Package for the Social Sciences (SPSS) and R, which is designed for statistical computing and graphics (Field, 2018; Ong & Puteh, 2017). The choice of software depends on the amount of data being analysed, with R being particularly suited for large datasets due to its ability to run on multiple platforms, including UNIX, Windows, and macOS. In contrast to SPSS, R offers a wider range of statistical and graphical techniques. Other quantitative data analysis software tools include Python, Strata, and Apache Spark.

Researchers can also use specific tools to analyse qualitative data such as texts, graphics, audio, or video. For instance, NVivo is designed for qualitative data that includes large volumes of

32

text-based information (Mortelmans, 2019; Zamawe, 2015), while other options include MAXQDA and Atlas.ti. It is important for researchers to be aware of the availability of these tools, which can make data analysis easier. Researchers should also consider which tools they intend to use based on whether their data is qualitative or quantitative, starting from the data collection stage.

The final step in data analysis is data visualization, which involves presenting data in a visual format to aid comprehension. Data visualization is founded on the principle that "a picture is worth a thousand words", (Park et al., 2022; Kirk, 2019). Rahlf (2019) explains that visualization allows the researcher to communicate the essence of the data in a clear and succinct manner. By using illustrations, readers can quickly comprehend the main findings of a study. Therefore, researchers can benefit from a comprehensive data literacy training program that includes data visualization skills.

2.5.5 Data preservation

According to Jharotia (2018), preservation involves the curation, conservation, and safeguarding of a document, while also ensuring that it remains accessible and usable for future purposes. After carefully analysing and selecting data, it is crucial to preserve and archive it in a suitable format and location. It is advisable to preserve all data collected, even if the researcher does not plan to use it. As a result, it is essential to establish policies, regulations, and strategies that focus on protecting and prolonging the lifespan and authenticity of the collected data.

Data preservation refers to the process of safeguarding and maintaining the security and integrity of data. The aim of data preservation is to retain data for a specified duration for future use, which may involve archiving or depositing it in a data repository (Groenewald & Breytenbach, 2011). Data preservation can be carried out in a local or non-public location. Effective data preservation requires proper management, necessitating the appointment of a responsible individual or team. Technical actions and procedures are typically involved in the data preservation process, with their specifics dependent on factors such as the type, volume, and intended use of the data. These technical actions may include migration of data to a suitable format and medium, secure data storage and backup, creation of metadata, and archiving. In addition, to complete the preservation process, data may need to be licensed for reuse, retention schedules established, and access controls specified (Rafiq & Ameen, 2022).

In anticipation of data sharing and reuse, researchers are faced with the task of making informed decisions regarding the preservation of their research data. A data literacy program tailored for researchers ought to provide a platform for training on data preservation. An accomplished data literate researcher must possess knowledge of all the essential activities that guarantee the longevity of their research data. Such individuals must also possess the requisite skills for identifying various data formats, choosing suitable data repositories, and devising effective data naming strategies that enhance discoverability.

2.5.6 Data sharing

The UK Data Archive research data lifecycle model includes a phase referred to as data sharing, which involves granting access to research data for the purpose of research collaboration Urbano, Cagnacci & Initiative, 2021). In this study, data sharing and data publishing are used interchangeably. Bossaller and Million (2023) identify publishing or sharing data as a crucial phase in the research data lifecycle, characterized by the ability to prepare, release, and disseminate high-quality data to the public and other agencies. Penev et al. (2011) further describe data publishing as the act of releasing research data in a published format, either online or in print, to enable use and reuse by other individuals.

The dissemination of research data has become an imperative aspect of the open data movement, as data serves as the fundamental basis of research, necessitating its sharing within the practitioner community (Bierer, Crosas & Pierce, 2017). Open science has experienced an upward trend in acceptance, whereby research data is openly shared through public repositories. This practice is widely supported by numerous institutions, funding agencies, and journals (Campbell, Micheli-Campbell & Udyawer, 2019:95). The publication of data may be carried out as a standalone product or in conjunction with the scholarly articles it supports. Certain research funders and publishers, particularly government agencies, require that data from all funded studies be made available for public consumption (Couture et al., 2018; Pasquetto, et al 2019; Kimbroughand & Gasaway, 2015).

For data to be shared, it is essential to include a data citation that specifies the terms of access and where to locate it. To enhance the discoverability of data and its accessibility, the data repository should make its metadata available online (Beretta, et al., 2021). Nevertheless, access to data is subject to authorization, as confidentiality concerns may lead to restrictions on its accessibility, despite its public availability (Williams & Pigeot, 2017). It's in this case

that there is need to train researchers in order to be aware of what is involved in data literacy. As Palsdottir (2021), Data literacy is therefore a prerequisite for the sharing of research data.

2.5.7 Data re-use

Data publication and sharing are not ends in themselves (Pasquetto et al., 2019). The purpose of the two phases or practices supersede what they are. In the era of open science and open data, research data sharing or publication is also aimed towards allowing other researchers to have access to others' data for re-use. For clarity purposes, Pasquetto, et al. (2019:3) describe data reuse as when someone else other than the originator, retrieves data in order to re-use. They emphasise that "when a repository consists entirely of datasets contributed by researchers, available for use by other researchers, then subsequent applications of those datasets would be considered reuse" (ibid). In other words, data re-use would mean taking research data that was collected to accomplish one purpose and use it to accomplish another.

Data re-use elevates the researcher into a scientific community where they are not only willing to share their data but also allow other researchers to use it. Furthermore, there are some benefits associated with making research data reusable which include higher citation rates and an increase in eligibility for funding. Data reuse therefore depends largely on data preservation and avenues of sharing (Tenopir, 2015). While preserving data, the researcher should use formats that would allow future access and reuse by other researchers. Furthermore, avenues used to share data should be channels that allow access and re-use. Therefore, it is a vital stage not only within the lifecycle of research but within the realm of research as a whole (Khan, Thelwall & Kousha, 2023; Wallis, Rolando & Borgman, 2013).

According to Hilgartner and BrandtRauf (1994:359), data should not be conceptualised "as the end-products of research, but as part of an evolving data stream". Because of this reason, creators of data should allow the reuse of data beyond the initial purpose of their creation (Dallmeier-Tiessen, et al., 2014. Zimmerman (2008: 634) insists that "for data to be reused, they must be able to travel beyond the location in which they were produced". A case in point is research in the medical field which widely thrives on data re-use. With the conviction that their research is meant to influence the world, and they work in a community of researchers, every single researcher needs to be trained on how to give rights and access to their data for re-use. Similarly, they need to be trained on how to access other researchers' data for re-use.

In summary, being the primary custodian and manager of research data, every researcher must be aware of the different stages that his/her data needs to go through. Every stage that a

researcher's data goes through has significant implications that each research must be aware of. Learning how to carry out data planning, data collection, data processing, data analysis, data preservation, data sharing and data re-use is not typically an explicit part of postgraduate and faculty training. In cases where this happens, there is a lack of an in-depth analysis of each stage for a researcher to be fully aware of what is needed in each stage. This means that a comprehensive data literacy programme for both the postgraduate students and faculty members covering these stages would be ideal hence equipping them with the most necessary skills as a 21st-century researcher.

When data has been published or shared and permission for re-use is granted, it means that other researchers may access it and use it. They may re-use it to corroborate the already existing findings or, through further cross-examination and analysis or they may generate new insights. At the re-use stage, data becomes the new raw material that a researcher is ready to work with to generate findings. However, data re-use must take place within guiding principles such as FAIR data principles. FAIR data there is another aspect that researchers need to be trained so that to understand what is involved.

2.6   Data literacy and FAIR data principles

Other than ensuring researchers are well versed with the lifecycle of data, there are other principles which are equally necessary that every researcher should be aware of. Data preservation which envisages giving metadata to datasets, data sharing and data re-use, is aimed towards ensuring adherence to good use of data commonly referred to FAIR data principles (Llebot, Castillo & 2023; Elouataoui, El Alaoui & Gahi, 2022).

Findability, Accessibility, Interoperability and Reusability of data depend on well thought and planned research data practices that every researcher in the 21st century should be knowledgeable of. Adhering to FAIR principles benefits researchers by maximizing the impact of their study, boosting the visibility and citations of their work, and improving the reproducibility and dependability of research (Carballo-Garcia & Boté-Vericad, 2022). Consequently, application of FAIR data principles does not only interest and attract new research partners but also drives and leads to new research domains (Carballo-Garcia & Boté-Vericad, 2022; Wilkinson, et al., 2016).

The FAIR principles were published in 2016 and since then they have been widely cited endorsed and adopted (Jacobsen et al, 2020). FAIR is an acronym which stands for Findability, Accessibility, Interoperability and Reusability of data. FAIR principles, therefore, provide a

convenient basis for data sharing in order to maximise use and reuse. The principles make clear the kind of characteristics that data needs to have to augment reuse, by humans and machines (European Commission, Directorate-General for Research and Innovation, 2018). According to Research Data Alliance (2020:4), FAIR principles are a "set of related but independent and separable guiding principles and practices that enable both machines and humans to find, access, interoperate and re-use research data and metadata".

The FAIR principles define how to organise research results so they can be viewed, interpreted, shared and reused more easily. Major funding bodies, including the European Commission, encourage and implement the use of FAIR data to maximize the value and impact of their investment in research. To underscore how data literacy and FAIR data principles are related, and why it is necessary for every researcher today to be aware of their application in research, the discussion of each of the principles is done below while expounding on what each of them entails.

2.6.1 Findable

To be able to share data with the expectation that it will be reused, it must be made findable. Any data created within the realm of research especially with the expectation of sharing it within the community of researchers and providing reusability needs to be made findable. To make data findable, the researcher must endeavour to describe it using relevant and unique metadata hence distinguishing it from other data. Data should also be "registered or indexed in a searchable resource". In making data findable, the datasets ought to be given a persistent and unique identifier. Other ways of making data findable include providing any attached code that permits using data as well as "research literature that provides further insights into the creation and interpretation of data" (Collins et al., 2018:19).

For researchers whose interest is their research to reach a wider community of people, then they are mandated to make their data findable. Researchers are encouraged to make their data findable as it is one way of expanding their impact among their peers and beyond.

2.6.2 Accessible

After making data easily findable by using relevant metadata to describe it (Garnett, et al. 2017) the researcher is urged to make it accessible. Data accessibility is only possible if authorisation has been granted. However, researchers should be made aware that accessibility in FAIR does not mean Open Access (European Commission, Directorate-General for Research and Innovation, 2018). As Open Access to data would mean free access to data, accessibility in

37

FAIR means making it possible to access data by providing clearly stated conditions and guidelines for one to be able to access (Harris, 2012). In the event that there are technical protocols or mechanisms for data to be accessed, then all these must be made clear in the easiest way possible. The researcher should therefore clearly define who and how the actual data can be accessed. The provided metadata should be clear in specifying the conditions under which data access would be permitted.

2.6.3 Interoperable

Interoperability involves facilitating the reuse of data by ensuring it can be seamlessly integrated with other datasets, applications, and workflows, both by humans and computer systems (Carballo-Garcia & Boté-Vericad, 2022). Technically, data interoperability means that data is "encoded by a standard that can be read on all applicable systems" (Collins et al., 2018:20). Therefore, data interoperability implies data re-usability. For data to be interoperable, researchers should ensure that their data and metadata are described within the parameters of FAIR principles. In order to make data interoperable, researchers are encouraged to use a broadly applicable language which is formal, accessible as well as which is shareable. While creating the possibility of data being interoperable, the researcher should be aware that it is necessary to use controlled vocabularies and a well-defined framework to describe and structure (meta) data in order to ensure findability and interoperability of datasets.

Researchers can employ open formats and technologies to improve the interoperability of their data, allowing for easier reuse by other researchers. Furthermore, when applicable, they may use relevant metadata standards or community-approved schemas, controlled vocabularies, keywords, thesauri, or ontologies (Carballo-Garcia & Boté-Vericad, 2022).

2.6.4 Reusable

Data reusability significantly depends on its proper description, which includes facilitating good citation. Other conditions under which data can be used should be made clear. To achieve data reusability a "clear and accessible data usage license" is to be provided. This includes conditions which clearly spell how the data should be accessed and used by both humans and machines.

The researcher has the responsibility to ensure the findability, accessibility, interoperability and re-usability of their data. In the changing publication environment where it is a growing requirement by publishers as well as research funders that data should be openly shared, the awareness and application of FAIR data principles has become more relevant. Furthermore,

38

the increasing availability of online resources means that data needs to be created with longevity in mind. Researchers should therefore be aware that by providing other researchers access to their data, they are facilitating knowledge discovery and improving research transparency. Unfortunately, there is an imbalance in the uptake and practice of FAIR data principles between countries in the global north and those in the global south. As implementation of FAIR data principles in the global north seems to be on the rise, it is different for those in the global south. For instance, lack of education and training of researchers in Africa has been highlighted as one of the challenges hindering full adoption, implementation and practice of FAIR data principles (van Reisen et al, 2020). To remedy this challenge, this study proposes the implementation of a data literacy program for researchers in universities as it is a breeding ground for researchers. The programme should consist of training researchers on application of FAIR data principles.

The next section of this chapter will review why there is need to introduce data literacy training in universities, why the focus should be on postgraduate students and faculty members, and why the library is considered to be well experienced in spearheading the implementation of the programme.

2.7   Data literacy programme for university researchers

As it was pointed out (see Section 2.2.2), data-driven decision making has gained significant interest in various sectors of society. The emphasis on using data and evidence to inform practice is on progressive trend. The education sector is not an exemption. There is a steady and increasing interest among academicians or academic institutions to embrace the use of data. In their study on data literacy in education, Mandinach and Gummer, (2012) point out that data has become increasingly important in education. They are of the opinion that academic institutions must, therefore, be the "driving force in improving data literacy" considering its vital role in decision making (ibid: 34). However, this study considers academic institutions as the key "driving force in improving data literacy", for the sake of empowering researchers (post-graduate students and faculty members) with research skills that are relevant in the 21st-century realm of research.

Today, in recognition of the fact that data drives discovery, decision making, and innovation, there is a critical need to develop data literacy skills in students and researchers worldwide. Universities, as academic institutions, outline research as one of the key pillars or goals and therefore data literacy would be reinforcing and imparting necessary skills for its (university's)

community of researchers. Graduate students and faculty members are considered to be researchers because it is presumed that they often engage in research. They play a critical role as members of the research community in a university setting. Their participation and involvement in research activities bring them into contact with data through data collection, processing, analysis or generally, the management of data as a process. For that reason, the academic world is beginning to recognise the importance of preparing and arming members of their respective research community with data skills for the workplace and society as a whole, which is increasingly becoming data-centred (Mandinach & Gummer, 2012).

In the milieu of research reproducibility and responsibility, researchers are finding themselves under the obligation to share data sets of their studies for re-use. Requirements from granting agencies and scholarly publishers emphasise on the need for researchers to share data. Unfortunately, the know-how required to fulfil the funders' requirements are not strongly emphasised and incorporated as part of postgraduate training. Researchers are left to figure out how to manage data on their own (Carlson, et al., 2015).

The 2015 Research Data Access and Preservation (RDAP) summit organized by the Research Data Alliance (RDA) and the University of Minnesota focused on how libraries are building data literacy instruction services for undergraduate students, graduate programs and researchers across the disciplines. During one of the panel sessions, members focused on describing various data literacy-related programmes in different institutions of higher learning geared towards various groups that are involved in research. According to the panel, today's researchers are under immense "pressure from federal agencies, scholarly publishers, disciplinary societies and their peers to administer their data in ways that enable them to be discoverable, understandable and used by others. However, the knowledge and skills required to fulfil these expectations are not often included as a part of higher education" (Carlson, et al., 2015:14). The summit produced a report, published in a book, *Data Information Literacy*, detailing why graduate students and faculty members need to be helped in building data literacy skills (Carlson et al., 2015).

General practice in many universities is that postgraduate students are required to write a research project (thesis/dissertation) and submit as a requirement for completing their degree. On the other hand, faculty members get assessed based on the outputs of their research engagements. These put more pressure on universities to implement a programme that considers offering training to researchers on how to deal with data (Rasul & Singh, 2010).

There are previous studies which reveal a level of disparity when it comes to data literacy competencies among researchers despite their importance (Pothier & Condon, 2020; Raffaghelli & Manca, 2023). Figure 2.2 from Carlson et al. (2015: 53) study, summarises the comparison that was done between postgraduate students and faculty ratings of the importance of data literacy competencies. The figure shows how essential or not important is each competency to each group of participants (students or faculty). Respondents were to rate the importance of each competence. Each competence had choices of ratings on the scale of 1-5 where, 5 = essential; 4 = very important; 3 = important; 2 = somewhat important and 1 = not important.
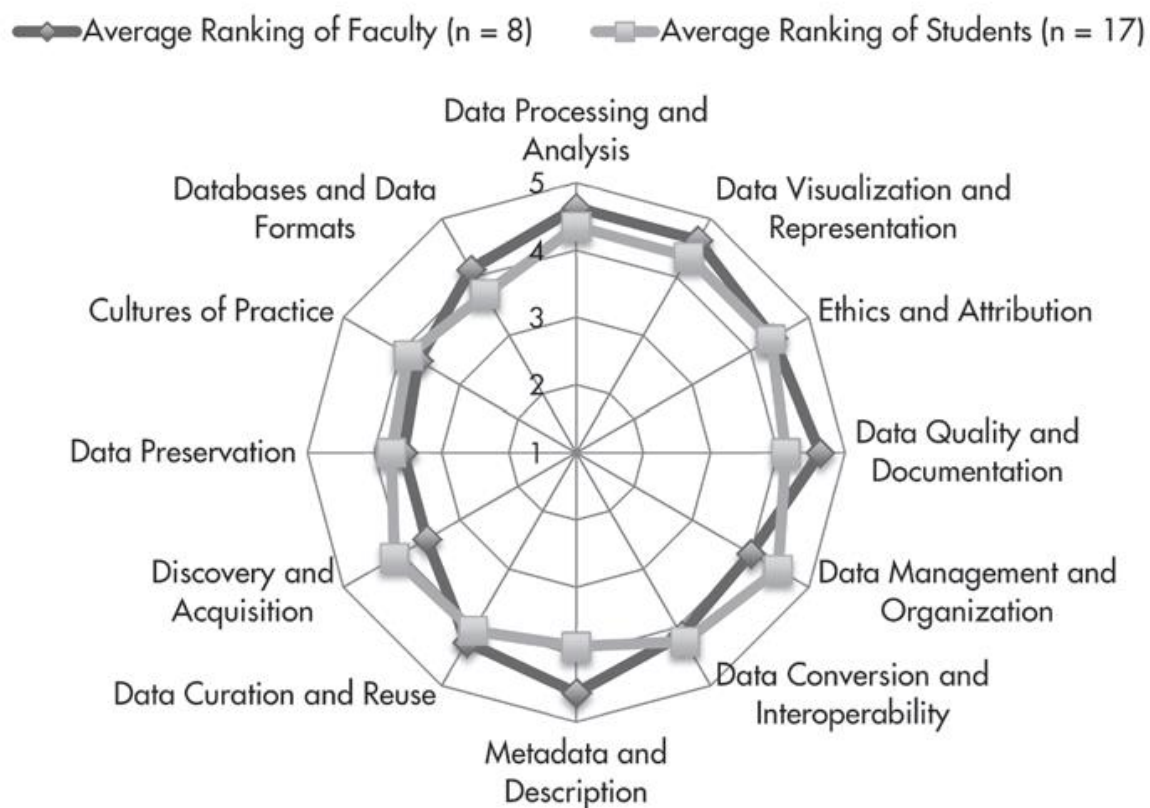
Figure 2-2: Comparison of faculty and student ratings of the importance of DIL competencies
Source: Carlson et al., (2015: 53).

The study singled out twelve basic competencies. Generally, the two groups of participants in the study, students and faculty, rated the value each competency as either "important," "very important," or "essential." Even though the analysis revealed some variance in response, the rating revealed a fundament need for data literacy to the two groups of researchers in a university (Carlson et al., 2015).

41

Considering the focus groups for this study, the following sections (2.7.1 and 2.7.2) will detail why post-graduate and faculty members in a university are singled out as groups of researchers that need data literacy training.

2.7.1 Postgraduate students

Postgraduate students are singled out as key members of the research community in a university and being on the frontline of research processes in the university as they handle various types of research data in the course of their studies (Rasul & Singh, 2010; Carlson et al., 2015). Their engagement in research leads them to collecting data, processing data, analysing data and managing data. This provides us with a starting point of examining their data literacy skills. In Kenya, it is a requirement for all postgraduate academic programmes to have a research component. One can only graduate with a master or doctorate after having written, defended and submitted a research thesis/project/dissertation which carries a total of 90 credits (CUE, 2014). Other than the final thesis/project/dissertation, postgraduate students are also required to publish at least two articles before graduation. Furthermore, they are encouraged to participate in conferences by presenting research papers.

These requirements point to the status of postgraduate students as researchers in Kenya. They are part of the research community in a university hence disserving any attention that other researchers get including research services offered or to be offered by a university. Considering the voluminous amount of data, they have to deal with in their research activities, postgraduate students deserve to be considered as beneficiaries of any data literacy program.

To emphasise the challenges postgraduate students, experience while handling data, Carlson et al. (2015) in their report point out on how most of the graduate students are given the task of data management without prior preparation or training. Furthermore, they cited challenges such as students' unfamiliarity with practices and techniques such as data documentation and different options of data storage available. The study also highlighted students' lack of understanding of the value of data as a challenge. According to the findings of the report, these are key takeaways indicating why graduate students need sufficient training in data literacy.

Equipping postgraduate students with data management skills should, therefore, be viewed as an exercise in posterity because these are lifelong skills. As emphasized by Wiley and Kerby (2018) and Sabzwari, Bhatti and Ahmed (2012), the early cultivation of effective data management practices is paramount for graduate students. These habits are essential as they pave the way for their development into future researchers. The authors insist that many

42

graduate students who work in laboratories or research teams are responsible for managing research data on a daily basis, yet they often lack formal training in data management. Therefore, it is important to provide these students with training opportunities to learn how to handle data effectively.

Some multi-disciplinary studies corroborate Wiley and Kerby's (2018) view. Conducted across various institutions, these studies commonly indicate a limited awareness or comprehension among postgraduate students regarding data management. According to findings from a study by Adamick et al. (2013) at the University of Massachusetts-Amherst, which was aimed at assessing the needs of graduate students, most of the graduate students being members of the research community of the university, need data management skills just as much as the faculty members. A survey done at Syracuse University to assess postgraduate students' data literacy awareness revealed that the majority of students lacked understanding of core skills required for efficient data management (Sharma & Qin, 2014). A similar study at Oregon State University found that postgraduate students were willing to practice good RDM but lacked the necessary expertise. According to a study by Doonan, Akmon and Cosby (2019:15), which assessed how social sciences graduate programs in the USA were including data management and data sharing in students' formal training, there was a "deficit of training in both data management and data sharing".

An underlying point that arises from these studies is that lack of basic data skills is depriving graduate students a 21st-century lifelong skill which is necessary for their research life (be it while still studying or after their studies). According to Carlson et al. (2015), lack of a structured program embedded in graduate studies is leading most of the postgraduate students to seek help elsewhere on how to manage data. This indicates the need for graduate students to be provided with adequate training which would enable them to be data literate (Giese, et al., 2020; Smith, 2008; Fischer, Rohde & Wulf, 2007; Barrie, 2004; Zamorski, 2002).

2.7.2 Faculty

Faculty members form part of the community of researchers in a university. Other than teaching, the faculty members engage in research which contributes to their professional development (Johnson & Steeves, 2019; Rahimi & Weisi, 2018a, 2018b; Wald & Harland, 2017). Their engagement in research plays a significant role in their academic ranking. They, therefore, remain to be core stakeholders in increasing research-related revenues in a university by participating in research themselves (Huenneke, Stearns, Martinez & Laurila, 2017).

In the Standards and Guidelines for Universities, the Commission for University Education in Kenya singles out research as one of the key requirements in the development and ranking of a faculty member. For a faculty member to be hired and promoted in Kenyan universities, they must have published some articles in peer-reviewed scholarly journals, among other requirements. Other requirements include their teaching experience, the total number of postgraduate students one has supervised, community engagement activities, as well as research funds one has managed to attract. The requirements increase with the seniority of the position. In measuring and evaluating universities' performances using webometric rankings, research publications, citations and web visibility are considered important indicators, hence faculty members are encouraged to engage in research (Weng'ua, Rotich & Kogos, 2017: CUE, 2014).

What this means is that, faculty members have to engage in continuous research where they deal with various forms of data. Considering the magnitude of data, they are going to get exposed to in their various disciplines of study, faculty members should be one of the major groups to be considered for data literacy training (Piracha & Ameen, 2018). In a study by Wiley and Mischo, (2016) disciplinary differences among faculty members reveal significant and diverse challenges which require focused data management attention. The study recommends the need for academic libraries to take charge and implement data management services. Another study by Adika and Kwanya (2020) revealed existing gaps in RDM literacy levels amongst lecturers after analysing the required skills by faculty members at Strathmore University in order to support RDM. The study recommends RDM faculty training as a mitigation measure.

It should not be assumed (as it is commonly the case) that the ability to do research depends only on the researcher's inherent intelligence and tacit intellectual abilities. It also depends on the researcher's development facilitated by systematic training programs aimed at equipping their research skills. The gap in data literacy skills among postgraduate students and faculty members in a university context as researchers, should therefore drive the need for implementation of systematic development programs which should cover key data management areas aimed at providing necessary skills needed by researchers. These groups of researchers deal with data and lack of adequate skills could be a challenge in their research endeavours hence that needs to be addressed. One way of responding to these needs is tasking the responsibility of training researchers in data literacy to a department within the university that would ensure researchers have attained the most relevant skills.

Considering their involvement in research activities, academic libraries have been called upon to fill the gap by use of their information literacy training experience to educate researchers on matters to do with data handling and management. This study singles out the library as the most relevant department that could coordinate training of researchers in data literacy considering its (library's) past experience in offering Information literacy programs and its engagement with researchers. In a university setting, academic libraries play a significant role in supporting teaching, learning and research activities. Based on libraries' previous involvement with research activities, they are singled out as being in a better position to implement data literacy initiatives for researchers as well as mobilising resources that could support the growing demands in the data age (Carlson & Johnston, 2015). Section 2.8 of this chapter will lay the case for the role of academic libraries in spearheading data literacy programs for researchers in universities. The literature in this section will provide information on the readiness of university academic libraries in developing data literacy programs.

2.8   Academic libraries as drivers for data literacy programmes in universities

Most of the universities have a directorate of research and post-graduate studies or generally an office or department that coordinates research in universities. The directorate is responsible for various activities concerning research, including monitoring research activities, planning and organising research seminars, trainings and workshops. It is also responsible for developing and enforcing research-related policies. However, research directorates or departments do not work in isolation but in tandem with other departments. They collaborate with other departments in ensuring research services are timely provided to users. The role of academic libraries in collaborating with the research directorate or department in promoting and supporting research activities cannot be underestimated. This has been one of the transformations that has shaped some academic libraries in the past few years as captured in some studies (Cox et al, 2019; White & Cossham, 2017; Pham & Tanner, 2015; Corrall, 2014).

Libraries are leading in offering Research Data Services (RDS) considering some funders and publishers requirements (Farooq, Heppler & Ehrig-Page, 2022; Pryor et al., 2014). Some studies have shown the magnitude of demands that researchers are experiencing today while dealing with data. Other studies show that some researchers are willing to share and give access to data, which is becoming a requirement today (Klain & Shoham, 2019; Schmidt, Gemeinholzer, & Treloar, 2016). Unfortunately, they (researchers) lack the expertise to enable them to comply with these requirements. They need to be assisted by being trained on how to create metadata and how to locate datasets. They also need to be trained on how to locate where

to store or archive data. Sewerin's (2015) study points out that in some countries such as the USA it is a standard practice for researchers to include a DMP while applying for publicly funded research grants. Based on this, academic libraries in the USA play an important role in supporting faculty needs in line with such requirements.

Some studies have previously provided solutions to these challenges by indicating that libraries should be willing to provide support in such areas as a transformation of their traditional role in providing research and reference services (Klain & Shoham, 2019; Si et al., 2015). According to Cox et al. (2019: 2), "driven by the 'data deluge', funder policy and open scholarship, libraries, in collaboration with other professional services and researchers, have developed a range of advisory and technical services to support Research Data Management".

According to Koltay (2017), the reason for bringing libraries and librarians on board in the processes of research projects that are data-intensive is to provide efficient help, especially in data management. He stresses the fact that "the goal of data literacy is similar as it intends ultimately to enable adequate research data management in a wide sense" (ibid: 9). The provision of data literacy training as a service provided the library is guaranteed to reshape the future role of libraries especially academic libraries (Merrill, 2011). Today, the library is taking on a more advanced role in the evolution of research data management activities. As introduced in this section, graduate students and faculty members are key members of the research community. They, therefore, deal with data in the entire process of research. According to Merrill (2011) researchers, today are dealing with datasets that have become easier to acquire while the tools to manipulate and present data have moved from the spreadsheet to the web browser.

In a study on "Investigation and analysis of research data services in university libraries", where a total of 50 libraries from, America, UK, Australia, the Netherlands, Canada, China and Sweden participated, libraries were found to be focal points in coordinating and disseminating data literacy among researchers (Si et al, 2015). The study, therefore, recommended that libraries' role in research should be strengthened (Si et al, 2015). Akers and Doty (2013) concur with this recommendation by asserting that academic libraries are the foundation of research support, especially data conservancy. They argue that the contemporary research world exposes researchers to unavoidable scientific data challenges. Therefore, researchers should have access to a central entity that allows better coordination of data activities.

A study by Guss (2016) acknowledges the place of the library in supporting researchers in confronting data-intensive research challenges. The study affirms that academic libraries have always dealt with the management of data and they are better placed as coordinating centres in universities. Koltay in his two different studies, that is Koltay (2016b) and Koltay (2017a), believes that academic libraries should provide information literacy education by supporting postgraduate students and individual faculty members to raise awareness on different issues in research. He points out intellectual property and copyright awareness as well as matters to do with privacy and confidentiality as issues to be addressed by academic libraries. However, he affirmatively states that even in the context of information literacy education, attention must also be given to data quality and data competencies.

Sewerin (2015)'s study acknowledges that today there is an explosion in the production of data during research which comes with a myriad of complexities. They raise poignant challenges including data management, data curation, data preservation as well as long-term storage of data. The study concludes that considering their previous experience with researchers' needs and practices, libraries have the ability and capacity to assist with these challenges.

The singling out of the library as a well-placed department in a university to spearhead data literacy is supported by previous practices of various academic libraries that have either implemented the program or are in the trial stages. Section 2.9 of this chapter highlights some of the libraries as case studies of data literacy programs.

2.9 Case studies of data literacy programmes by academic libraries

Various studies have previously been carried out to showcase efforts by some academic libraries in handling data-related activities in support of researchers in dealing with research data. For a library to offer quality research data services to its community of researchers, there is a need for a well-trained and knowledgeable community of librarians (Jones, Pryor & Whyte, 2013). It is assumed that their ability to deal with data will translate into transferring the same knowledge and skills to researchers. Two different studies were conducted by Tenopir et al. (2014) on "research data management services in academic libraries and perceptions of librarians on research data management". The two studies aimed to find out whether academic library directors and librarians in the United States of America (USA) and Canada had embraced research data services in their respective libraries. Findings from both studies indicated that there were notable discrepancies between Research Data Services (RDS) and some of the policies developed by some of the libraries. The findings revealed either there were

47

no RDS offered or provisions of RDS were at the developmental stage. Tenopir et al. (2014) study then recommended that there was a need to create awareness among the librarians especially the top management of the library. The studies also recommended that librarians be trained in order to enhance the development and implementation of RDS for the sake of serving researchers better.

While some studies have made a call to action to librarians to take have an active role in data science (Masinde, et al., 2021; Moore, Smith, Schultz-Jones & Marino, 2019; Martin, 2016), others have emphasised on the role of librarians in RDS. For instance, Carlson and Kneale's (2011) study looked at the role of an embedded librarian in research. The study notes and points out how important an embedded librarian is, and the critical role they play by journeying and giving support to researchers through RDS throughout the entire period of the research project.

Another study was conducted at Purdue University Libraries and the University of Illinois at Urbana-Champaign by Carlson et al. (2011). The study examined research data needs of postgraduate students. The study also was aimed at finding out the role played by university librarians in enabling data sharing and data curation. The study found out that by then, there were insufficiencies in data management. Majority of the respondents who were researchers revealed their lack of skills in metadata creation, data ethics, data preservations and basic database skills. According to the study's recommendations, there was a need to develop and implement a research data literacy training programme which would enhance the good management of data among researchers.

The few studies mentioned above, reveal a scenario of lack of data related services in selected libraries at the time. The studies indicate that, though there is a gradual increase in the awareness and implementation of open data and big data initiatives, existing gaps in universities indicate lack of data literacy programmes geared towards helping researchers in handling data. Therefore, as an academic unit in a university, the library's role in the research endeavours of an institution cannot be overlooked. And one way in which its impact can be felt is by implementing a comprehensive data literacy training program. Some academic libraries have attempted and developed data literacy instructional programmes. Majority of them are prototype programs to address researchers' data-related needs.

This study singles out seven institutions in North America and Europe that have in one way or another mooted data literacy programmes. The choice of these seven case studies is based on the institutions' training of researchers in data literacy while focusing on either postgraduate

students, faculty or both. Furthermore, the selection of the seven case studies is based on the collaboration between academic libraries or librarian and disciplinary faculty in developing and implementing a data literacy program in these institutions. According to the case studies, models of implementation vary from one institution to another. Some have embedded "data literacy" programmes in the already existing library training programmes, others are running it as an independent programme while others have embedded it within-subject curricular hence working in collaboration with different university faculties.

2.9.1 The Massachusetts Institute of Technology Libraries

The Massachusetts Institute of Technology (MIT) libraries help faculty and researchers from the institute to manage, store and share data produced (Carlson et al., 2015). The MIT library has created a subject guide manual titled *Manage Your Data*. The manual explains all the steps and procedures in the data lifecycle, that is from the scratch with different services like why manage the data, data management plan, file organisation, data security and backup (Tripathi, Shukla & Sonkar, 2017; Nayek & Sen, 2015).

The libraries develop, organise and offer seminar sessions such as Managing Research Data 101 around the year. The seminars are open to all researchers, especially postgraduate students and faculty members. The training program covers topics such as "documentation and metadata, data security and backups, directory structures and naming conventions, data sharing and citation, data integration". Researchers are also trained on how to convert data into file formats that allow for long-term access. Furthermore, participants are taken through some of the best practices for data retention and data archiving (MIT Libraries, n.d).

2.9.2 The University of Virginia Library

In 2007, the University of Virginia Libraries came up with the Scholar's Lab and Research Computing Lab as services geared towards supporting researchers who were dealing with large data sets. In 2010, the two entities were merged. While the library focused on providing reference and project-based services, the Lab focused on providing training on data management. The trainings which were conducted inform of seminars and workshops covered a variety of topics including the development of web application and text digitisation (Carlson et al., 2015).

However, the two also had collaborative projects that led to the development of a new service model. Other than respective traditional services provided by each, new services were introduced such as "data management and analysis, computational software support, and

49

knowledge of emerging technologies". Specific areas that were focused on included, research software support, data collaboration, and research communication (Carlson et al., 2015:16).

2.9.3 Syracuse University

Syracuse University developed the Science Data Literacy Project. The project was meant to train and empower postgraduate students with research skills in data collection, data processing, data management and evaluation. As researchers, postgraduate students were also to be trained on how to use data for scientific inquiry (Nayek & Sen, 2015: 40). This led to the development and implementation of a credit-bearing course, *Science Data Management*. The course introduced students to basics in scientific data description, how to manage data, data visualisation, and data curation (Carlson et al., 2015: 16).

2.9.4 Purdue University Libraries

Purdue University Libraries have been active in the development and provision of data services and trainings. Collaboration between the library and the faculty of Earth and Atmospheric Sciences department led to the development of a course titled Geoinformatics (Carlson et al., 2015). The course was designed for and targeted graduate students. The course aimed to provide a holistic approach to spatial data and training students on how to handle various activities that take place at every stage of the data lifecycle. Other areas of focus included teaching students and making them competent in areas such as data conversion, data manipulation, data analysis, data visualisation, metadata, and data re-sharing (Carlson et al., 2015).

2.9.5 Edinburgh and Minnesota University libraries

The EDINA Research Data Management Training (MANTRA), is an online course (Carlson, Johnston & Westra, 2015). It is specifically developed to help researchers who deal with digital data. The University of Edinburgh's data library and the University of Minnesota libraries jointly designed the course to provide data management course for structural engineers' students and faculty (Carlson, Johnston & Westra, 2015; Nayek & Sen 2015).

2.9.6 University of Toronto

Map and Data Library of the University of Toronto is a special library which mainly deals with maps and datasets. This library website provides a large variety of data on Canada and US by hosting itself and also providing other authentic sources of data like statistics Canada website, National Historic Geographic Information System (NHGIS) website among others (Nayek & Sen, 2015). It provides different data sets and services such as

© University of Pretoria

i. Statistics of census data, spatial data including maps, census of agriculture.

ii. How to cite the data and statistics, like citing maps and atlas, geospatial data, citing rules of machine-readable data, bibliographic citations for data files.

iii. Data visualisation tools, recommended sources, tutorials, visualisation principles...

iv. Guide of STATA which is a statistical package of data management (Nayek & Sen, 2015).

2.9.7 Bielefeld University

Wiljes and Cimiano work at Bielefeld University and since 2013, they taught research data management as an interdisciplinary course. The number of participants interested in the course has been increasing gradually from the time of its inception in 2013 when there were only 9 to 91 in 2018 (Wiljes & Cimiano, 2019). The methodology of the course is participatory as it follows a competency-based approach where practical exercises were encouraged. From the experience of the two instructors, the course incorporates has more of hands-on demonstrations in class. Class participation allows group discussions as well as individual presentations. These methodologies enable participants to connect the gained knowledge with their own research projects, ensuring their comprehension of how to utilize the acquired research data management skills. During the individual presentation sessions, participants are allowed to carry along their own laptops where they "test software for research data management (e.g., a Wiki, version control software, backup software, electronic lab notebooks)" (Wiljes & Cimiano, 2019: 3).

The course is divided into 15 topics. The topics cover different of practices in RDM and include, Introduction to research data management, Good scientific practices, Data, Information, Knowledge, Data backup, Data archiving, Organizing data; Documentation + Metadata, data publication and data sharing, Copyright law and licenses, Finding and re-using data, Sensitive data and privacy protection, Data management services at Bielefeld University, Tools 1: CMS + Wikis, Project management software, Tools 2: Cloud storage, Version Control Systems (Git), Tools 3: Electronic Lab Notebooks, Data management plans and Open Science (Wiljes & Cimiano, 2019).

The above-discussed projects indicate just a few of the academic libraries that have managed to integrate data services in their traditional services. As indicated, the majority of them focused on postgraduate students and faculty members, considering their involvement in research and their interaction with research data. Whether these programmes meet the threshold of data

51

literacy framework or standards is a matter that deserves another evaluation depending on a data literacy framework or standards and guidelines which unfortunately do not exist as for now. However, the initiatives, as documented, reveal

    i)   Various needs of researchers as far as data is concerned

    ii)   The diverse data related areas that need to be captured and addressed during the training

    iii)   The active role of the library and librarians in spearheading data training

    iv)   Avenues of collaboration between libraries or between libraries and faculties

The case studies also show the need to capture diverse aspects of data training, that is, from the most basic to the most advanced skills. For instance, some of the libraries have included training researchers on basic introduction to data management, to the use of sophisticated computer software in the management of data throughout its lifecycle. However, the above analysis only captures case studies from the global north, which shows considerable progress in data literacy initiatives. Section 2.10 will look at some of the data management related initiatives in Africa, especially within the academic context.

2.10   Data literacy initiatives in Africa

It would be too early to introduce the nomenclature "Data Literacy in Africa" as it appears to not yet have found space and attention among data or literacy-related studies on the continent. By the time of carrying out this study, there was no single study on data literacy in Africa according to some of the electronic research databases consulted by the researcher. The researcher did an online search in electronic databases such as ERIC, JSTOR and EMERALD using the search terms ((Data()literacy or data()training) and (Africa or South() Africa or Kenya or Tanzania or Uganda)). Unfortunately, no relevant results were retrieved. The same search was done on Google Scholar using the same search terms and no literature was found. The few studies that could be related to data literacy in Africa are those that highlight Research Data Management (RDM) services. It is on this basis that the researcher opted to search for information on "((Research Data Management or RDM) AND (Africa or South Africa) OR Kenya OR Tanzania OR Uganda)) to try and retrieve relevant information about strides made so far on the continent with regard to training and equipping researchers with 21st-century data skills.

According to Si et al. (2015), Research Data Management is the administration of data that is produced during a research project. Following data lifecycle, RDM involves controlling data

through various stages of the lifecycle. It is purposed to make the research process effective in order to avail research and meet requirements of either the research funder or university. RDM activities are aimed at facilitating quality research data by employing various techniques and strategies that would enable effective data storage, accessibility, security, future re-use of data and ethical use of data. Makani (2015:347) points out that researchers benefit a lot from the "effective management of research data, including data discovery, reuse, validation and verification". Patel (2016) concurs by adding that researchers benefit from effective research data management activities hence making most of their research acceptable.

The need to establish RDM services in universities today is becoming a need considering research being one of the key pillars besides teaching and learning. And with researchers (postgraduate students and faculty members) being encouraged to engage in research and publish in peer-reviewed journals, they come across research funders' requirements with which they need to meet before their work is published. In the recent past, due to funders' requirements, universities have scaled up their RDM services (Makani, 2015; Wilson & Jeffreys, 2013). Universities and researchers are therefore obliged to administer their data better or risk missing out on research funding (Makani, 2015). Researchers are therefore advised that for the purpose of a good scientific practice they need to practice a good Research Data Management (Wiljes & Cimiano, 2019). A data literacy program would, therefore, entail introducing and training researchers (faculty and postgraduate students) in RDM (Koltay, 2017a).

Several studies (such as the case studies highlighted in section 2.9) have shown that many countries in the global north have taken leadership in data literacy through Research Data Management activities aimed at enhancing researchers' data handling skills. These activities or initiatives are carried out either at a national level or institutional level or both. Some of the countries that have made positive progress towards RDM initiative include Australia, Germany, United Kingdom, Canada, Netherlands, New Zealand, Ireland and the USA (Cox et al., 2017; Henty, 2014; Lewis, 2010). According to Koltay (2016), the rise in RDM activities in these countries could be attributed to research funding requirements.

The few studies on RDM in Africa are predominantly either from South Africa or about RDM practices in South Africa. According to Kahn et al (2014:296), this is because "in South Africa, those involved in research are also increasingly aware of the importance and value of curating

53

and sharing the research data produced through public funding and RDM policies." Some of these studies will be reviewed in our subsequent discussions ahead.

Studies outside South Africa include a study by Avuglah and Underwood (2019) which assessed RDM capabilities at the University of Ghana (UG). There is also a study by Mushi, Pienaar and van Deventer (2020) whose aim was to identify and report on the implementation of RDM services at the University of Dodoma in Tanzania. The goal of the study was to enlighten researchers as well as the university management on the need for collaboration in the implementation of RDM services to realise the accessibility of data to the international community. Chigwada, Chiparausha and Kasiroori (2017)'s study also evaluated "how research data are being managed in research institutions in Zimbabwe". Findings from the study confirmed a lack of proper RDM as one of the main challenges as researchers were managing their own research data with no prior training.

Compared to other African countries, there are wider and consistent RDM services for researcher's advocacy in South Africa. The country's National Research Fund (NRF) which hosts the South Africa Data Archive (SADA) is the lead government agency that plays a pivotal role (Kahn et al., 2014). The levels of advocacy have created a positive awareness and capacity about RDM among researchers in South Africa (Kahn, 2014). Some of the initial developments towards embracing RDM are documented in van Deventer and Pienaar (2015) study. The study is their personal journey and campaign towards creating RDM awareness the realisation of entrenching RDM services in various institutions in South Africa. According to the study, their journey led to the establishment of the Network of Data and Information and Curation Communities in South Africa (NeDICC). One of the key roles of NeDICC has been to help in promoting "awareness/best practices relating to digital preservation, dissemination and use of research data" (van Deventer, & Pienaar, 2015:37). This is a key component in data literacy and awareness. The network has managed to organize "seminars, workshops and conferences" in order to help researchers and other stakeholders on the importance of RDM. van Wyk *et al*. (2017) concurs that NeDICC has also enhanced RDM initiatives in South Africa by engaging practitioners in the field.

South Africa has also seen the implementation of national data repositories such as the South African National Park (SANParks), the National Health Information Repository and Data Warehouse and the Data-Intensive Research Infrastructure for South Africa (DIRISA). Other than only handing big data, these repositories also do get involved in offering training to

researchers on how to handle data. As national repositories, they provide services and tools that researchers both in and outside the country could make use of while carrying out their research (Kahn et al., 2014). For instance, in May 2018, in conjunction with NRF, DIRISA developed and released DMP tool which was to be tested by researchers from around Africa. This was one of its kind pilots in Africa.

There has been a steady rise in developments in RDM in Higher Education Institutions (HEIs) considering the number of researchers in these institutions, the amount of data dealt with and funders' requirements they have to adhere to. Some Universities in South Africa, such as the University of Pretoria (UP), the University of South Africa (UNISA), the Witwatersrand University, and the University of Cape Town (UCT) have made strides in implementing RDM programs and are currently offering RDM services. They are reported to engage researchers through training and advocacy about RDM. The training opportunities which are offered with the purpose of empowering staff to support RDM are offered through workshops, in-house training, and knowledge sharing platforms, self-learning and networking. Most of these initiatives are housed and spearheaded by the respective universities' academic libraries.

For instance, the University of Pretoria (UP) in South Africa is providing RDM services to researchers through the library. The implementation of the service has undergone a steady chronological development since 2007 when the preservation and retention of researcher data policy was developed (van Wyk, 2017). By 2016 the university had developed and approved a formal RDM policy (van Wyk, 2017). Some of the other universities offering RDM programs where they train researchers in RDM and offer RDM related services include the University of South Africa, Stellenbosch University, and the University of Cape Town.

There have also been developments of education curricular programs in various universities in South Africa aimed at enhancing data skills. The University of Pretoria has introduced a Master's degree in Information Technology in Big Data Science. The University of Witwatersrand, Johannesburg offers a BSc Honours in Big Data Analytics. The University of Cape Town is also offering a Masters in Data Curation. All these programs might not be directly related to training researchers in data literacy but they are offering some solutions to data-related challenges that researchers encounter today. They respond to the emerging data needs of the society where there is a high demand for skilled data scientists. Other efforts aimed towards data literacy by the University of Pretoria include the establishment of a Carnegie funded Master degree programme in Information Technology (M.IT) and Continuing

Professional Development (CPD) which integrate RDM training. The two programs were aimed at training librarians in Africa and equipping them with IT skills.

Further steps have been taken on the continent to increase research data literacy among researchers in Africa. For instance, in 2018 CODATA, ICTP and EAIFR organised a 10-day workshop at the University of Rwanda, Kigali. The goal of the workshop was to train researchers in Research Data Science (RDS). Participants were introduced to "principles and practice of Open Science and research data management and curation, the use of a range of data platforms and infrastructures, large scale analysis, statistics, visualization and modelling techniques, software development and data annotation". RDM, data visualization and introduction to R are among the topics that were covered during the workshop.

Similarly, the CODATA-RDA Research Data Summer School held at the University of Pretoria took place from 13 to 24 January 2020. While covering topics such as (but not limited to) RDM, data visualization, Open and Collaborative Research and Information Security, the summer school aimed to provide early career researchers with foundational data science skills "to enable them to work with their data in an effective and efficient manner" as dictated by the 21st-century research needs. The summer school covered notable areas considered to be essential to a 21st-century researcher. Participants had the opportunity to experience a practical introduction to the topics covered during the sessions with some theory accompanied by extensive hands-on training.

In Kenya, it could be said that not much has taken place as far as data literacy or implementation of RDM is concerned considering the lack of studies on RDM and its implementation. In a recent study, Adika and Kwanya (2020) analysed the required skills by faculty at Strathmore University in order to support RDM. The study also assessed the level of RDM literacy among faculty at Strathmore University. According to the finding from the study, faculty members had some applicable skills in most of the RDM capacity areas. The respondents indicated that they were "able to plan for, search, find, organise, store, secure and share research data competently" (Adika and Kwanya, 2020:15). However, the study unveiled some existing gaps in RDM literacy levels among the lecturers. Majority of them were lacking skills in data sharing, data security and data legislation. The study recommended RDM training as a way of mitigating the existing challenge.

Some of the previous studies in RDM in Kenya are by Jao et al. (2015); Olum (2013) and Family Health International- Kenya (2005). The studies were aimed at promoting RDM services and activities in the health and migration sectors. Findings from another study by Ng'eno and Mutula (2018: 42) which was a literature review analysis of RDM in agricultural research institutes showed that there was "lack of coordinated RDM strategies in the research institutes" hence leading to "loss of data and difficulty in accessing, reusing and sharing research data". The investigation also revealed a lack of knowledge and skills related to research data management inside agricultural research institutes, limiting the use of research data in Kenya's agricultural research institutions. Findings from Ng'eno and Mutula's (2018) study implicitly point to gaps in data literacy knowledge being raised in this study. In view of the identified challenges, researchers need adequate training in RDM in order to be data literate (Borgman, 2012).

Analysis of data literacy initiatives in Africa reveals two key points. One, nearly all African countries, except for South Africa are still lagging in the area of data literacy or RDM. This is quite unfortunate considering that this is the 21st century and data is quickly becoming the new currency as seen before. Majority of the countries have no national research data repositories. Secondly, all is not lost as other African countries could learn from initiatives in South Africa. Universities in Africa have a chance to initiate data literacy trainings for their researchers by tapping into the experience and expertise of their academic libraries and librarians. As Universities ponder on the next move, their general emerging practices in the realm of data literacy. These practices could act as guiding principles for universities as they plan to develop and implement data literacy training programmes. The following section (2.11) reviews some of the emerging practices, especially in academia.

2.11   Data literacy: emerging practices in academia

Though data literacy is still an embryonic service in many universities, it is emerging that it is an important service necessary for researchers in the 21st Century. Some best practices can be pointed out from the little that is currently being offered as mentioned in preceding sections of this study. The practices embody an extensive collection of evidence-based suggestions for institutions seeking to establish data literacy programs for researchers (Wanner, 2015).

2.11.1 The role of the library in implementing data literacy

Anchoring data literacy in the library in a university setting seems to be a common practice (Al-Jaradat, 2021). What stands out is that, other than providing access to information resource

and sources, the library plays a key and practical role in providing research services. These services include research data management services (Briney, Coates & Goben, 2020; Brown, et al, 2018; Delaney & Bates, 2018; Cox et al., 2017).

Hamblin (2005) carried out a study to find out some of the research-related activities listed on various library portals. According to the findings many of the selected libraries were providing separate space for researchers, developed research collections, offered training in database search to researchers, managed the submission of electronic theses and managed institutional repositories, had set up communication and collaboration tools for research communities, provided editing services and research data management instruction services. They also had a research librarian who kept contact with researchers especially postgraduate students (Hamblin 2005).

In the 21st century, academic libraries face a significant strategic challenge when it comes to handling data. Their pivotal role in supporting research is more relevant than ever and should not be underestimated. Academic libraries are increasingly finding new opportunities to contribute, such as raising researchers' awareness about data, establishing archival and preservation services for institutional data through repositories, and creating a distinct professional field known as data librarianship (Frederick, 2019; Brown, et al, 2018; Friedlander & Adler, 2006).

The provision of Research Data Services by libraries takes many facets as shown by some studies (Borghi & van Gulick, 2022; Chawinga & Zinn, 2021; Bunkar & Bhatt, 2020; Koltay, 2016). The provision of this service to researchers varies from library to library and may include services such offering research data management training, management and administration of institutional repositories, helping researchers with the creation of data management plans. Some libraries also offer assistance to researchers by providing data mining tools, data visualisation tools and creation of metadata. Other services include helping researchers to understand the intellectual property issues related to research data (Koltay, 2016; Flores et al., 2015; Tenopir et al., 2015b; Linde et al., 2014).

Undoubtedly, data-intensive research has a global visible impact on library service and practice. Considering the changing trends, libraries have the responsibility to take up more notable initiatives that would cement their role in the research data space. While outlining new areas of involvement for libraries such as "policy formulation, repository development,

curriculum innovation, professional updating, postgraduate training, resource selection", Corrall (2012:19) points out that, libraries have to play a role in providing research advice, data advocacy and curation profiling.

2.11.2 Data librarianship

The provision of research data services in libraries requires skilled professionals (Fuhr, 2022; Tenopir et al, 2017; Brodsky, 2016). With the evolution and adoption of ICT in the library, coupled with the ever-changing needs of library users, the role of traditional librarianship has been transformed. In the era of big data and open data, librarians are called to play a vital role in the research life of their users. In this context, they could play a frontline role in spearheading data literacy programmes by training students and faculty members, due to the magnitude and complexity of research projects they are working on (Simons & Searle, 2014; Corrall, 2012).

Carlson and Kneale (2011) carried out a study on the role of embedded librarianship in the context of research. According to the findings of the study, faculty members who participated insisted that students were lacking data literacy skills hence the need for an instructional data literacy programme especially for students who were carrying out research. However, even with the admission that there is need for data literacy instruction programs for students, the faculty members were reluctant to take up the responsibility of teaching data management skills to their students hence preferred delegating the duty to librarians. What this means is that a 21st Century librarian must be different from a "traditional" librarian. As an embedded librarian, (Andrews, 2015) they should be able to take up such roles and clearly indicate which skills should be included in a data literacy programme for researchers in order to remedy the situation.

In the same context, Koltay's (2015:6) study, "Data literacy for researchers and data librarians", adds that the new role of the librarian in offering data literacy services is an evolving one considering the skills needed by students and researchers. He calls for the development of effective instructional content, methods, and formats that would impart data competencies. It is therefore clear that data literacy services offer librarians a chance to transform their influence in research especially in academic institutions. This view is echoed by Schultz-Jones, Moore and Marino, (2019); Semeler, Pinto and Rozados (2019) and Sewell and Kingsley (2017) who emphasize the need for librarians to undergo training to enhance their research data literacy skills, given the increasing demand for research data services from students and researchers.

2.11.3 Collaboration between the library and academic departments

In the words of Pham and Tanner (2015:3), collaboration in the education context is described as "a joint working, learning and sharing process that specifically focuses on the activities of teaching, learning and researching among educational participants, in which knowledge can be activated and transferred". The library and the teaching fraternity have various levels of collaborations in a learning institution. In a study on how academics and library staff can collaborate in order to enhance services, Pham and Tanner (2015:3) highlight specific areas of collaboration which include "building subject collections, organising transition programs, developing learning skills for students, embedding information literacy skills and research skills into the curriculum, teaching and tutoring or working on research projects". For example, a working partnership between the University library of South Florida St. Petersburg (USFSP) and the faculty members has achieved in delivering "quality information literacy programs". Furthermore, the library at USFSP has a program aimed at assisting faculty in RDM (Burress, Mann & Neville, 2020:1).

The need to develop or create a data-literate community of researchers presents another opportunity of collaboration between the library and other departments. From the case studies presented, most of the initiatives reviewed involve a level of collaboration between the university library and the faculty. Such collaborations give way to the development of data literacy modules that are tailored towards the students' or/and faculty members' research needs. In this line, Carlson and Kneale's article concludes, that such collaborations between faculty and librarians are highly recommended: "as the best practice for teaching data information literacy skills" (Carlson & Kneale, 2011:653).

It is therefore worth noting that to establish a collaboration between the library and academia, in the realm of data literacy services, and fostering a dynamic partnership is crucial. This collaboration should extend beyond the traditional boundaries, involving joint efforts in curriculum development, research projects, and the integration of information literacy skills.

2.11.4 Embedding data literacy instruction into the curriculum

Embedding Information literacy in the universities' curriculum has been hailed as one of the reasons for its success and wide adoption. At the University of Bedfordshire, a collaboration between the faculty librarian and the Department of Psychology led to the embedding of information literacy into the Psychology curriculum during the years 2007-2008. The use of blended learning during the program and incorporation of a variety of teaching and assessment

60

methods were unutilised (Robertson et al., 2012). A study by Azubogu and Madu (2019) advocates for the integration of information literacy training into the curriculum in Nigerian.

Borrowing from embedding practices of Information literacy across the curriculum, data literacy can be integrated into a variety of courses and disciplines (Calzada & Marzal, 2013). One of the emerging themes around data literacy as documented in the case studies above is the need to embed data literacy instructions into the curriculum. Some of the proponents who advocate for embedding data literacy instructions into the curriculum go further in arguing that leaners at all level of learning need to be introduced to data literacy. That there is a need to expose data literacy concepts to all groups of learners at an early age (Wanner, 2015; Derakhshan & Singh, 2011; Proctor, Wartho & Anderson, 2005). They add that, though libraries play a key role by organising training schedules, the library cannot exhaust all the areas of concern. They insist that library time is not sufficient to take a learner through a comprehensive understanding of RDM.

According to Stephenson and Caravello (2007: 527), incorporating data literacy training across the curriculum increases chances of comprehensively attaining data literacy competencies. They add that this could be done in various ways such as "incorporating individual consultations, class orientations, reference desk instruction, incorporating data literacy activities into other library instruction initiatives, using hands-on activities and so forth". Furthermore, the use of creative, practical pedagogical techniques as well as case studies could be applied.

Another advantage of introducing continuing data literacy training in all stages of learning according to Stephenson and Caravallo (2007), is to develop skills for ongoing research in academia and workplace for lifelong learning. They insist that "if students are expected to use data skills as they move into a higher level, evaluative, and interpretive research and coursework, they will need multiple opportunities in which to develop data literacy" (ibid: 535).

2.11.5 Data literacy for lifelong learning

The concept of lifelong learning is more associated with information literacy than data literacy considering the available literature (Solmaz, 2017; Lau, 2006; Serap, 2003; Candy, 2002). However, closely related is the *data-informed learning* concept. Maybee and Zilinski, (2015:3), describe data-informed learning as "an approach to data literacy designed to address shortcomings associated with generic skill-based models by having students learn to use data

within the context of disciplinary learning". The concept envisions embedding data learning in educational programmes at all levels of schooling with the hope of equipping learners with lifelong skills. The concept is closely related to data literacy considering its approach to data literacy and lifelong learning goals. It is quite apparent that the need for one to be data literate should not just be for the moment in time but rather for posterity. Maybee and Zilinski, (2015:3) add that data literacy and data-informed learning "support lifelong learning using data in the context of learning and prepare learners to use data in their professional and personal settings".

Data skills are meant to help one be competent in handling data-related assignment both in class and outside the class. According to Stephenson and Caravallo (2007), data literacy skills are transferable right from classrooms and to other spheres of life like the workplace.

The concept of data literacy for lifelong learning is undoubtedly related to the previous discussion on public data literacy. While the former focuses on the development of skills and competencies required for individuals to effectively engage with data throughout their educational and professional journeys, the latter is concerned with the general public's understanding and use of data. Despite differing emphasis, both concepts share the main goal of providing individuals with the necessary skills to access, comprehend, and critically evaluate data in a variety of circumstances. Maybee and Zilinski (2015) define data-informed learning as a paradigm for integrating data literacy into educational programs, promoting lifetime skill acquisition and application. As a result, while data literacy for lifelong learning may have a greater educational focus, its alignment with the larger goals of public data literacy emphasizes the interconnection and mutual reinforcement of these concepts in promoting data literacy throughout various spheres of life.

2.12 Possible challenges in the implementation of data literacy programmes

As indicated before, data literacy is a fairly new area of both practice and study. Not so much has been studied regarding challenges related to its implementation. However, challenges associated with research data management could be closely related. Some of these challenges include; RDM skills gap, fiscal support and inadequate infrastructures (Xu, et al., 2022).

In a study, Ridsdale et al (2015:19) point out some barriers and challenges that data literacy programmes would face. The study points out that there is a misconception that people born post-1983 are technologically savvy hence they need no training in data competencies. That they "have inherent technological skills and abilities" (ibid: 13). This kind of misconceptions

leads to overlooking of data needs of researchers hence no programme or initiative is developed towards helping them.

The report also raises the challenge of introducing data literacy at an advanced academic stage. That "it is difficult, to begin with, the basics with a professional audience with varying skill levels" (Ridsdale et al., 2015:19).

Qi's (2018) study alludes to the fact that despite the awareness of the need for data literacy, university libraries face some considerable challenges. The study notes that there is a lack of comprehensive theoretical framework designing specific paths for a data literacy programme. Even though some libraries are currently carryout data literacy programmes, mostly in the form of RDS, there is still a lack of a unifying framework providing standards and guidelines. Qi (2018) states that there are no consistent standards and corresponding rules in the industry. Typically, each library sets its own regulations and segregated systems, making significant and effective external collaboration and interchange difficult.

According to the study, data literacy implementation is suffering from insufficient understanding of its importance. There is a lack of "top-level design and policy support" especially from top-level management. Meaning that the organisational structure, be it in libraries or parent institutions have not yet had a buy-in into supporting data literacy training initiatives, Qi's (2018).

The lack of a suitable framework is a key barrier to the implementation of data literacy initiatives, limiting their efficacy and cohesiveness. While some libraries or institutions as captured in sections 2.9 and 2.10, are actively trying to implement data literacy programs, the lack of consistent standards and rules is a challenge to seamless collaboration and exchange among institutions. Addressing this framework weakness is critical to improving the effectiveness and sustainability of data literacy initiatives in academic settings.

Even though some institutions are currently conducting data literacy training, the problem is that they have not done a needs analysis to find exactly what the actual needs of the users are. Furthermore, they have not carried out an investigation to find out the possibility of inter-disciplinary or inter-institutional collaboration while doing data literacy training. And lastly, some of the libraries have not found out whether they have the right specialised personnel who could help in the training Qi's (2018).

2.13 Summary

As the world embraces the use of technologies, while world nations adopt knowledge economy, data literacy will remain to be one of the most significant literacies for citizens at large. The demand for data literacy programs or initiatives might not be expressed explicitly, especially in the global South. But as research funders requirements continue to be implemented, the volume of data continues to increase and technologies continue to evolve, there is no doubt that data literacy programs will have to become a normative part of scholarship.

Chapter 2 sought to explore the role of university libraries in data literacy. The chapter paints the picture of the researcher's needs especially in dealing with data while considering their various needs. The chapter starts by introducing the concept of data literacy as well as its development due to the rise of open data and big data, and how it is linked to other literacies. With limited literature addressing specifically the data literacy concept especially its implementation in universities as a service for researchers, the researcher consulted literature that addresses the library's role in providing Research Data Management services. The literature sought was also that which addressed the library's role in supporting various aspects of the research lifecycle with a special focus on data. The study reviewed literature on the support provided while adhering to the UK research data lifecycle model.

The literature reviewed also discussed the role of libraries and librarians' competencies about supporting data literacy; the vital role of university libraries and librarians in RDM, and the integration of technologies in the research lifecycle. The literature revealed that librarians globally are making effort to encourage data literacy but there are barely structured strategies or framework to guide effective implementation of data literacy. The most extant literature on data literacy was found in global north regions including the USA and Europe. There are limited resources addressing data literacy research data management in the African context.

In the analysis of the available literature, the researcher identified some gaps, which included, limited literature specifically addressing the concept and implementation of data literacy programs within university libraries, especially in the context of the global South and specifically in Kenya; a lack of structured strategies or frameworks guiding the effective implementation of data literacy initiatives within university libraries, particularly in African countries. The researcher also discovered that there is a prevalence of literature from global North regions, such as the United States and Europe, with few resources addressing data literacy and research data management in African contexts, including Kenya.This study

therefore, endeavoured to contribute to filling these gaps by exploring and providing evidence-based study for data literacy in Kenya. The next chapter presents the research methodology.

CHAPTER 3

3. THEORETICAL FRAMEWORKS

Chapter 2 presented relevant literature for the study. The focus here shifts to relevant frameworks for the study.

3.1 Introduction

In Chapter 3, the researcher presented theoretical frameworks for the study which served as the structure and support for the study (Grant & Osanloo, 2014) and pillars for the development of a foreseen data literacy framework. The chapter consists of concepts that helped in understanding the specific phenomenon in the study. The researcher presented the background for frameworks and models that were applied to the study before developing an initial conceptual framework for the study.

3.2 Background for frameworks and models

One of the objectives of this study was to develop a data literacy framework that provides a basic structure for the implementation of data literacy programmes in universities in Kenya. This section provides a general understanding of what frameworks and models are in relation to research. In the analysis, the researcher particularly focused on what theoretical and conceptual frameworks are and why they are considered as an important component in research before delving into data literacy-related frameworks.

In research, the use of theoretical and conceptual frameworks is prevalent as researchers try to relate or align their studies with theories that exist. In a single study, a researcher could choose to either use one of the frameworks or both (theoretical and conceptual). The use of theoretical and conceptual frameworks in research aims to explain the research's route while keeping it theoretically grounded. Using the two frameworks has the overall goal of making research findings more relevant, acceptable to theoretical constructs in the research field, and ensuring generalisability. A distinction between the two frameworks and how they are used or incorporated in research is provided in the following sections.

A theoretical framework that Grant and Osanloo (2014) describe as a "blueprint" or guide for research is an explanatory system or a perspective that frames a study. It is what connects a study to larger ideas beyond the particular. In research, a theoretical framework "serves as the structure and support for the rationale of the study, the problem statement, the purpose, the significance, and the research questions" (Grant & Osanloo, 2014:12). For a researcher

66

carrying out a study, it is a measure of the relationship and impact of their study to the specific context in which the study is to be applied. As Varpio et al. (2020) put it, it consists of concepts that help in understanding the specific phenomenon in research. According to Ravitch and Riggan (2017:13), one of the main reasons why researchers incorporate the use of a theoretical framework in their studies is to "confirm a gap in knowledge and to provide a justification for conducting a study".

The use or application of a theoretical framework has been deemed to be of particular importance especially to a study's literature review. A literature review of a study, for instance, consists of different yet interrelated parts considering the topic in discussion. In this case, according to Grant and Osanloo (2014), a theoretical framework helps in developing coherence among the various studies under review. Furthermore, a theoretical framework in studies is used as a foundation on which the constructs of a study as captured in the research questions, can be defined and located in the literature review. Thus, the researcher's views of the concepts under study can be established, explained, and justified (Varpio et al., 2020). Another reason why the use of a theoretical framework in a study is deemed important is that it sets the parameters within which literature can be reviewed. According to Kumar (2014), lack of a theoretical framework could lead to a researcher losing focus, getting overwhelmed with unnecessary reading and note-taking which could be irrelevant to the study.

The conceptual framework also known as a conceptual model, in contrast to the theoretical framework, is a structure or model constructed by the researcher in an attempt to elucidate the natural course of the phenomenon under investigation. It is a logical structure of connected concepts that helps provide a picture or visual representation of how ideas in a study relate to one another (Grant & Osanloo, 2014). It is the researcher's opinion on how the research problem should be investigated, based on the literature review and theoretical frameworks analysed in the study. According to Kivunja (2018), a conceptual framework is the researcher's "logical conceptualization" of their entire study project. In the conceptualisation, the researcher clearly defines the relationships between different variables in the study (Ravitch & Riggan, 2017).

The process of arriving at a conceptual framework is akin to an inductive process whereby small individual pieces (in this case, concepts) are joined together to tell a bigger map of possible relationships. Thus, a conceptual framework is derived from concepts, in so far as a

67

theoretical framework is derived from a theory. Schematically, this may be represented diagrammatically depending on the choice of the researcher.

In this study, the researcher will use both theoretical and conceptual frameworks. The use of the two frameworks will eventually be presented diagrammatically towards the end of the chapter as the initial framework.

3.3   Theoretical frameworks for data literacy

There is a steady increase in the realisation of data (re)use in higher education as an important research output with the emphasis being put on managing data for future use. Data re(use) entails some different activities, which include "returning to one's own data for later comparisons, acquiring datasets from public or private sources to compare to newly collected data, surveying available datasets as background research for a new project, or conducting reanalyses of one or more datasets to address new research questions" (Pasquetto, Borgman & Wofford, 2019:4). Similarly, due to interests in data, there is a general call for the development of basic data literacy competencies among researchers especially those within academic institutions (Maybee & Zilinski, 2015). The magnitude of data and the role data plays today in almost every aspect of life including decision making, demand that 21st Century global citizens should acquire or possess some data-relevant skills such as gathering, storing, processing and visualising data (Corrall, 2019c). An academic environment (such as a university) is seen as an ideal place to champion the development of data literacy skills. Entrenching data literacy programmes in the curricula at the university level is therefore considered to be a step in the right direction (Koltay, 2017; Carlson & Bracke, 2015), as that would target researchers (post graduate students and faculty members) who intend to impact the society through their research engagements.

To argue and justify the need for the implementation of data literacy programmes in institutions of higher learning, the researcher will present some theories and models which will help in guiding an investigation into data literacy among researchers. These theories and models will form the pillar upon which a data literacy framework will be built. The theories are the Radical Change theory and the Stakeholder theory while the models will include the Intellectus model and the Bielefeld University RDM training model. The Radical change theory is adopted in this study as it explains the changing environment that dictates the need for data literacy among researchers (Dresang & Kotrla, 2009). It frames the changing context for data literacy development. The Stakeholder theory and the Intellectus Model are socioeconomic theories,

68

concerned with the relationship between social and economic factors within society. Their use in this study is justified by the emerging trends in social and behavioural sciences where there is a move towards understanding research problems from an interdisciplinary perspective within the context of society (Corrall, 2019a). And lastly, there is the Bielefeld University RDM training model which is adapted from the training programme implemented by a German university having identified the data literacy needs of researchers.

3.3.1   The Radical change theory

The initial development of the Radical Change theory is associated with Eliza T. Dresang dating back to the mid-1990s. However, the first link between Radical Change theory and data literacy was made by Professor Sheila Corrall from the University of Pittsburgh (Corrall, 2019a). The main reason for the development of the Radical change theory was to try and explain how the changing digital age affects information resources and information behaviours among the youths while applying the digital age principles (Dresang, 2005). Its development and application as a digital theory was aimed at demonstrating changes in the information world. This is well illustrated in the book, Radical Change: Books for Youth in a Digital Age, by Dresang (1999). Koh (2015:2) notes that since its inception, the radical change theory has "influenced numerous researchers and professionals by equipping them to understand and better serve digital youth". Koh (2015) singles studies by Pantaleo (2008) and Hassett (2005) which are some of the earlier studies that applied the radical change theory. There is also a later study by Pantaleo (2017) that applies the radical theory framework.

The Radical Change theory acknowledges the digital age's "radical changes" in terms of forms and formats as well as perspectives, and boundaries in the youths' literature and their impact in terms of interactivity, connectivity, and access to information. In summary, the Radical Change theory "serves as a lens through which to examine, explain, and use contemporary literature for youth growing up in the Digital Age" (Koh, 2015:3).

Even though the Radical Change theory was developed purposely to explain some of the changes that were occurring in some of the literature meant for the young people navigating the technological evolutions affecting writing and authorship, it has been impactful in other areas too. The Radical change theory has been adopted as an explanatory framework in various contexts such as in explaining changes experienced in information behaviour and resources (Knox, 2014; Rubel, 2014) as well as trying to make clear the place of "intellectual freedom in librarianship" (Childs, 2017).

69

Essentially, the Radical Change theory is grounded on three key digital age principles; interactivity, connectivity, and access. According to Dresang (2005), the purpose of the use of the three principles is to elucidate changes experienced in information resources for young people and their information behaviour. Dresang's (2006) digital age principles can be applied to explain aspects such as data literacy in the current information-filled society.

By coming up with the Radical Change theory, Dresang and McClelland (1999) had identified and acknowledged that some radical changes had and were still taking place with regard to young people's information behaviour in the digital world. The ICTs evolution and their rapid adoption were impacting young people's information-seeking behaviour. They pointed out what they considered as changes that were evident in youth literature: "changing forms and formats, changing perspectives, and changing boundaries," (Dresang & McClelland, 1999:161). The three types of changes correspond to what was developed as the three principles of the digital age by Elizabeth Dressang, that is "interactivity, connectivity, and access" (Dresang 1999).

The general schema of this theory offers a favourable view for investigating some of the radical changes in the data world. This is what Corrall (2019a) identifies as "human-data interaction and data literacy" among researchers in the complex pluralist context. The context is described in more detail below.

3.3.1.1 The complex pluralist context of data

The evolution of Information Communication Technologies (ICTs) has brought with it various radical changes that have impacted human interaction with data. ICTs have, for instance, inspired the development and growth of globalized networks of trade, communication and transport. With ICTs, there is movement or transfer of information from one end to another which has become a fundamental component in the socio-economic setup of any society today. According to Ridsdale (2015: 8), there is an increase in "collection, analysis, sharing, and use of data" as a result of ever-emerging technologies. ICTs have consequently led to the data revolution that researchers are experiencing today.

As pointed out in Chapter 1 of this study, the data revolution is real. The 21st Century researchers are experiencing huge amounts (large quantities) of data in their various engagements with/in different studies. The reality is parallel for organizations, institutions, companies and businesses which are depending on data in decision making in relation to their strategic and operational directions. In their encounter with data, they find themselves dealing

with increasing volumes of data hence demanding critical analysis. Daniel (2015:907), notes that "the nature of data available is changing, and the changes bring with them complexity in managing the volumes and analysis of these data".

According to Pentland (2013) and Patil and Mason (2015), we live in a data-driven society considering the fundamental increase in data as compared to the previous years. The scientific world has witnessed the rise of Open Data and Big Data hence leading to a fundamental change in how we perceive, work with and consume data today (Monino, 2021; Berman, 2017; Blackburn, Alexander, Legan & Klabjan, 2017; Connelly et al., 2016). The research community in one way or another is getting influenced by the emerging world of open and big data. These two phenomena (Open Data and Big Data) offer both challenges and opportunities to researchers ranging from workforce training issues, ethics and privacy concerns as well as new chances for public-private partnerships (Acharjya & Ahmed, 2016). However, the big challenge is the competency of a researcher when handling and dealing with data (Berman, 2017).

The 21st Century researcher is struggling to deal with huge volumes of data and other related requirements (Younas, 2019; Kvale & Stangeland, 2017). According to Kvale and Stangeland (2017:728), "the amount of digital data currently produced in research, combined with incentives for open science" raise new challenges for researchers today. They add that requirements for data management plans by funders as well as journals and publishing platforms demand that 21st-century research should be data literate. Ridsdale, et al. (2015:8) believe that there is a need to shift from just being a "data-rich to an information-rich and knowledge-rich" society considering the described context. The transition demands data literate researchers and citizens in order to cope with the "radically" changing research environment.

3.3.1.2 Radical change theory and data literacy

As mentioned earlier in section 3.2.1, the first person to use or link the Radical change theory to data literacy is Professor Sheila Corrall from the University of Pittsburgh (Corrall, 2019a; Corrall, 2019b). Corral has written and presented papers including hosting roundtable discussions on the need to develop data literacy competencies among the community of researchers as well as the society in general (Corrall, 2019a; Corrall, 2019b; Corrall, 2019c; Corrall, 2019d). According to Corrall (2019a), the Radical Change theory provides a suitable framework to describe, explain and analyse what she terms as the "complex, pluralist context

for data literacy development". She cites studies such as Critchlow and van Dam (2013); Hey, Gannon and Pinkelman, (2012); Bell, Hey, and Szalay, (2009) which highlight how the ever-increasing volumes of data have led to what is considered to be the "fourth research paradigm" as evidenced by data-intensive research.

Critchlow and van Dam (2013) in their book Data-Intensive Science, look at the impact of data-intensive science on research and describe what they consider as upcoming tools that will help scientists make breakthroughs in the future. They examine some of the best practices and gaps in these techniques for addressing issues that data-intensive science faces. Hey, Gannon and Pinkelman, (2012) point out in their work how data-intensive science is shifting the boundaries of research in different ways. On the other hand, Bell, Hey and Szalay, (2009) in their study acknowledge how intensive data science will be to the endeavours of many researchers or scientists. However, they warn that this comes with the demand for some specialized skills that each researcher must be ready to attain in order to maximise opportunities that come with it. They conclude their study by recommending that "making optimum use of such services will require some radical rethinking in the research community" (Bell, Hey & Szalay, 2009:1298).

In one of her presentations, Corrall (2019a) points out how the data deluge has resulted in the fourth industrial revolution due to the explosion of data. She is of the view that data explosion has impacted every stratum of the society including businesses, governments and the rest of the society's sectors as far as decision making is concerned. Fundamentally, thanks to the radical changes, data has become a valuable asset for organisations as well as individuals (Corrall, 2019a). This view is augmented by Wiljes and Cimiano (2019:1) as they highlight that "with the advent of digital data, computer-based analysis and the internet, the way scientific work is done has been revolutionised". Therefore, they advise that to be at the same speed with the "rapid transformation, the world needs competent and informed researchers who can apply new methods of research data management effectively and responsibly," (Wiljes & Cimiano, 2019:1)

In Corall's analysis of the Radical Change theory in relation to data, there are three digital age principles (interactivity, connectivity and accessibility) which, according to her, form the "complex pluralist environmental context for data literacy development". The complex environment is a result of the evolution of ICTs which have radically changed interaction with data. ICTs, for instance, allow for data interactivity which has influenced data presentation and how users interact with the presented data. Data interactivity is associated with data

visualisation which allows users to manipulate some aspects of the data set or the presentation of the results allowing users to gain insight from the data at the level they require (Park et al., 2022). Data interactivity gives the user some level of control over the presented data and an opportunity to get answers about the different questions they may have about the data.

The evolution of ICTs has also led to data connectivity which is another digital age principle. The connection of diverse data sets and applications, including data from distinct identity spaces, is known as data connectivity. This allows multiple parties to collaborate with data controls, assuring safe and effective activation across the ecosystem. Data connectivity encourages more connection and collaboration between researchers and better research practices that lead to new findings and better decision-making. In a time of reduced monetary investment for science and research (Tchamyou, 2019), data connectivity that allows sharing among researchers turns out to be a better way for researchers to share resources. This is a radical change in research today as it allows researchers to build upon the work of others rather than repeat already existing research. Similarly, data connectivity allows researchers to share data, allowing them to conduct meta-analyses on the present research issue. Meta-analyses are useful for identifying greater trends across topical ranges. This ensures the continued production of these types of analyses. Data connectivity has been made possible today with the evolution of ICTs.

Furthermore, the adoption of ICTs in research has resulted in a radical change in data accessibility hence impacting researchers' interaction with data. Data accessibility denotes removing barriers that inhibit access to data and fully leveraging the data contained in different databases today. Modern research technologies have greatly increased the amount of scientific data being generated. Data accessibility is important at every level of the research process because it allows researchers to pool data and conduct analyses with more confidence in their findings. The more data a researcher has, the more statistical power they have to back up their findings and anticipate data quality concerns.

According to Corrall's (2019a) analysis, there are four radical changes that bring about the complex pluralist context of data that researchers experience today. These changes are related to the evolution and development of ICTs, the datafication process, the penetration and influence of the data revolution and the changing role of librarians in supporting research.

The evolution and development of ICTs signified or ushered in the first radical change. Technologies have enabled new ways of handling large sets of data, accessing data, archiving

73

data and even sharing data. The new technologies have radically disrupted the traditional way of handling and managing data. In a study aimed at establishing some of the determinants of data sharing, Sayogo and Pardo, (2013) established that there is a lack of data management skills as one of the key determinants in the management of data. The study linked a lack of data management skills to the robust advancement in ICTs that they considered as evolving at a high speed. According to the study, "advances in computing, information and communication technologies produce" have had a significant impact on scientific research" hence there is a need to empower researchers with necessary skills in order to cope with the changing environment in research. Evolving technologies, as Tenopir et al (2011:9) point out, have made research to be "increasingly data-intensive and collaborative". The same point of view is emphasised by Muriithi, Horner and Pemberton (2016). Additionally, the new ICT evolutions and capacities have introduced new technologies in research that aid in the manipulation and exploration of large volumes of data (Park et al., 2021).

The second radical change is that data has tilted the status quo in research, higher education, and society, hence making them part of or beneficiaries of a datafication process (Corrall, 2019c). Datafication is a buzzword which, according to Mayer-Schoenberger and Cukier (2013:56), "is the transformation of social action into online quantified data, thus allowing for real-time tracking and predictive analysis". van Dijck (2014:198) adds that it involves the transformation of previously unobservable processes or activities into data that can be observed, traced, examined, and improved. In the wake of datafication, not just as a radical concept but also a radical practice, data literacy becomes a fundamental requirement for both researchers and anyone else who would like to make sense out of data. As noted by Corrall (2019), it is also required for effective participation in citizen research, open government, community development, intelligent healthcare, and social networks. As a result, data literacy is in high demand among all sections of society. However, this study narrows its focus on data literacy for researchers in the context of a university.

The third radical change concerns the level of penetration and influence of the data revolution. According to Corrall (2019), the impact of the data revolution can be felt in all areas of work, activities and services in libraries. Corall's view conquer with earlier studies by Connelly et al (2016) and Kitchin (2014) which highlighted that it is as a result of this revolution that 21st-century libraries have been turned into data-intensive organisations courtesy of the ever-increasing "developments in digital scholarship, bibliometrics and altmetrics, open science, linked open data, learning analytics and data-based decision making". The institutional

management of the library must, therefore, realise the need to invest heavily in data literacy due to the radical change in the role of the library in supporting researchers with data literacy services (Corrall, 2019a; Tang & Hu, 2019; Koltay, 2017a; Koltay, 2017b).

The fourth radical change courtesy of the data revolution is the new role played by librarians in research. With libraries repositioning and offering new services such as research data management, it means that there is a need for a radical change in librarianship as a profession (Koltay, 2016). The future librarian must be data literate (Rice & Southall, 2016). Librarians need some basic competence in dealing with data, only then will they be able to stand a chance of effectively contributing to the library of the future. According to Koltay (2015) while engaging librarians, some libraries are gradually introducing data literacy services as part of their RDM services. This state of affairs confirms the assertion that for libraries to remain relevant in the 21st century, the expertise of their staff, in this case, information professionals, must be put into use and start offering research data services (Christensen-Dalsgaard et al., 2012).

According to Corrall (2019), the role played by librarians in offering data-related services range from "helping students and faculty find and use external social data sets, to advising on management and sharing of original research data in the context of funder requirements for data management plans". Corrall (2019) adds that today some libraries "have extended their data literacy support to cover other areas and involve more library staff via data literacy training". Data literacy is closely related to research data services (Koltay, 2016; Searle et al., 2015) hence, if librarians wish to support data-intensive research, they would also need to be able to provide data literacy training. The justification of librarians getting involved in data literacy as key stakeholders will be discussed later in this chapter as an element of the Stakeholder theory. A comprehensive discussion on the role of the library in data literacy will also be discussed later in this chapter under the *Intellectus* model.

Considering the radical changes experienced by researchers in the handling of data today, as highlighted above and in line with the response to the needs of researchers in universities, academic libraries have a key role to play in ensuring their satisfaction (Shelly & Jackson, 2018; Koltay, 2017). Academic libraries are a key component in a university's wide range of activities including supporting teaching and learning, and research. Considering that universities, as institutions of higher learning, operate in constantly evolving complex and competitive environments, they have a responsibility to adapt to local, national, and

75

international economic, political, and social shifts (Chankseliani, Qoraboyev & Gimranova, 2021) such as the world of complex pluralistic data. Charged with the responsibility of developing the world's future human resources/employees (Köksal, 2019; Talley, 2017), institutions of higher learning must adopt new and innovative ways of responding to the radical changes experienced today. Developing researchers that are highly skilled in the responsible use of data is one of the reasons why data literacy should be one of the 21st-century skills that should be prioritised (Koltay, 2017).

The strength of the Radical Change theory lies in its recognition and acknowledgement of the radical technological evolution and its widespread impact in various praxis of life. Technology has become an ultimate disruptor in research. As research continues to experience a high-tech makeover, the 21st-century researcher needs to acquire new and relevant skills in the wake of their encounter with huge and complex sets of data. Furthermore, radical change in data technology characterised by big data, open data, data reuse and FAIR data exerts pressure on concerned players (stakeholders) to ensure researchers are well equipped with the necessary skills in handling evolving changes in research. The researcher, therefore, applies the radical theory to explain some of the contemporary changes in data management. As a theory, it is an important tool for concerned stakeholders in assessing the radical changes in the research environment in order to alert those concerned to be aware of the evolving needs of a researcher in relation to data management.

Even though the Radical Change theory categorically highlights and points out on the changes experienced, it does not offer an explicit solution in terms of what needs to be done and how to offer a solution to the changes experienced. As a theory, it only situates the problem. In the next section of the study using the Stakeholder theory, the researcher will discuss some of the key players that need to be involved in the development and implementation of a data literacy program in order to cope with the evolving changes in the research environment.

3.3.2   The Stakeholder theory

The development of Stakeholder theory is attributed to Edward R. Freeman which he documented in 1984 (Freeman, Phillips & Sisodia, 2020; Jones, Wicks & Freeman, 2017; Freeman, 2015). The elaboration of the theory is captured in one of his books (Freeman, 1984). In the book, according to Wood, Mitchell, Agle and Bryan (2018), Freeman introduced the concept of a stakeholder as an alternative perspective on "who matters" or who is the most significant in the context of a business, company or organisation. An elaborate development of

the Stakeholder theory was later done by Mitchell et al. (1997) in a study, where a stakeholder typology was developed, classifying different groups of stakeholders in an organisation as, dormant, discretionary, demanding, dominant, dangerous, dependent and definitive stakeholder.

The Stakeholder theory was developed for the business world. It has widely been used as a socio-economic theory in organisational management and business ethics. There is also a wide adoption of the Stakeholder theory not only in the field of business ethics as it is also widely used as a framework in corporate social responsibility approaches. As a theory, it accounts for various constituencies (individuals or groups) that in one way or another have a bearing on an organisation such as employees, suppliers, local communities, creditors, and others (Lin, 2018). In his book, Freeman (1984:46), describes a stakeholder as "an individual or a group that can affect or is affected by the achievement of the organization's objectives". Freeman's definition considers stakeholders as individuals or groups with a specific interest in an entity. As per the definition provided by Project Management Institute (1996), stakeholders are individuals and organizations that are actively involved in the project or whose interests may be influenced, either positively or negatively, by the project's execution or its endpoint.

Since there are various groups of stakeholders in an organisation or company, the basic fundamental question at the heart of the Stakeholder theory is which groups of stakeholders deserve the highest attention from the management and which one does not. What is essential to understand about the Stakeholder theory or concept is the "identification and prioritisation of stakeholders" based on three concepts; power, legitimacy, and urgency. Power, legitimacy, and urgency are social science concepts characterised as stakeholder attributes (Mitchell et al., 1997: 869). The three attributes are considered as factors that determine the kind and level of attention management will give to various stakeholders (Yang, Wang & Jin, 2014; Mitchell et al., 1997).

According to the Stakeholder theory power as an attribute is described as the stakeholder's ability or potential to enforce their will within a given social relationship despite resistance (Mitchell et al., 1997). It is further described as the likelihood that one stakeholder in the organisation's setting, despite resistance from management or other stakeholders, would be in a position to carry out their own will (Miles, 2017; Mojtahedi & Oo, 2017). With power, a stakeholder can influence what they desire in an organisation.

Urgency is "the degree to which stakeholder claims call for immediate attention" (Mitchell et al., 1997: 869). In an organisation, for instance, the needs of some stakeholders deserve to be responded to with immediate effect. For example, if the core business of a university is academic progress, then students are the definitive stakeholders whose needs deserve to be attended to with utmost urgency. The same applies to faculty members.

The Stakeholder theory recognises the legitimacy of each stakeholder. Broadly, legitimacy is described as actions or conduct of an individual considered to be suitable within the parameters of a "socially constructed system of norms, values or beliefs" of an organisation. A legitimate stakeholder is therefore one whose way of conduct is seen to be "appropriate, proper, and desirable" to the organisation, within the framework of the social system (Parent & Deephouse, 2007: 2). Suchman (1995) identifies three forms of legitimacy (pragmatic, moral and cognitive). Although legitimacy as a concept is considered to be a generalized perception, according to Mitchell et al. (1997), that is the actions and conduct of a stakeholder fitting within some systematically and socially constructed customs, principles, beliefs and definitions.

A stakeholder analysis of data literacy actors

The stakeholder theory is appropriate for this study because it allows the researcher to identify interested players and analyse their possible engagement in data literacy implementation in universities. According to Corall (2019a), the Stakeholder theory provides a suitable framework that fits a data literacy study because of its historical application in identifying direct and indirect actors in a phenomenon. The implementation of a data literacy programme in a university is determined by the collaborative efforts of various stakeholders. While some of them will be direct stakeholders in the implementation (librarians) and beneficiaries (postgraduate students and faculty members), some will be indirect stakeholders. Secondly, it is because of its description of different stakeholders while categorising them by their attributes and indicating specifically "who really counts for the firm" and needs attention (Bonnafous-Boucher & Rendtorff, 2016:3). Practically for this study, in view of Someh et al. (2019) study, due to the complex-pluralist context of the 21st century of the big-data landscape, the Stakeholder theory would be an ideal theory for the study in terms of identifying players and actors (stakeholders) in the implementation of data literacy in a university. In the context of this study, therefore, this theory helps in identifying individuals, groups and organisations with salient interests in data literacy in universities.

Despite the Stakeholder theory being predominantly a socio-economic theory, according to Leisyte and Westerheijden (2014), there has been a steady uptake and application especially in higher education since the late 20th and early 21st centuries. The understanding is that there are individuals or groups of people or organisations with a legitimate interest in education. Students, teachers and governments are just a fraction of different stakeholders who have the right to intervene in education matters considering, the power, urgency and legitimacy they command in these sectors (Bjorkquist, 2011).

As a new concept advocating for data competencies and gradually taking shape, achieving data literacy demands collaboration between various stakeholders. One of the findings from the Ridsdale, et al. (2015) study identified the aspect of collaboration among various stakeholders in order to achieve data literacy in society. They write that the "best practices for teaching data literacy education include collaboration between educators, organizations, and institutions to ensure goals are being met by all stakeholders" (Ridsdale, et al. 2015:2). They further acknowledge the diversity of experts in a university setting, such as students, faculty and librarians, who in their capacity are exposed to and interact with data at some point. The study advises that these are the different groups of people to be considered as priority partners while developing and implementing a data literacy programme (Ridsdale, et al. 2015).

Stakeholders could be categorised as either internal or external stakeholders (Marques et al., 2019). To implement a data literacy programme in a university, the two categories should be strongly considered. Internal stakeholders are senior members of the organisation with control over resources. They wield much of the influence within the setup of the organisation. External stakeholders can be considered as informal members with no direct control over resources in the organisation. However, there is a chance that they can impact the project positively or negatively (Aaltonen & Kujala, 2010).

For this study, students and faculty are considered to be among the key internal stakeholders. They are the definitive stakeholders of data literacy training implementation initiatives. This study singles out postgraduate students and faculty members as key stakeholders because of their involvement in research. As pointed out in Chapter 2 section 2.7.1, post-graduate students are required to conduct research as an academic requirement before they are awarded a relevant degree in their respective programmes of study. It is also a requirement that faculty members carry out research, whether in their respective disciplines or as a member of an interdisciplinary team. Considered to be members of the research community in a university due to their

involvement in research activities, postgraduate students and faculty, need to be empowered with skills that will help them in handling or managing and sharing their research data. According to Carlson, et al. (2015), both the graduate students and faculty members need to be helped in developing data literacy skills, if data literacy is then regarded as a 21st-century skill. This study, for instance, singles out the two groups of stakeholders (postgraduate students and faculty members) as recipients of data literacy programmes. Hierarchically, according to the stakeholder theory, the researchers would be considered to be the definitive stakeholders of data literacy implementation. As definitive stakeholders, they need to be given attention not only because their needs "represent a legitimate claim" but because they are "likely to exercise power because of a sense of urgency" in their claim (Flak, Nordheim & Munkvold, 2008).

Another group of stakeholders, who would need to be considered in the implementation of successful data literacy programmes, would be the librarians. In Chapter 2 of this study, we highlighted the impact that the ICT revolution has had on the current librarian's role. The librarian's role in the provision of research services has been transitioned into that of a key player in the research lifecycle. In a study on the role that academic librarians could play in data literacy instructions to meteorological students at the University of Oslo, Department of Geosciences, Frank and Pharo (2016) started their study with an assumption that considering their traditional role of providing information literacy training, academic librarians are better suited to provide data literacy training too. However, the study's findings indicated that librarians as information professionals would need to improve their data skills as they will be entering a new domain. In this context, there appears to be a great need for librarians to initiate data literacy programmes in the training of postgraduate students and faculty members considering the magnitude and complexity of research data they encounter (Corrall, 2012; Simons & Searle, 2014).

One of the reasons why librarians are considered key stakeholders in data literacy is because of their already existing expertise in training users in information literacy and the impact they have on students (Koltay, 2017). In a preceding study, Koltay (2015) advises that, as an evolving role, librarians have an opportunity to exert their influence in scientific research especially in academic institutions by offering data literacy-related services.

The implementation of a data literacy programme and its success demands the buy-in of key officers and personnel within the organisational structure of a university (Chión, Charles & Morales, 2020). In a setting of an institution of higher learning such as a university, other than

faculty, other staff members, especially in administration and management, would be considered as stakeholders in the achievement of a data literacy programme. This cluster of stakeholders has the power to influence resource allocation towards the implementation and success of the programme. As the general administration of a university, they could be considered as key stakeholders in data literacy following the need for infrastructure to support data literacy (Yang & Li, 2020). For instance, a mutual collaboration between the library and the ICT department is necessary for the purpose of having an ICT infrastructure that supports data literacy programmes (Koltay, 2017). In the opinion of Corrall (2019a), there is potential for numerous collaborators whose knowledge can enhance data literacy provision and create consistency as long as universities intend to educate professionals who will continue engaging with data in their daily lives after their academic experience. This comprises not only librarians but also established individuals and organisations interested in data related activities.

Following Corrall's (2019a) view while in agreement with Mandinach and Gummer (2013), it could be concluded that the implementation of data literacy programmes, therefore, requires the involvement and participation of external stakeholders. The need to include external stakeholders is based on the fact that training researchers in data literacy is not a short-term endeavour especially in the context of a university as this would be a continuous programme with continuous activities. Furthermore, data literacy training is meant to prepare postgraduate students for interaction with data not only during their research process while at the university but also how to handle data in the workplace and posterity.

In this case, there are some potential external collaborators and experts with an interest in data literacy who would be part of the strategy of establishing data literacy programmes or initiatives in universities. Though their involvement might not be as direct as that of internal stakeholders, their involvement is necessary. For instance, research and development funding bodies or grant agencies could play a key role as partners in data literacy development. They could be international or local agencies. However, as Koltay (2017:3) puts it, "the involvement of different stakeholders (research funders, data managers, research institutions and publishers) differ by region and by country".

In Kenya for instance the National Commission for Science, Technology, and Innovation (NACOSTI, 2021) and the National Research Fund (NRF, 2019) are considered to be interested in research that is taking place in universities hence they are external stakeholders with an interest in data literacy. The two are government agencies that give research permits and

81

allocate funding for research respectively. As government agencies, they provide policy and guidelines related to the whole process of research data management. Any researcher must therefore be educated or trained in order to understand some of the laid down requirements. Another government agency that might fall under the category of external stakeholders and could be interested in data literacy as a programme in Kenyan universities is the Commission for University Education (CUE) in Kenya. This is the body mandated with the responsibility for providing standards and guidelines for the establishment of universities as well as university academic programmes in Kenya. It is through the standards and guidelines provided by the CUE that most of the Kenyan universities through the library have managed to implement information literacy programmes.

Other external stakeholders are the publishers especially those that require researchers to deposit, archive or publish their data alongside their manuscripts at the time of submission. The insistence from some of the publishers is that just as the literature, data needs to be well archived for the purpose of findability, re-use and citation. For that reason, they require authors to publish their data alongside their publications. Considered as a good scientific practice today, Cousijn et al. (2018:2) point out that the published data makes data to be "FAIR", that is Findable, Accessible, Interoperable, Reusable and "accessible from the primary article". To adhere to all these requirements, a researcher would need to be trained on how to manoeuvre and adapt to the new regulations.

Furthermore, organizations or data centres that provide data services in the country and even beyond would-be key stakeholders in establishing data literacy programmes in universities. Some of the research funders insist on making data available. According to Cousijn et al (2018:2) "funders and research institutions increasingly will require full primary data archiving and citation" which is a service provided by data centres if the university does not have the required capacity to provide such a service.

The strength of the Stakeholder theory lies in the acknowledgement that for the successful implementation of a project or programme, there must be some key players (stakeholders) involved. This theory also helps in identifying the two categories of stakeholders (internal and external). Though the hierarchical typology of the theory as developed by Mitchell et al. (1997) provides some unique identification and classification of stakeholders, this study will not apply it in total other than the identification of researchers (postgraduate students and faculty members) as definitive stakeholders. However, the Stakeholder theory does not point out the

82

capacity of stakeholders in the implementation of data literacy. The Intellectus Model or the Intellectual Capital model which will be discussed next fills this gap.

3.3.3 The 'Intellectual Capital' model

The Intellectual Capital model commonly referred to as the Intellectus model was developed from the understanding of what intellectual capital is. Intellectual Capital is the accumulation of an organisation's intangible resources or assets that are not included in the balance sheet (Bryl, 2019). The model was developed as a social-evolutionary approach to social issues in 2003 (Bueno et al., 2006). Traditionally, the intellectual model has been conceptualised to consist of three principal components; human capital, structural or organisational capital, and relational (or market/customer) capital (Bryl, 2019; Bueno, Salmador & Longo-Somoza, 2014; Karchegani, Sofian & Amin, 2013). In the business world, the three principal components are considered to define the strength or weakness of an organisation.

Human capital refers to the professional ability, competence and expertise of the organisation's human resources or employees (Bryl, 2019; Bueno, Salmador & Longo-Somoza, 2014). It includes "knowledge, skills, innovativeness and the ability to meet the task at hand." It encompasses all that is associated with people who are the employees of the organisation is classified as human capital. According to Karchegani, Sofian and Amin (2013:566), it comprises different elements of the employees such as their "tacit knowledge, skills, experience and their attitude". It is the overall implicit or explicit abilities or skills that either individuals or a team of personnel in an organisation possess.

Organisational capital is the collective knowledge or ability of an organisation. It is also characterized as investments in innovative organizational structures used by modern organizations. These structures offer employees greater decision-making autonomy, cross-functional job roles, and ongoing learning opportunities (Stucki & Wochner, 2019; Torres, 2019). According to Barbieri et al. (2021:1) it includes, "knowledge resources rooted at the organizational level, such as processes, databases, policies, culture, vision, mission, and value statement."

Relational capital which is also known as the client or customer relation is further classified into two: business capital and social capital (Paoloni, Cesaroni & Demartini, 2018). "Business capital express the value that represents for the company the relations hold with the main agents connected with it basic business" while "social capital is made up of the intangibles generated

83

by the relationships which the organisation maintains with the social agents of the environment" (Bueno et al. 2006:398).

In this study, the Intellectual Capital theory or model is applied in the evaluation of both concrete and potential abilities, capacity and roles academic libraries could play in advancing data literacy competencies for researchers. Through the lens of this model, this study recognises that the library can implement a data literacy programme (Koltay, 2017). The intellectual model, therefore, encourages academic libraries to step out of their comfort zones and utilise their potential (Burress, Mann & Neville, 2020; Moran, 2019). The model calls on organisations to leverage and capitalise on their existing intellectual ability and be innovative and venture into new activities, take up new roles and perform different professional tasks to meet the client's needs (Swart, 2006). The identification and singling out of the library as the intellectual capital that would take lead in the implementation of data literacy is based on its past involvement in similar programs and its evolving role in the life of a researcher as explained below. Even though the the term "intellectual capital" is traditionally associated with the collective knowledge, skills, and expertise within an organization or institution, in this context, the conceptualization of the library as intellectual capital implies recognizing the library not just as a repository of physical books but as a dynamic hub of intellectual resources, information, and expertise. It is a central entity that possesses the intellectual capital necessary to lead initiatives related to data literacy. What this simply means is that libraries have what it takes to run a data literacy programme.

For a long time, university libraries have been revered for their role as gateways of information or as information hubs. As a core pillar in academics and research in a university, university libraries tend to play a big role in supporting the university's vision and mission. There has been a significant transition in the roles played by university libraries today compared to past roles. Libraries are taking up new roles considering the new trends shaping library science. According to Foo et al. (2002:6), as libraries position as learning and research centres in a university, academic libraries are proving to be playing a key role in providing a competitive advantage for a university. Today, libraries' traditional role which they were widely known for, that is, making available information resources to support teaching and learning, has significantly been challenged and changed. Campbell (2006:75-76) asserts that libraries are no longer mandated only to collect, stock and circulate books. Instead, they have metamorphosed into providing other services that facilitate "research activities, such as bibliographies, reference services and information literacy" training. And as pointed out in Chapter 2 section

84

2.11.1, data literacy is gradually becoming a core service to be rendered by the library considering its intellectual capital.

In the wake of changing information needs of a user this study singles out the library as a key department with the capacity to develop, implement and promote data literacy programmes and activities. The current needs of a researcher demand a one-stop-shop that would provide tailor-made services that would simplify their research engagement. Therefore, libraries, according to Verbaan and Cox (2014), have a long-standing experience considering their existing influence in open access and provision of information literacy training, hence well suited to implement data literacy in universities.

As academic libraries, they are well placed to play a key role in different stages of a research lifecycle. The study by Tenopir et al. (2014:89) which was based on two surveys, found out that other than offering various RDM services, a number of libraries in Canada and the United States of America were engaged in research data services, ranging from assisting with "data management plans to preparing and preserving research data for deposit in data repositories". In his study where he tries to "identify the new roles for academic libraries in supporting data-intensive research" Tibor (2019:97) writes that academic libraries have their work cut out due to data-intensive research; that they should provide services to researchers who work with research data. He points out that even though some of the data services are fairly new, "others heavily build on librarians' traditional, well-known skills" (Tibor, 2019:97).

According to Pryor (2014), academic library services have undergone significant transformations, one of them being offering Research Data Services (RDS). Given the huge volumes and complexity of handling research data, interpretation and implementation of funder policies as well as adherence to open scholarship, libraries are better placed to offer advisory and technical services to support researchers in RDM. Cox et al. (2019) express it as follows: RDM, particularly the concept of Data Management Planning (DMP) starting from the project's inception, suggests that libraries have a role in assisting with information management across the research process and the entire data lifecycle.

While declaring the favourable position academic libraries have in helping researchers navigate through data-related challenges, the Association of College and Research Libraries (2018) pointed out the ability of the libraries' human capital. That librarians in these libraries have the ability to foster interdepartmental communication and collaboration. This kind of collaboration is seen to be effective in research data management. Furthermore, librarians play a vital role

because, in their day-to-day service to researchers, they tend to familiarise themselves with researchers' data needs. They are also advocating inventive publishing models which include open access publishing. The Association of College and Research Libraries adds that by virtue of offering information literacy services, some librarians are already including a component of data literacy within information literacy instructions as well as working towards attaining essential skills on how to manage data (Association of College and Research Libraries, 2018).

The strength of the Intellectual Capital model is in the recognition of the human capital as well as the organisational capital in the context of an organisation. It is based on this that this study chooses it as a model that can be applied in the context of the implementation of data literacy in a university. In the context of this study, the researcher can use the model to identify the human capital as well as organisational capital in the implementation of a data literacy programme in a university. From the discussion, academic libraries and librarians have been singled out as having prior knowledge considering their involvement in the implementation of information literacy programmes for library users in universities. Furthermore, there is an existing rapport and network between librarians and library users. Organisationally, the library as mentioned before has a hierarchical structure with various experts who can assist in the implementation of data literacy. It has the capacity to develop a policy framework, based on its experience, that would steer the implementation of data literacy. However, the Intellectual Capital model does not explain how the library and librarians would implement a data literacy programme. Furthermore, it does not indicate what is to be included in the programme and how it is supposed to be implemented. This study, therefore, borrows from the Bielefeld University RDM training model and is discussed in the next section.

3.3.4 The Bielefeld University RDM training model

The Bielefeld University RDM training was developed by Wiljes and Cimiano (2019:3). The faculty members at the University developed it as an interdisciplinary course, purposely to fill the then existing gap where students were not considered as a priority in many universities where RDM training was taking place (Wiljes & Cimiano, 2019). The reason for having it as an interdisciplinary course was based on the fact that students were not receiving training in data management; hence, they were to be imparted with general skills which they could later apply in their research work, no matter the discipline. Bielefeld University RDM training, therefore, became one of the first institutions to recognize researchers' needs and decided to develop a program to help meet those needs.

To successfully implement the course which was not anchored in any specific degree program at the university, they collaborated with other experts such as the research data managers at the university who would help in co-teaching the course. At the onset of the course, they started with small numbers of students. However, the numbers kept on increasing gradually until they had to move from one room to another in search of space. The goal of Wiljes and Cimiano in developing the course was to empower students with skills that would help them navigate the data landscape by themselves. In the implementation of the course, a competence-based approach was used. The approach allowed the use of practical exercises (Wiljes & Cimiano, 2019).

The Bielefeld University RDM data training provides a model for the implementation of data literacy in universities. The model, according to Wiljes and Cimiano (2019) points at what should be done at the initial stage of the implementation, such as identifying researchers' data needs, developing the programme and marketing the programme among potential beneficiaries. In this case, the target beneficiaries are the researchers (postgraduate students and faculty members). The model singles out some of the most essential areas that could be included in a data literacy program. Some of the key areas to be covered should include but are not limited to data collection planning, data collection methods, data processing tools and methods, data analysis, data visualization, data publication, data preservation and data re-use. The FAIR data principles and their implication in research should be considered as part of it too.

The Bielefeld University RDM model provides this study with an opportunity to identify some of the key areas such a programme should focus on as well as the pathway to take in the implementation of the programme. However, the researcher is cognisant of the fact that while the Bielefeld University RDM programme was implemented in Germany, what this study hopes to implement is to be done in Kenyan universities. These are two different regions. The uniqueness of a context matter and the research will be very much aware of this while developing a data literacy framework for universities in Kenya.

3.4   An initial framework for data literacy

The development of a data literacy framework is aimed at having in place a basic structure that will help in the implementation of a data literacy program for researchers. The framework should provide key pillars upon which implementers can choose and adapt what is suitable considering their context, discipline and mode of implementation. Koltay (2016a:303) makes it categorical that "data literacy is closely related to research data services that include Research

Data Management (RDM)". These tenets should form the background in the wake of the development of a data literacy framework regardless of the context, subject area or any other category. For that reason, this study aims at developing an initial data literacy framework as shown in Figure 3.1. Other than the key variables that need to be considered, the framework captions theoretical frameworks and models that influence its development as well as implementation.

The development of a data literacy framework is therefore meant to provide a standardised approach towards the implementation of data literacy to achieve a data literate community of researchers. The design of such a framework is strongly dependent on existing practices towards data literacy by either academic institutions or government agencies and organisations with an interest in data.

Practices in different universities and organisations as previously discussed reveal various approaches and activities aimed towards enhancing or imparting data literacy skills, whether explicitly or implicitly. From the case studies presented in Chapter 2, section 2.9, there is an indication that in some of the universities where data training is taking place, the responsibility of the training is vested in the library as the library plays a major role in research. However, each seems to have a different approach hence the need for a harmonised or well-structured framework in approaching data literacy. Various data literacy competencies have been proposed though there are some similarities cutting across. Variations in this view, however, depending on the context, purpose pursued and even academic discipline.
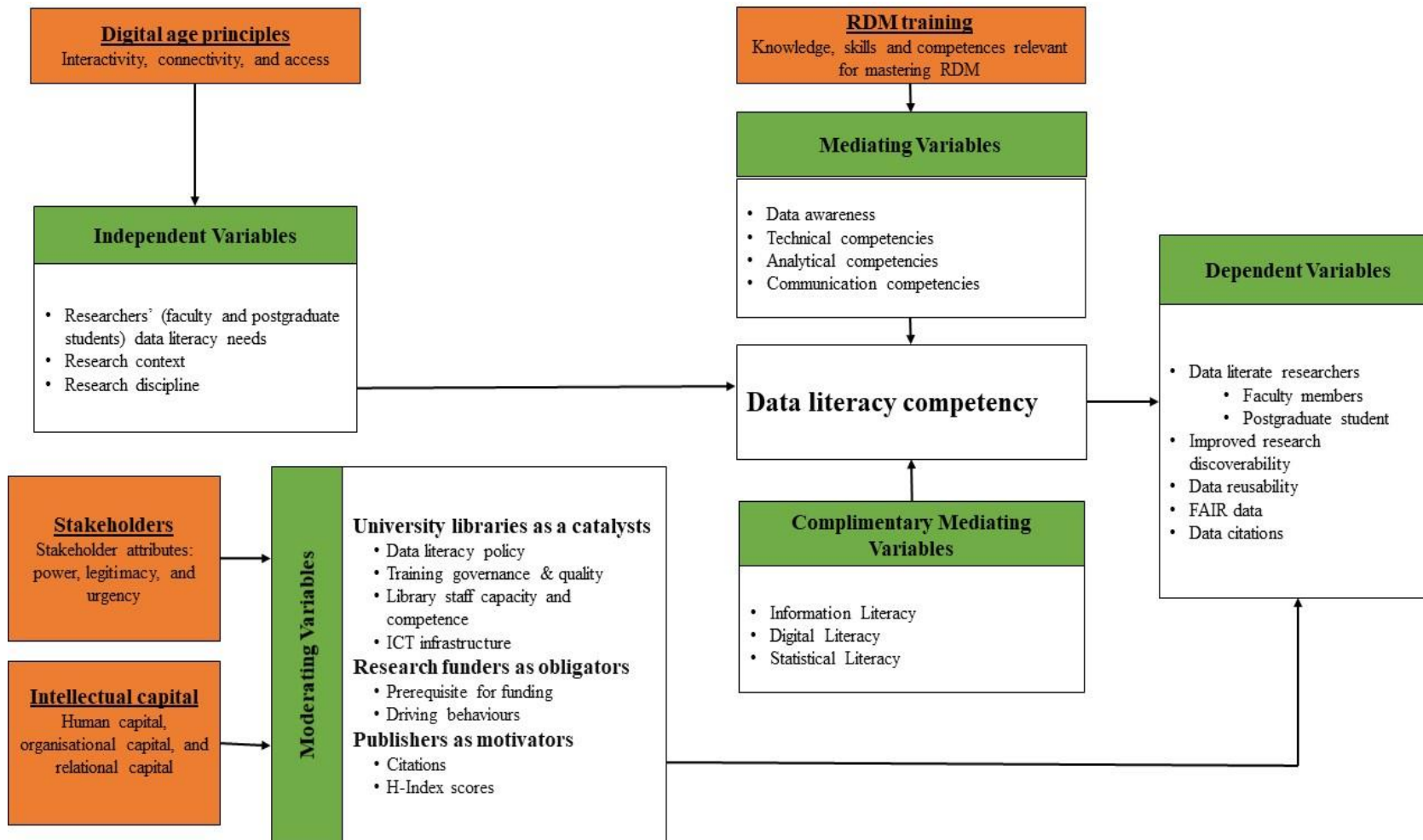
Figure 3- 1: An initial data literacy framework

This study proposes an initial data literacy framework as captured in Figure 3.1. The initial framework, was primarily conceived based on an extensive review of existing literature, aligning with the objectives outlined in the study. This literature-driven development ensured that the initial framework was grounded in established theories, models, and best practices in the field of data literacy. However, it is crucial to note that the framework was then further refined and enhanced through the incorporation of primary data obtained from respondents during the course of the study.

Key components of the framework are captured under various variables (coloured in green). These are the independent variables, moderating variables, mediating variables, dependent variables and complimentary variables. Variables are features or factors that are fundamental to the topic or problem being examined and may affect the study's conclusion (Mertler, 2019).

The independent variables of the study include researchers' (faculty and postgraduate students) data literacy needs, their research context as well as a research discipline. The independent variables as shown in the figure are necessitated by the digital age principles (interactivity, connectivity and access) and their impact on research and researchers. As per Carlson et al.'s (2011) statement, the landscape and execution of research and scholarship are undergoing significant transformation due to the emergence of accessible high-bandwidth networks, the ability to store vast quantities of data, and the availability of advanced tools for data analysis and visualization. The principles situate the background as to why researchers need data literacy training today. Furthermore, the context of research, as well as study discipline, have varied requirements. For instance, data management and handling in social sciences might differ from that of natural and applied sciences. Data literacy needs for researchers in different study disciplines have, therefore, undergone gradual change and variation hence requiring researchers to acquire some advanced skills in data handling.

To achieve a data literate community of researchers, there is a need to develop a data literacy program that provides relevant competencies in data awareness, data technical skills, data analytical skills and data communication skills as captured under the mediating variables. In this study, mediating variables provide a pathway or a bridge for intervention between independent variables and dependent variables. To be data literate one should attain competencies in data awareness, technical skills in using data-related technologies, data analysis skills as well as data communication skills. The four categories of competencies would therefore provide a backbone for the data literacy curriculum to be developed. The Bielefeld

90

University RDM training programme provides a model that could be adopted in the implementation of data literacy training. The Bielefeld University RDM training model was developed by Wiljes and Cimiano (2019).

In order to achieve data literacy, there is a need to involve various stakeholders who would play various roles. The different stakeholders, whether internal or external, are associated with different attributes (power, legitimacy and urgency) depending on the role they play in implementing data literacy in universities. According to figure 3.1, the Stakeholder theory responds to the question, "who are the key players in the implementation of data literacy". For the study, the stakeholders are considered to be part of the moderating variables. Moderating variables are meant to strengthen or alter any significant relationship between the independent and the dependent variables. Considering the role of different stakeholders, they could affect the achievement of a data literate community of researchers in a university. Other stakeholders that would influence the implementation of data literacy would be research funders and publishers considering some of their requirements towards researchers and authors. However, in the context of this study, faculty members and postgraduate students are considered stakeholders in the implementation.

Furthermore, the moderating variables point to the fact that the implementation of data literacy must be spearheaded by a group of stakeholders that have the competence as well as experience. This study singles out the university library and librarians as catalysts in its implementation. The library as a department is singled out for having the intellectual capital (human capital and organisational capital) to influence and spearhead the implementation of data literacy. It is assumed that, following its experience in developing and implementing information literacy programs, the library can develop a data literacy policy framework, has library staff with the capacity to run the program and has the technical infrastructure to support the achievement of a data literacy programme. According to Wong (2010), in the design of a data literacy framework, one cannot avoid the influence of information literacy. He suggests that the development of a data literacy framework needs to borrow a leaf from an information literacy programme. The structural organisation of the library in running an information literacy programme, the expert knowledge of the librarians and especially their interaction with users, the ICT infrastructure and policy development ability gives the library an upper hand in implementing data literacy.

The researcher points out that other than the full implementation of the data literacy program, the researcher will also need to have some skills in information literacy, digital literacy and statistical literacy. The three categories of literacies are necessary and can contribute towards achieving full data literacy hence considered to be complimentary mediating variables. A researcher with prior knowledge in the three mentioned literacies will find it equally beneficial or easy in attaining data literacy. Some components from the three literacies could be incorporated in data literacy training.

The development of an initial framework is based on the literature consulted by the researcher while incorporating various practices from different institutions. However, the final framework to be proposed by the study will be developed and be presented in the final Chapter of the study.

3.5   Summary

The two theories and two models provide context for a framework upon which to develop and implement a data literacy programme at universities in Kenya. They lay a foundation upon which to justify why there is a need for data literacy competencies.

The Radical Change theory awakens this need by laying bare the current situation, environment and challenges experienced by researchers today. It points out that data deluge is real and challenging to researchers. There is an increase in the amount of data that researchers are dealing with as well as an evolution of ICTs which are getting integrated into research. Then there is the complex world which is leaning towards a data-decision driven society. However, the implementation of such a programme to mitigate existing radical changes experienced by researchers requires a team of stakeholders.

From the discussion, the implementation of data literacy in universities requires singling out key stakeholders both internally and externally. Other than postgraduate students and faculty members who were mentioned as the definitive stakeholders in the whole exercise, librarians form another group of salient stakeholders in the implementation of a data literacy programme. Externally, this study singles out government agencies such as research funders and publishers as partners in the implementation of a data literacy programme.

It was argued, in this chapter, that academic libraries develop the capacity to influence and spearhead the implementation of data literacy. The Intellectual Model was used to reveal the capacity and potential of the libraries. Current trends in library and information science show

that library services have significantly changed in the recent past with many libraries offering information literacy and research data management services. Libraries have the capacity in terms of personnel, technology and organisational structure. For instance, there is a gradual involvement of librarians in data-related initiatives to help researchers throughout their research journey. With evolving research needs, librarians in some academic libraries have creatively designed ways of offering RDM services. The involvement of an academic library in data-related services is based on organizational capacity as well as the relevant technological infrastructure available in the library. Experience, which includes training users in information literacy, also gives the library an upper hand as the best suited in implementing data literacy.

The call, therefore, is that the 21st-century academic library must transition and move with the changing times in terms of service delivery. Taking note of the radical changes and taking advantage of the already existing expertise in the library as well as the level of acceptance and trust, academic libraries should start offering data literacy training to researchers despite their diversity. Otherwise, if academic libraries will not quickly adapt to the new changes and seize the opportunity of reinventing themselves, chances are that they "might become redundant or irrelevant in the new competitive environment of information delivery" (Tait, Martzoukou & Reid, 2016:7-8).

The researcher presented two theories and two models that support the study, demonstrating that the study is grounded in established ideas. Furthermore, the researcher presented an initial data literacy framework based on the reviewed literature and aspects of the frameworks and models identified. The revised framework is presented in Section 8.7.

In the next chapter, the researcher presents the methodology that was used to carry out the study.

CHAPTER 4

4.   RESEARCH METHODOLOGY

A tangible roadmap for implementing and operationalizing the concepts introduced in the previous chapter is provided in Chapter 4, which transitions from the theoretical underpinnings discussed in the previous chapter. It outlines the methods, techniques, tools, and techniques used to accomplish the study's specific objectives, bridging the gap between theory and practice.

4.1   Introduction

In the preceding chapter the researcher presented two theories and two models in relation to data literacy in view of developing of a framework for the study. Furthermore, the chapter presented an initial framework developed by the researcher, drawn from the literature review of the study, guided by the study's variables. In this chapter (Chapter 4), the researcher presented the study's methodology. Research methodology refers to the strategies, techniques and tools the researcher employs in carrying out a particular study with the aim of achieving specific objectives of the study. The chapter presented the research paradigm or the philosophical worldview for the study, research approach, research design, the study's target population, sampling techniques and the study's sample size. The chapter also presented the pre-testing, validity and reliability of the study, methods of data analysis and finally some of the ethical considerations the study had to adhere to.

4.2   Research paradigm

A philosophical research paradigm of a study expounds the researcher's philosophical worldview in relation to the study. The concept of a paradigm can be traced in the works of Thomas Kuhn in 1962 (Kuhn, 1962). Thomas Kuhn was an American physicist and philosopher, who used the word paradigm to mean a philosophical way of thinking (Kaehne 2017; Shin 2020; Orman 2016). In its later development and use in different disciplines, the word paradigm adopted the meaning of a "worldview" in educational and social research (Ngulube, 2019; Shah & Al-Bargi, 2013; Taylor & Medina, 2011). Kivunja and Kuyini (2017:26) define worldview as "a set of shared views, perspectives, school of thought, or thinking that informs the interpretation or meaning of results of the study".

Various scholars have defined or described the meaning of a paradigm, however, without losing its meaning of a worldview. Guba (1990:7) describes it as a "basic set of beliefs that guide action". To Chilisa and Kawulich (2012:1), it is "a way of describing a worldview that is

94

informed by philosophical assumptions about the nature of social reality", and as a result, researchers can pose questions and apply "appropriate approaches to systematic inquiry". They add that considering its application in the context of a study, it is meant to help in elucidating the reason why researchers choose or plan to choose and use any of the three research approaches (quantitative, qualitative or mixed methods). Bertram and Christiansen (2014) on the other hand identify it as a specific worldview that delineates what research is appropriate and how it should be conducted.

There are various reasons why a researcher would choose and apply a specific paradigm. The choice could be based on either their discipline orientations, research communities they belong to, their interactions with various schools of thoughts or their past experiences (Brown & Dueñas, 2020; Lukenchuk, 2017; Chilisa & Kawulich, 2012). And as such, Chilisa and Kawulich (2012) point out that there is no solitary paradigm that is highly regarded to use or better than the others. For that reason, it is, therefore, the responsibility of the researcher to decide on which one to select as long as the chosen paradigm addresses issues under investigation.

There are several paradigms associated with social science, with each having its own unique ontological and epistemological perspective and they include positivism, constructivism, transformativism, and pragmatism (Kivunja & Kuyini, 2017; Willis, 2007). However, in this section of the chapter, the researcher highlighted and briefly explained some of the most common paradigms associated with social science studies. After the discussion, the researcher pointed out the paradigm that was adopted for the study.

4.2.1   Positivist paradigm

Positivism has its origins in the enlightenment renaissance and is founded on the idea that experience is key for human beings from whom we can learn about the natural world (Houghton, 2011). The initial proposal of positivist paradigm as a research paradigm was by Auguste Comte (1798-1857), a French philosopher. According to Comte as cited by Kivunja and Kuyini (2017) the basis for understanding human behaviour should be experimentation, observation and reason based on experience. Comte maintains that this is the only genuine way to increase knowledge and human understanding.

The goal of positivism as a study paradigm is to provide interpretations based on empirical findings. As a worldview, it assumes that to investigate a phenomenon, the researcher must gather data that supports the study's theoretical framework and be able to test the formulated

hypotheses (Park, Konge & Artino, 2020). According to Kivunja and Kuyini (2017), studies that apply positivist paradigm employ what researchers consider as "deductive logic, formulation of hypotheses, testing of those hypotheses, operational definitions and mathematical equations, calculations, extrapolations, and expressions".

As a research paradigm, however, positivism has been criticized by scholars who pursue research in human sciences as well as some social sciences such as education, psychology and sociology for not being able to meet some of the key expectations. According to critics, the paradigm's quest for predominantly "hard evidence" falls short of acknowledging elements experience encountered by human beings in their daily life. Pham (2018) agrees with this perspective, asserting that one of the primary challenges of employing this paradigm in social research endeavours is its limited capability to effectively gauge aspects associated with individuals' intentions, attitudes, and thoughts. These concepts often lack explicit observability or measurement through sensory experience or without supporting evidence. Furthermore, the paradigm imposes some limitations when investigating some abstract social research concepts developed around human relationships (Hammersley, 2013). Similarly, there is no assurance that research design will ever be completely free of human partiality. However, this does not exclude the paradigm from being used by some social scientists.

4.2.2 Constructivist paradigm

The constructivist paradigm's goal, as pointed out by Kivunja and Kuyini (2017:33), is to recognise the "subjective world of human experience". In its application, a researcher who uses it makes an effort to understand the subject under observation's viewpoint rather than the observer's viewpoint. The emphasis of the paradigm's application is to understand the respondents and their interpretation of their context because according to constructivists, reality is socially constructed (Jung, 2019; Kamal, 2019; Taylor, 2018). This is why sometimes this paradigm is referred to as the Interpretivist paradigm (Rehman & Alharthi, 2016).

Constructivists hold that given that it is a product of people's minds, reality is subjective and that is the reason why people have different perspectives about the same thing (Peck & Mummery, 2018). The assumption is that each person derives meaning or understanding of the world through their interactions with it. In research, the constructivist paradigm is associated with qualitative approaches. While emphasising the place of the subject in research, Creswell and Clark (2018) reiterate that research is shaped by the individual participant's understanding

and perception of the phenomenon and for that reason, participants' perspective about a phenomenon under investigation are critical.

Furthermore, constructivism seeks to comprehend human experiences and the way in which meaning can be drawn from experience. Emphasis is placed on taking note of both the collective and individual experiences in the society which includes their feelings, how they communicate and how they think (Pham, 2018).

As indicated before, constructivists' approach considers that reality is socially constructed (Alverson, 2009), by the people, within a social context and that is the reason why those who advocate for its adoption conclude that the world is a societal construct (Andrews, 2012). They are of the view that research cannot overlook people's subjective experiences and views which as they form part of their social construct (Doyle et al., 2020). Constructivists are interested in determining how social constructions occur and seek to comprehend the world through the lens of people's experience.

The constructivist approach, seeks to comprehend how individuals socially shape their environment by investigating the emergence of objective societal elements and people's perceptions of them (Denicolo, Long & Bradley-Cole, 2016). The paradigm is then considered as most suitable to apply in qualitative research based on understanding of the phenomenon under investigation. However, its proponents do employ some mixed research approaches in the collection of data as well as its analysis owing to the sensitivity of people's experiences (Given, 2008). In the process of carrying out research while applying this worldview, researchers inductively develop a theory or what Creswell (2013) terms as a pattern of meanings.

The constructivist paradigm's strength lies in its understanding and consideration of people's experiences in the society. With this, it allows what is considered as the natural way of collecting data where researchers gather data from respondents within their settings. Cohen, Manion, and Morrison (2007:25) observed that constructivists' focus on a phenomenon as one that preserves contextual integrity by stressing how people being studied think and feel. However, the paradigm has been identified as being disadvantageous as it requires a lot of time as well as resources (Denicolo, Long & Bradley-Cole, 2016). For instance, a researcher will have to travel and meet respondents where they are in order to gather data. This in itself can be expensive and time consuming. That is the reason why some critics conclude that it presents the researcher with some difficulties in the analysis and interpretation of data (Easterby-Smith,

Thorpe & Lowe, 2002). These are some of the reasons why this paradigm was not considered for this study.

4.2.3 Transformative/Critical paradigm

In literature, the transformative or critical paradigm is termed as a participatory paradigm, that advocates for social transformation for instance feminism, freedom and equality (Mertens, 2017). As a school of thought, the paradigm arose in the late 19th century. Proponents of this paradigm were concerned that positivists assumptions were not adequately addressing the needs of those who are marginalised in the society. According to Al Riyami (2015), social justice issues were not adequately addressed because existing worldviews did not recognize the need for action to assist those considered to be on the periphery of society.

At the heart of transformative paradigm is the view that research should include a reform action plan that can transform participants' lives. In their review of the paradigm, Kivunja and Kuyini (2017) point out that the paradigm delineates its investigations in issues of social justice which closely affect people's lives. It addresses social, political, and economic differences that are characteristic of injustices, conflicts, struggles, and general power structures at any level. Since transformative paradigm seeks to alter politics in order to address social oppression and improve social justice in the situation, it is considered as a highly valued worldview in research especially in setting social structures. Its goal is to dispel myths, deceptions, and inaccurate knowledge in order to empower people who can influence action in the society. The paradigm points to research designs aimed at influencing people's transformation (Chilisa & Kawulich, 2012). Creswell (2014:10) reiterates this point of view by indicating that a researcher who uses this paradigm should be concerned with critical issues in the society that affect people's daily life, such as "empowerment, inequality, oppression, domination, suppression, and alienation".

The transformative paradigm is commended for facilitating power balances in society. It is aimed towards restoring equality, social justice and economic inclusion (Pham, 2018; Taylor & Medina, 2011). The paradigm's influence on people's realities stems from its emphasis on the idea that "reality is shaped by culture, politics, economics, race, gender, ethnicity, and disability" (Chilisa & Kawulich, 2012:18).

4.2.4 Pragmatic paradigm

Pragmatism arose as a philosophical movement towards the end of the 19th century in North America (Maxcy, 2003). The discussion that led to its development as a paradigm incorporated the views of both academics and non-academics such as Charles S. Peirce, Oliver W. Holmes

Jr., Chauncey Wright, Nicholas St. Johns Green. Others included John Dewey, William James, George H. Mead and Arthur (Kaushik & Walsh, 2019:2). Pragmatism as a philosophical movement, resulting from the discourse, developed with an agreement from the proponents' "rejection of traditional assumptions about the nature of reality, knowledge, and inquiry" (Kaushik & Walsh, 2019:2). They negated the view held by proponents of positivist and interpretivist paradigms that a single scientific method is sufficient enough for a social science inquiry (Kaushik & Walsh, 2019).

According to Kivunja and Kuyini, (2017) proponents of pragmatic paradigm were of the opinion that it was impossible to obtain the 'truth' about the real world primarily through the application of one scientific process as promoted by positivists and interpretivists. They added that the two paradigms (Positivist and Interpretivist) could not predict and address social reality. As a result, the paradigm is not bound by any particular procedure or theory. It does, however, make an attempt to demonstrate various methodologies that could be used in a variety of research contexts (Hesse-Biber & Johnson, 2015).

While applying it in research, pragmatic paradigm theorists believe that a single-paradigmatic research approach is insufficient Kivunja and Kuyini, (2017). As a result, they sought more pragmatic and pluralistic research approaches that would encourage the combination of different strategies. The combination is intended to shed light on participants' behaviour, the beliefs that underpin those behaviours, and the potential consequences of those behaviours. This eventually led to the pragmatic paradigm, which promotes the use of mixed research methods in order to fully comprehend human behaviours in a practical way. A key tenet of pragmatist philosophy is that knowledge is always gained through experience. The emphasis is that our social experiences shape our perceptions of the world (Kaushik & Walsh, 2019).

4.2.5   Summary

In this section, four research paradigms were examined. It is clear from the analysis that there is no single 'correct' worldview. It is up to the researcher, therefore, "to identify the most appropriate paradigm for informing their research design" in order to best meet the study's objectives or study questions (Chilisa & Kawulich, 2012:2). In summary, while each paradigm is associated with some strength and weaknesses, it is undoubtable that each of them has a distinct role to play in research depending on the context, phenomenon under study and informants taking part in the study. The paradigms have one thing in common: that they provide researchers with a comprehensive perspective and viewpoint to address critical social issues.

Today, an integrated use of paradigms in research is still necessary because it ensures high quality standards of studies in terms of their validity, relevance, reliability and actionable outcomes. A researcher could also choose a paradigm that tends to lean to such an integration of methods. In this case, the pragmatic paradigm was chosen as the study's worldview. The choice of this worldview was based on the fact that at the heart of pragmatism is the escalation and application of knowledge gained from a particular study in question to other situations. Pragmatists argue that knowledge gained through a mixed-method approach can be partially, if not completely, applied to other situations. The main benefit is that it makes social science research more transferable. The knowledge gained can be analysed to determine which parts of the knowledge can be applied to a similar situation in another context. The researcher also considered it to be relevant because the paradigm suits smaller and time-constrained investigations. Critical to the choice of the pragmatic paradigm are the philosophical underpinnings of the paradigm (ontological assumptions and the epistemological considerations in pragmatism).

Pragmatism is consistent with the ontological concept that reality is changeable and shaped by human experiences (add citations). In the context of this study, the pragmatic paradigm recognizes that data literacy and its application in university libraries are not set or absolute. Instead, it acknowledges that our understanding of data literacy and its practical applications are influenced by interactions between academics, library personnel, and the changing technological context. This is consistent with the premise that knowledge is acquired by practical experience, indicating that the reality of data literacy services is determined by the experiences and perceptions of individuals involved. By adopting this ontological perspective, the study emphasizes the need to understand the dynamic nature of data literacy needs, the increasing function of academic libraries and the dynamic roles of library professionals.

The pragmatic paradigm adds to the study's epistemological issues by highlighting the value of practical knowledge obtained through mixed research methodologies. The study's goal is to get a thorough understanding of the topic under investigation by integrating quantitative surveys, qualitative interviews, and document analysis. This approach recognizes that various research methodologies provide distinct insights into the complicated nature of data literacy in university libraries. Furthermore, the pragmatic paradigm encourages flexibility in methodology, allowing researchers to tailor their approach to the specific setting and study concerns. Thus, by adopting pragmatism, the project attempts to generate practical knowledge

based on the realities of data literacy difficulties encountered by academics and libraries in Kenya.

The other paradigms were found unsuitable for the present study for various reasons. The positivist paradigm, which emphasises objective measurement and testing of hypotheses, was not suitable because of its limitations in capturing subjective experiences and social constructions, which are essential to understanding data literacy needs and organisational contexts within private university libraries. While the constructivist paradigm on the other hand focuses on understanding subjective experiences and social constructions, it was not chosen because conducting in-depth qualitative research often necessitates extensive resources and time, which was not feasible for this study given its scope and timeframe. Lastly, the transformative or critical paradigm, which emphasizes social justice and transformative action, did not fit this study because it primarily sought to investigate data literacy needs and organizational infrastructure within private university libraries, rather than explicitly addressing broader societal inequalities or power structures.

4.3   Research approach

A research approach is described as a research methodology that lays out various research techniques, steps and plans the researcher intends to employ in order to meet the study's objectives. The problem of the study, the specific concerns being addressed, the researcher's experience as well as the study's population, do influence the choice of a research approach (Creswell & Creswell, 2018). Conventionally, quantitative, qualitative and mixed methods research approaches, have been identified as the most common research approaches.

Practically, the three should not be taken to be so different from one another as they sound or appear. For instance, qualitative and quantitative methodologies ought not be regarded as rigid categories or as totally opposed to each other. Instead, they should be regarded as merely opposite ends of a continuum, as a study is more likely to be qualitative than quantitative, or vice versa (Williams & Vogt, 2011). Mixed methods research, therefore, falls somewhere in the middle of the spectrum because it incorporates elements from the two approaches. The researcher discussed each approach in more detail in sections 4.3.1, 4.3.2 and 4.3.3.

4.3.1   Quantitative research approach

In the social sciences, quantitative methodology is the prevalent research approach (Coghlan & Brydon-Miller, 2014). It encompasses a set of techniques that a researcher adopts in order to investigate a phenomenon by examining numerical patterns. The approach entails

the collection and conversion of data into numbers in order to draw statistical inferences. It relies on the "collection and analysis of numerical data to describe explain, predict or control variables and phenomena of interest" (Creswell & Creswell, 2018:147). On one hand "some of the numeric data is intrinsically quantitative", for instance when one is asked how much they earn (Coghlan & Brydon-Miller, 2014). On the other hand, the numeric structure is imposed, for instance when the researcher asks a responded to rate a phenomenon on a scale of 1 to 10.

Different scholars have highlighted their understanding and application of quantitative research approach. According to Creswell (2013:4) the aim of this approach is to "test theories by examining the relationship among variables" that are measurable. To Coghlan and Brydon-Miller (2014), quantitative research aims at generating new knowledge and understanding about the social world. Burrell and Gross (2018:2) point out that that social scientists study phenomena or occurrences that affect individuals, referred to as a sample population, using quantitative research approaches. It uses scientific inquiry and data from observations or measurements to investigate questions about the sample group. However, they call for the objectivity of the researcher when a study involves a sample group. In his study, Mertler (2019) explains the goal of quantitative research, stating that it is to establish links between variables, characterize existing circumstances, and occasionally make an effort to explain causal relationships between variables. He adds that quantitative research consists of pre-established procedures and strategies that aid in increasing the objectivity of researchers. He emphasizes that the focus on objectivity is what allows the quantitative researcher to extrapolate research study results beyond the specific context that was examined (Mertler, 2019).

In quantitative research, the collection of data is done using questionnaires, structured observations, or experiments. Data collection in quantitative studies allow researchers to conduct statistical analyses ranging from simple to highly advanced where they cluster data in averages and/or percentages, demonstrating patterns in the data or correlations across aggregated data (Mertler, 2019; Creswell & Creswell, 2018).

The positivist paradigm is commonly associated with quantitative research. This alignment is viewed in the context that quantitative research is built on a philosophical assumption that the world is relatively unchanging; that it is possible to quantify, comprehend, and even make broad generalizations about it. A quantitative research approach perspective, according to Gay, Mills and Airasian (2019) entails direct observation and measurement. They are of the view

that for any researcher to make any conclusion about the world and its phenomena, direct observation and measurement must be involved otherwise the entire exercise will be meaningless. The same view is held by Fraenkel, Wallen and Hyun (2012) that quantitative research is founded on the conviction that feelings should be separated in research. That the world as a whole is made up of facts that can be explored through observation and analysis.

4.3.2 Qualitative approach

Qualitative research is a natural process of inquiry that involves seeking an in-depth understanding of a social phenomenon in its natural context (Creswell & Creswell, 2018). It is said to concentrate on seeking a deeper understanding of a phenomenon while relying on people's experiences in the context of their daily lives. It focuses on the "why" rather than the "what" of social phenomena. It aims at looking at the whole scenario of people's lives rather narrowing down variables and analysing it numerically. Instead of relying on logical and numerical procedures to investigate human phenomena, qualitative research methods employ a wide range of inquiry methods such as discourse analysis, case study, open-ended interview, participant observation, ethnography, introspection, historical research grounded theory, biography, comparative method, casuistry, focus group, literary criticism, meditation practice, (Cibangu, 2012; Mohajan, 2018).

Various studies have attempted to describe what qualitative research entails. According to Creswell (2013), it is an approach used when conducting a detailed investigation that require acquiring and analysing textual or verbal data. It emphasises the use of words and narrations which includes what the respondents said and the voice of the researcher in describing what they observed and experienced. The approach is distinguished by "inductive approaches to knowledge development aimed at generating meaning" (Leavy, 2014:109). The method is used to probe and comprehend a specific social phenomenon; to unravel the meanings that people attribute to activities, contexts, or occurrences; or to obtain a thorough understanding of certain aspect of the society (Leavy, 2014). Subjective experiences of participants and need to have a deeper understanding of their situation, condition and behaviour are at the heart of qualitative research, which every researcher must emphasize. When a researcher's primary goal is to investigate, describe, or explain a phenomenon, this method is considered appropriate. Some of the methods used by a researcher to collect qualitative data include observation, interviews using interview, narratives and analysis of documentary sources, artefacts, as well as group discussions (Groenland & Dana, 2020; Mohajan, 2018; Paradis et al., 2016).

103

Bogdan and Biklen (2007), summarise, five characteristics associated with qualitative research approaches as follows:

i. Qualitative research is naturalistic. This is considered as unique factor qualitative research as it takes place in people's natural settings. Researchers collect data directly from the research context of their interest. The interest of the researcher is to observe people's behaviour in their natural context. It is the natural setting that people live and operate in that serves as the source data for the study.

ii. Qualitative research is descriptive in nature. Words are used to represent qualitative data instead of numbers. The intention is to provide a comprehensive description of the phenomenon under study which could not possibly be done by converting the narration in to numerical sum-ups. In the results, the researcher incorporates quotations taken directly from the collected data so as to show corroboration and support the findings.

iii. The interest of qualitative researchers is not only in the product of the study but also in its process. To the researchers, the outcome of a study is as important as the how and why things happen the way they do.

iv. Qualitative data is analysed inductively. Qualitative research is not concerned about the formulation of hypothesis and the collection of data meant to prove or disapprove the research. Instead, data is collected, synthesised with the aim of producing generalisations. This process is known as inductive reasoning. The researcher benefits from this method because new information emerges. Furthermore, the researcher picks up new ways of understanding whatever is being studied. In the analysis of qualitative data, the researcher uses inductive analysis where themes are generated.

v. Qualitative research is concerned with the way human beings make sense out of their daily lives. It is about people's individual perspectives on their lives.

According to Creswell and Creswell (2018), the constructivist paradigm, which underscores the social construct reality is commonly associated with qualitative research approaches. This is due to characteristics of qualitative research, which include the recording and analysis of events, as well as the uncovering of "deeper meaning and significance" in people's behaviour and experience. When analysing data, the goal of qualitative researchers is to gain a deeper and broader understanding of human perspectives and not just to obtain information that can later be extrapolated to other populations.

4.3.3 Mixed methods approach

Another research approach, as mentioned earlier, is the mixed research methods. Mixed research methods, which was established around the year 2000 (Lund, 2012), has steadily grown in popularity among researchers. It arose as a result of the shortcomings of both quantitative and qualitative research methods (Caruth, 2013). It is sometimes called the "third methodological movement," according to Venkatesh, Brown, and Bala (2013:22). Even though many studies and scholars have tried to describe or define what it is, most of them agree on the fact that it entails the gathering, analysis, and integration of quantitative and qualitative data in one project. For instance, Johnson, et al (2007:123) describe it as "the type of research in which a researcher or team of researchers combine elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration".

In one of their studies, Wisdom and Creswell (2013) indicate that as a methodology of research, mixed research approach enhances systematic incorporation of quantitative and qualitative into one inquiry. They emphasise that the fundamental principle of this methodology is that such integration allows for a more comprehensive and complimentary use of data than separate quantitative and qualitative data collection and analysis (Wisdom & Creswell, 2013:1). Creswell and Clark (2018) describe it as a method that involves collecting, analysing, and combining quantitative and qualitative methods in a single study or series of studies to better fully comprehend a specific problem of the study. Leech and Onwuegbuzie's (2009:268) study concurs with Creswell and Clark by indicating that mixed methods research approach entails the "collection, analysis and the interpretation of quantitative and qualitative data" in one study or a sequence of studies that look into the same underpinning phenomenon. Shorten and Smith (2017) refer to it as a method in which the two data sets (quantitative and qualitative) are collected and analysed in a single study.

The common denominator in these descriptions is that it is a method that incorporates quantitative and qualitative elements. Due to the reason that the use of mixed method approach takes advantage of the strengths from qualitative and quantitative approaches, this has led to its wide adoption and use by researchers. It is credited with allowing researchers to investigate various points of view and discover connections between research questions as a research approach (Brierley, 2017; Creswell, 2009). As an approach, it is characterised by the researcher

105

collecting quantitative and qualitative data while involving philosophical assumptions and theoretical frameworks.

The adoption and use of mixed methods approach is therefore based on the assumption that, integrating quantitative and qualitative data will "yield additional insight" (Creswell & Creswell, 2018:4). It offers a more comprehensive picture as a study approach by acknowledging trends and generalizations, as well as in-depth knowledge of participants' perspectives. Furthermore, it supplies strengths that compensate for the shortcomings of quantitative and qualitative data.

4.3.4   Summary

To accomplish the purpose of this study which is to carry out the feasibility of offering standardised data literacy services at selected private university libraries in Kenya, the researcher chose to use mixed method approaches. One of the reasons for the selection of this approach was to allow for the application of various procedures to strengthen the study of certain aspects while also allowing for data convergence. The method enabled the researcher to study the specific phenomenon from a wider and deeper point of view, resulting in richer and more diverse data that was drawn from multiple approaches. Furthermore, it was anticipated that the combination of quantitative and qualitative approaches would provide a wide-range understanding of the research problem which was better that using a single approach. In conclusion, the approach fits within the application of pragmatist paradigm which had been selected as the study's worldview.

4.4   Mixed method research designs

Creswell and Clark (2018) state that if a researcher decides on a mixed method as the study's approach, what follows is the selection of an appropriate design that best addresses the research problem. A research design, in this case, entails the use of specific techniques when collecting data, data analysis procedures, interpretation of the data and the presentation techniques the researcher employs in their research. Creswell (2014:12), describes research designs as "types of inquiry within qualitative, quantitative, and mixed methods approaches". A researcher employs the use of research designs in research to meet the set objectives. They direct the methodological choices that a researcher makes during the study as they lay the groundwork for the line of reasoning that the researcher intends to use while making interpretations at the end of the study.

Considering the critical role research designs play in a study, Creswell and Clark (2018: 51) point out that they are key to any study as they inform the method decisions that a researcher must consider during their research process and establish the rationale by which they make interpretations at the end of the study. It is critical to emphasize that the selection of a given design is dependent on a variety of criteria, including the aim, techniques, and the benefits and drawbacks connected with each design (Creswell & Clark, 2018).

According to Nagpal, Kornerup and Gibson (2020) there are four research designs that are associated with the mixed method research approach and they are, "triangulation design, embedded design, explanatory design, and exploratory design". The triangulation design could be further sub-divided into four further variants. The following section discusses the four designs.

4.4.1 Convergent design

Also referred to as triangulation/parallel/concurrent design, convergent design is a common and well-known design widely used by researchers who adopt mixed research methods approach. This design, as well as its underpinning goal of bringing different methods together, has indeed been widely discussed and examined by different scholars (Uwe, 2022; Creswell & Clark, 2018; Creswell & Creswell, 2018; Turner, Cardinal & Burton, 2017; Cajal et al., 2015). In using this design, the researcher collects the two different types of data separately but complementarily, on the same topic in order to gain a better understanding of the research problem. The goal of using this design is to bring together the various strengths of quantitative and qualitative methods while avoiding overlapping weaknesses that emerge from using only one of the two (Turner, Cardinal & Burton, 2017). Convergent design is credited as a design that enables researchers whose aim is to make a comparison or contrast between quantitative statistical results and qualitative findings. It also suits researchers who intend to corroborate or expand quantitative findings with qualitative data (Carter et al., 2014).

The application of triangulation design which is also known as "one-phase design" entails employing quantitative and qualitative methods with equal merit at the same time during the study. According to Alexander, (2020), in the use of a triangulation design, even though the collection of two data sets is done separately using different data collection methods, the collections are done concurrently (single-phase timing), hence the reason why it is also known as the "concurrent triangulation design". The researcher's aim is to have clear understanding

107

of the research phenomenon. The two sets of data are merged by combining the different results through data transformation (Turner, Cardinal & Burton, 2017).

According to Creswell and Clark (2018) there are four variants of convergence design. They are listed by Thomas (2014:145) as "the convergence model, the data transformation model, the validating quantitative data model, and the multilevel model". Thomas (2014) goes on to give a distinction between the four indicating that the difference between convergence model and the data transformation model is on the manner in which the researcher tries to merge the two different types of data (at the interpretation stage or while doing analysis). On the other hand, a researcher may adopt the use of validating quantitative data model in a study with the aim of having a better understanding of the study. By including a few open-ended qualitative questions, the researcher validates and expands on the quantitative findings of a study. The researcher in this model collects both types of data using a single survey instrument. A brief analysis of the four variants was done in the following section.

i. Convergence model

The convergence model is considered as the conventional model of a mixed methods triangulation design (Demir & Pismek, 2018). According to Hong et al. (2017), in this design, quantitative and qualitative data are collected and analysed in parallel or complementary ways during the same stage of the study. The researcher collects and analyses quantitative and qualitative data on the same phenomenon separately, and then the various results are converged (by comparing and contrasting the various results) during the interpretation. This model is used by researchers to compare results or to validate, confirm, or corroborate quantitative results with qualitative findings. The goal of this model is to arrive at valid and well-supported conclusions about a single phenomenon (Demir & Pismek 2018).

ii. Data transformation model

The data transformation model entails collecting and analysing quantitative and qualitative data sets separately. However, it differs slightly from the convergence model in that, following the initial analysis, the researcher employs procedures to transform one data type into the other data type (Bazeley, 2018). This is accomplished through the quantification of qualitative findings or the qualification of quantitative results. This transformation enables combination of data during the analysis stage which becomes easier at interpretation (Allen, 2017). In a study, for example, the researcher may decide to deduce qualitative themes from qualitative data and then dichotomously score those

themes as present or not present for each participant. The quantified scores are then compared to the quantitative data, which is analysed with correlation coefficients and logistical regression to identify correlations among categories like gender (Allen, 2017; Bazeley, 2018).

iii. Validating quantitative data model

If a study has quantitative data and the researcher intends to validate and expand them, they will use a validating quantitative data model. This starts by including some qualitative open-ended questions in a predominantly quantitative questionnaire (Ndanu & Syombua, 2015). Two different data sets are collected by the researcher using a single survey instrument. The qualitative items do not produce a rigorous qualitative data set because they are an add-on to a quantitative study. They do, however, offer additional valuable statements to the research which can be used to validate and complement quantitative results (Creswell & Clark, 2018).

iv. Multilevel model

Multilevel research is another name for the multilevel model. According to Creswell (2007), while using a multilevel model, the researcher ensures that different quantitative and qualitative methods are used to address different levels within a system, and the results from each level are combined into a single overall interpretation.

Triangulation design is credited with some strength which make it an option for some studies. By using several techniques to address eminent research challenges, triangulation design tries to improve the process of empirical research (Creswell & Clark, 2018; Hastings, 2010). It is also claimed that the use of triangulation design, which incorporates various models, helps prove that the advantages of one approach compensate for the disadvantages of another (Yeasmin & Rahman, 2012). The triangulation method, as a mixed method design, is very clear and straightforward to new researchers. It clearly demonstrates mixed method characteristics, such as the collection of two sets of data, their analysis, and interpretation. It is also thought to be an efficient design because the collection of both types of data is done simultaneously and approximately within the same phase of the study (Hastings, 2010). Furthermore, in a multi-research study, a researcher can separately and independently collect and analyse each data set using the procedures that are associated with each data type.

Despite being one of the most popular designs among the four, convergence design is considered to be one of the most challenging designs. Because of the simultaneous collection of data and the fact that each type of data is to be accorded equal weight, Hastings (2010) points

out that the use of triangulation design demands a high level of expertise and can be expensive. Nevertheless, researchers can address this challenge by ensuring they allocate sufficient time and resources to collect both sets of data variances. Furthermore, in a study where there is only one principal investigator, they should make sure that they are well trained and prepared in the application of quantitative and qualitative techniques (Creswell & Clark, 2018).

4.4.2 Embedded design

This is a mixed research methods design based on the concept that a solitary data set is insufficient; that in a study, while collecting data, participants respond to several questions and different types of data are gathered (Creswell & Clark, 2018; Edmonds & Kennedy, 2017). In this design, though there are two data sets, one of them is meant to support the other. Researchers adopt the use of embedded design in studies when the intention is to include qualitative or quantitative data in a predominantly quantitative or qualitative study. For instance, incorporating a qualitative component into a quantitative study. Edmonds and Kennedy (2017) cite embedded design as being ideal for experimental or correlational studies where one type of data is secondary to the other.

The embedded design combines diverse data sets. When using it in a study, the researcher incorporates one set of data in a methodology made up of another set of data (Edmonds & Kennedy, 2017). The design allows researchers to collect two sets of data (qualitative and quantitative) using the one-phase or two-phase strategy where one type of data ends playing a supporting role in the current design. However, even in the wake of applying or using the embedded design, the collected quantitative and qualitative data should be based on different questions in the study (Creswell & Clark, 2018). Experimental and the correlational models are the two commonly used forms of embedded designs.

There are some advantages associated with the use of embedded design. The design is considered suitable for researchers who do not have enough time or resources to devote to extensive collection of both quantitative and qualitative data (Yu & Khazanchi, 2017). In the collection of the two sets of data, only one is prioritised. Similarly, because one methodology has less data than the other, the design is thought to be more practically doable for graduate students (Creswell & Clark, 2018; Edmonds & Kennedy, 2017).

However, there are also some challenges associated with the embedded design. One of them is that the researcher must state why either of the data (qualitative or quantitative) data is being

110

collected in the wider quantitative or qualitative investigation. Secondly, when the researcher chooses to use quantitative and qualitative methodologies with the aim of responding to different research questions in one study, there are chances that they will encounter some difficulties when it comes to the integration of results. Thirdly, the majority of studies that have used embedded design are those that have embedded qualitative into quantitative (Creswell & Clark, 2018; Edmonds & Kennedy, 2017). According to Creswell and Clark (2018) there are very few examples that have done it the other way around.

4.4.3 Explanatory design

The primary goal of explorative design is to explain or expand on quantitative findings using qualitative data. The design suits a study in which qualitative data is necessarily needed to explain important results, anomalous results, or unexpected outcomes in the findings (Creswell & Clark, 2018.) Researchers can use explanatory designs in various contexts. For instance, when a researcher intends to create groups of participants on the basis of quantitative results and then follow up with the groups through future qualitative research then the use of explanatory design becomes ideal. Furthermore, a researcher can apply the use of explanatory design when they choose to use quantitative subject characteristics to drive deliberate sampling for a qualitative phase (Creswell & Creswell, 2018).

Explanatory design takes place in two distinct interactive different stages, whereby the researcher starts by collecting and analysing quantitative data, "followed by the collection and analysis of qualitative data in order to explain or expand on the first-phase quantitative results" (Creswell & Creswell, 2018:65). It is assumed that the study is designed to build on the findings of the first phase while using the same participants. Since this design is quantitative in nature, researchers normally focus primarily on quantitative methods rather than qualitative methods.

According to Ghasempour, Bakar and Jahanshahloo (2014:87), "follow-up explanations model and participant selection model" are the two most commonly used variants of explanatory design. Although both models include the two phases, the link between the two phases differ in a way that one is focused on the findings to be investigated further and the other on the relevant participants to be chosen. They also differ in terms of the relative importance given to the two periods (Creswell & Clark, 2018).

One of the advantages associated with the design is that, it does not necessarily require a research team (Creswell & Clark, 2018). Since each phase involves the collection of one type of data, only one researcher can carry out the two different methods in distinct phases. The

111

writing of the final report is straightforward as it is done in two phases and this is also considered to provide readers with a clear delineation. Furthermore, since it frequently commences with a strong quantitative focus, the design is said to appeal more to quantitative researchers (Creswell & Clark, 2018).

Despite being considered as a straightforward design, some disadvantages associated with it have been pointed out. For instance, the design is more challenging depending on the time it takes to implement the two phases. Researchers are advised to be cognisant of the fact that qualitative phase can take longer time depending on the attention the researcher gives it. In case of its adoption in a study, researchers are advised to budget for adequate time for the qualitative phase (Creswell & Clark, 2018). Creswell and Clark (2018) point out the participant inclusion dilemma as another disadvantage of the design. This is based on the reality that, after the first phase, the researcher is faced with the challenge of deciding on who to include in the study's second phase as participants. At this point the researcher has three choices, that is, to use the same participants used in phase one, participants from the same sample and participants from the same population for the two phases.

4.4.4   Exploratory design

The similarity between exploratory design and explanatory design is that they both involve only two-phases. Phase one of this design begins with qualitative data in the investigation of a phenomenon and then progresses to a phase two which is quantitative (Creswell & Creswell, 2018; Creswell & Clark, 2018). According to Creswell and Clark (2018:67), "building from the exploratory results the researcher conducts a development phase by designing a quantitative feature based on the qualitative results". These developments connect "the initial qualitative phase to the subsequent quantitative component of the study" (Creswell & Clark, 2018:77). Since the design process begins qualitatively, the qualitative data is frequently given more weight.

Creswell and Creswell (2018:90-91) indicate that the most common variants of exploratory design are "the new variable development variant, the survey development variant, the intervention development variant and the digital tool development variant". As previously stated, each of them begins with the collection and analysis of qualitative then conclude with a quantitative phase. The variants, according Creswell and Creswell (2018:90) are "distinguished by what is developed in the middle phase of the design".

112

Exploratory design's advantages are similar to the ones of explanatory design because of its two-phase framework and the fact that it only collects one type of data at a time. For instance, the design is considered straightforward following its two separate phases. This makes it easy "to describe, implement, and report" (Creswell & Clark, 2018:89). Another advantage of exploratory design is that, the quantitative aspect in the design, despite it being predominantly qualitative makes the design appealing to those who have a bias towards quantitative methodologies (Creswell & Clark, 2018).

There are some similarities between some of the issues raised in exploratory design and explanatory design. For example, the application of the two-phase approach means that the design takes a bit of a longer time to implement compared to other designs (Creswell & Clark, 2018). This can be difficult for a researcher who is limited by time. Another weakness or disadvantage is that in most cases, exploratory studies use a small number of samples hence making it difficult to generalise them to a larger population (Dudovskiy, 2022).

In view of the discussed designs, the study adopted the convergence design specifically the convergence model. In this case, both qualitative and quantitative data were collected using a questionnaire and interviews. Questionnaires were used to collect majorly quantitative data while interviews with the university librarians and research/reference librarians generated qualitative data. However, the questionnaire also had some open-ended questions which generate qualitative data.

4.5 Population of the study

In research, the population of a study alludes to the wider group of elements to which a researcher would like to generalize; it encompasses every member of a specific category of people, events, or objects (Ary et al., 2019). It is a group in which the researcher is interested to study and to whom the study's findings should be generalised. Mugenda and Mugenda (2012) describes population as that group from which the research makes conclusions. Population is also described as collection of people, elements or events that a researcher analyses; or a set of elements from which a research sample is drawn (Babbie, 2004; Kothari, 2004).

Private chartered universities in the Nairobi metropolitan area made up the study's population. The Nairobi metropolitan area is around 32,000 km2 in size. It comprises of the capital Nairobi, and the surrounding counties of Kajiado, Kiambu, Machakos, and Murang'a. The region is considered as one of the fastest-growing metropolitan areas, with a 3.1% growth rate compared to the national and global norms of 2.5% and 1.2%, respectively. This could explain why the

majority of Kenya's private universities are concentrated in this location, which draws private investors.

In total, there are 12 private universities within the Nairobi Metropolitan area (CUE, 2021). Private universities that made up the population of study included, the Catholic University of Eastern Africa, United States International University, African International University, Adventist University of Africa and Daystar University. These universities were selected on the basis that they were the only chartered private universities within the Nairobi metropolitan area by then and the only ones offering PhD programs. The decision to focus on private universities was a pragmatic one, driven by the limited number of public universities in the area, practical constraints, and because at least one institution declined to participate.

Units of analysis within the selected private universities included postgraduate students, in this case the PhD students, fulltime faculty members, heads of university libraries and research/reference librarians. The assumption was that these groups of participants would provide reliable information because they were directly involved in collecting and/or managing research data and other related research activities. The researcher concentrated the study on the main campuses of the universities, keeping in mind that library services and resources on satellite campuses are limited. A study by Munene (2016a) which was aimed at comparing the state of services and resources at satellite campuses and main campuses revealed that satellite campuses lack some of the most basic services, resources and facilities such as libraries and internet access. Munene (2016a) found out that this state of affairs was playing as an impediment to excellent teaching and research.

In this study, these universities were referred to as A-E, but not in the order listed above. This is due to the researcher's desire that they remain anonymous. Table 4-1 depicts the distribution of the study's population. The figures in the table for each university were provided by university librarians, admission registrars, and human resource officers from each participating university.

As previously stated, the population of the study included University Librarians (those who head library and information services in universities). They were considered eligible for inclusion in the study because they are in charge of all library services. As university librarians, they have the ability to influence policy and the implementation of new services such as data literacy.

Full-time faculty members and PhD students were also included in the study as participants. Full-time faculty members at participating universities are classified as lecturers, senior lecturers, associate professors, and full professors in the study. As previously stated, the two groups were selected with the assumption that they are actively engaged in research activities. PhD students are expected to write and submit a thesis involving data collection, handling, and management before completing their postgraduate degree program. As mentioned in Chapter 2, (see Section 2.7.1) they must also publish at least two articles based on their thesis. Faculty members, on the other hand, are required to publish as part of their advancement through the academic ranks. In their efforts to publish, they become involved in data handling. This makes involving faculty members in the study ideal.

However, the researcher did not include part-time faculty members because, as part-time employees of the selected universities, their contracts do not obligate them to contribute to the research output of these universities. According Munene (2016b) since part-time move from one university to another, they are not allocated a permanent office space from which to use while carrying their research work, hence they cannot be subject to the research requirements similar to their fulltime counterparts.

Another group of librarians who work closely with researchers was also included in the researcher's study. They are usually referred to as research librarians or reference librarians. The title assigned is determined by university policy. The researcher believes that research/reference librarians provide research-related services to researchers, such as data management training.

Table 4-1: Population distribution

	Private chartered university	Fulltime lecturers	Postgraduate students (PhD)	Research/ reference librarians	University Librarians
1.	A.	237	138	1	1
2.	B.	134	135	1	1
3.	C.	40	62	1	1
4.	D.	17	50	1	1
5.	E.	120	182	1	1
Totals		**548**	**567**	**5**	**5**

Table 4-1 displays the overall population at the time when the study was conducted, from which a sample for the research was derived. Subsequent sections will outline the methodologies and techniques employed to obtain the appropriate sample for the study.

4.6   Sampling procedures

In research, because of the coverage of the population, researchers are advised to reduce the number of participants to a manageable size. When emphasizing the need of obtaining a sufficient sample size for research, Etikan, Musa, and Alkassim (2016) state that while using all members of the population would be ideal, it can sometimes be impractical due to population restrictions, emphasizing the importance of adopting sampling procedures. To manage any study, a researcher must narrow down from the population to a sample of the population while keeping time and resources in mind. Researchers would love to study a large group of people. However, challenges such as large populations, diversity, and a geographically dispersed population can make the exercise time-consuming and expensive, making it difficult. As a result, the researcher must obtain a population sample for the study. As Fraenkel, Wallen and Hyun (2011) point out, it is for this reason that it becomes necessary for a researcher to narrow down to a sample for a study.

A study sample is a subset of a population chosen for observation in a study (Creswell & Clark, 2018). To attain the required sample for the study, a researcher will have to employ some scientific sampling strategies or processes. Thus, sampling is the process of selecting individuals who will take part in a study (Etikan & Bala, 2017; Sharma, 2017). It allows a researcher to use procedures to ensure that the 'right' cases are included in a study. "They allow generalisation from the sample to the population since the sample is representative of the population. They are 'right' because they allow extrapolation from the sample to the population because the sample is typical of the population" (Flick, 2011:77).

In research, there are two main types of sampling procedures: probability and non-probability (Sharma, 2017). According to Ary et al. (2019:172) in a study, when "every member or element of the population has a known probability of being chosen in the sample" then that is probability sampling. It is applied in the studies where the investigator is capable of specifying ahead of time which population segments will be represented in a study. This is attained when samples are randomly selected. In this case, the selection of the samples suggests that each element of the population has an equal probability of being chosen to be part of the study. Ary et al. (2019:

116

172) single out "simple random sampling, stratified sampling, cluster sampling, and systematic sampling" as four major types of probability sampling."

On the other hand, non-probability sampling "includes methods of selection in which participants are not chosen by chance" (Ary et al., 2019:172). Etikan, Musa and Alkassim (2016) describe it as a sampling technique in which samples are collected in a way which does not give all participants or units in the population an equal opportunity to participate. The researcher cannot guarantee that every member or unit of the population is represented in the sample when using non-probability sampling. If it is implemented, this means that some people are likely not or there is little likelihood of being sampled (Sharma, 2017). Prior knowledge, expertise and judgement of the researcher are major component in ensuring the success of nonprobability sampling (Kohler, Kreuter & Stuart, 2019). Non-probability sampling is classified into three types: "convenience sampling, purposive sampling, and quota sampling" (Ary et al., 2019:177). Non-probability sampling techniques are said to be both time effective and cost effective.

Two non-probability sampling techniques were used in this study. The first is purposive sampling, in which the researcher selected elements that are most appropriate based on their ability and status as key stakeholders in data literacy implementation in universities.

The second technique is total population sampling. Given the circumstances of the targeted population, which included a diverse group of faculty members and PhD students from the universities chosen and since the researcher used an online questionnaire was there was no guarantee that each member of these groups would respond to the research questions. In order to achieve a reliable sample, the researcher had to allow the online questionnaire to be sent to the entire population of participants in the category of researchers (fulltime faculty members and PhD students). The two sampling techniques are discussed below.

4.6.1   Purposive sampling

It is also referred to as judgment sampling. It is a sampling technique whereby the researcher does a selective identification of specific participants from the target population. The participants are carefully chosen to provide essential information that would not be attained through other means (Taherdoost, 2016).

According to Palinkas et al (2015:1) purposive sampling "is widely used in qualitative research for the identification and selection of information-rich cases related to the phenomenon of interest". They add that other than the knowledge and experience the selected or sampled

participants have about the phenomenon of interest, their "availability and willingness to participate, and the ability to communicate experiences and opinions in an articulate, expressive, and reflective manner" (Palinkas et al., 2015:1) lead to their selection. When explaining the need of using purposive sampling, Bryman (2008) states that the rationale for selecting scenarios and human subjects is intrinsically deliberate. Bryman contends that random selection may fall short of giving the most informative cases or human subjects, resulting in sample bias and skewed research findings.

Purposive sampling is mostly used in qualitative study approaches (Kumar, 2014); the researcher will choose to use it considering the qualitative aspects within the mixed research approach the study will adopt. Primarily, this researcher will use a purposive sampling strategy to choose private universities within the Nairobi Metropolitan area that offer PhD programmes. The researcher will also use purposive sampling technique to target faculty and doctorate students judged to be mostly involved in research as well as university librarians who play an important role in implementing programs and services in university libraries. Finally, purposive sampling will be used in including research/reference librarian in the study considering their involvement in providing data related services to researchers.

Fraenkel, Wallen, and Hyun (2011) highlights that the use of purposive sampling allows the researcher to pick participants based on their prior knowledge of a community and the specific research aims. Even though Palinkas et al. (2015) stated that this type of sampling is associated with some limitations such as the researcher's judgment being incorrect, in this study, the researcher intends to overcome this challenge by including multiple and a wide spectrum of cases of participants in the study who will be selected randomly.

4.6.2   Full/Total population sampling

Having purposively sampled the five universities that participated in the study as well as the researcher, university librarians and research librarians, this study also used the full population sampling also known as the total population sampling. The total population sampling is a type of purposive sampling which involves the examining the entire population. Before using total population sampling, the researcher must ensure that the whole population has common characteristics or attributes.

The total population sampling was used only on the selection of the researchers (full time faculty members and the PhD students in the five universities). The application of this technique was based on the fact that the researcher was not in control of the population. The

118

researcher had to depend on gatekeepers to disseminate the questionnaire. The gatekeepers in this case were specific offices, departments or individuals in the universities who were to receive the link to the questionnaire and send to the participants. The researcher was not given direct access to the email addresses of the participants. In this case, the link was sent to the entire population of fulltime faculty members as well as the entire population of PhD students in the five universities.

4.6.3   Sample size

Regarding the need for a sample size calculation despite the mention of total population sampling, it's important to clarify that while total population sampling was employed for certain aspects of the study (i.e., distributing the questionnaire to all full-time faculty members and PhD students), it was not feasible to use this approach for all participants due to logistical constraints. For example, the researcher did not have direct access to email addresses and had to rely on gatekeepers within the universities to disseminate the questionnaire. Therefore, while total population sampling was used for some groups, a sample size calculation was necessary for other groups to ensure adequate representation and statistical validity.

To select the sample size for the study, the researcher used the SurveyMonkey sample size calculator (SurveyMonkey, 2021). The target sample is presented in Table 4-2. In the calculation, the researcher considered a confidence level of 80%, a margin of error of 5% and a z-score of 1.28, (which the number of standard deviations a given proportion is away from the mean). The formula used in calculating the sample size was;

$$Sample\ size = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + (\frac{z^2 \times p(1-p)}{e^2 N})}$$

Where:

N= population size

P (probability)=desired confidence level

e=error Margin

z = z-score

Table 4-2: Sample size for te study

	Universities	Fulltime lecturers	Postgraduate students (PhD)	Research/reference Librarians	University Librarians
1.	A.	98	76	1	1
2.	B.	74	75	1	1
3.	C.	33	46	1	1
4.	D.	16	39	1	1
5.	E.	70	87	1	1
Total		**291**	**323**	**5**	**5**

Table 4-2 exhibits the assembled sample of study participants. With the desired sample size achieved, the subsequent sections elucidate the researcher's data collection procedures and methods applied to the selected participants.

4.7   Data collection procedures and methods

The accuracy of data that have been collected determines the effectiveness of any research. Collected data and the procedures for collecting data have a great impacting in achieving the set objectives of the study. Muhammad (2016) describes data collection as the systematic technique of collecting information based on variables of interest in order to answer outlined research questions, test hypotheses, and evaluate outcomes. It is considered as a key stage in the research process because every study starts with a set of questions that must be answered or objectives that must be achieved by the end of the investigation. While research methodologies of different disciplines may vary, emphasis on honesty and accuracy in data collection cuts across the disciplines and consistency is maintained (Creswell & Clark, 2018). The reason for data collection is that the researcher aims at gathering quality data that after analysis would produce persuasive and dependable findings to the questions raised.

Depending on the research approach, an investigator decides on the kind of data to collect. In other words, the approach dictates the data collection methods to be used. This could either be quantitative or qualitative data. Ary et al. (2019) notes that the majority of quantitative data is collected using random sampling and structured data collection techniques; qualitative data are collected using non-structured data collection instruments. Data collection, therefore, requires thorough, correct, and systematic capturing of any potentially helpful information from respondents.

Muhammad (2016:208) identifies various methods used in collecting data and they include using, "questionnaires, interviews, focus group interviews, observation, survey, case-studies, diaries and activity sampling technique". While quantitative data collection involves the "use of a systematic standardised approach and employ methods such as surveys and ask questions" Muhammad (2016:203), qualitative data collection majorly uses focus groups, group discussions and interviews (Aborisade, 2013; Ary et al., 2019).

The researcher used both quantitative and qualitative data collection methods after adopting a mixed method approach and a mixed methods research design for the study. This enabled the collection of two data sets (qualitative and quantitative). As a result, the researcher collected data from participants through a questionnaire (see Appendix I) and interviews (see Appendices II and III).

4.7.1 Questionnaire

One of the data collection instruments that was adopted for the study, as indicated before, was the questionnaire. According to Muhammad (2016), a questionnaire is a research instrument that includes a set of questions and additional prompts designed to elicit responses from participants. They are widely used to collect quantitative data. However, as Muhammad (2016) points outs, they are not always designed to be used for statistical analysis of the collected data since some of the questions could require qualitative response. Many studies employ questionnaires because it is thought that they provide better anonymity, making it more likely to get correct data (Kumar, 2014). Furthermore, they are cost-effective, demand less effort from the surveyor compared to verbal or telephone surveys, and frequently provide standardized responses that facilitate straightforward data compilation (Muhammad, 2016).

When creating questionnaires, the researcher must guarantee that all the items captured in the questionnaire are simple and easy to comprehend (Kumar, 2014). Researchers are therefore advised to carry out a pre-test for the purpose of ensuring that there is clarity before the actual study.

In this study, the researcher created one set of questions for researchers (PhD students and faculty members). Both closed-ended and open-ended questions were used. Kumar (2014) stresses the importance of including both closed and open-ended questions in a questionnaire, stating that closed questions are highly effective in obtaining factual information, whereas open-ended questions are useful in eliciting survey participants' opinions, attitudes, and perceptions. Both aspects enrich and increase the value of the collected data.

121

Even with the adoption of a questionnaire as one of the data collection instruments for the study, the researcher was aware of some limitations of the instrument. For instance, as Kumar (2014:182) points out, when one uses a questionnaire, there is "lack of opportunity to clarify issues" for instance when some of the questions cannot be understood by the respondents. Because the concept of data literacy is relatively new, the researcher was aware of the possibility of this limitation occurring. To address this challenge, the researcher supplemented the questionnaire with extra details for questions that appeared complex, aiming to enhance respondents' comprehension. This included providing definitions for intricate concepts.

4.7.2 Interviews

Use of an interview schedule as a data collection tool in research "involves asking questions and getting answers from participants in a study" (Muhammad, 2016:211). An interview is designed to capture and analyse people's perspectives, experiences, beliefs, and ideas on a particular topic. When a researcher chooses to use an interview in a study to collect data, they expect, respondents to provide detailed information. When compared to quantitative methods such as questionnaires, one of the key characteristics of interviews is that they offer an in-depth understanding of a phenomenon. Interviews are also useful for gathering data, especially when a study is dealing with what could be considered to be sensitive issues that the subjects do not feel comfortable discussing openly in a group setting (Barrett & Twycross, 2018). Barrett and Twycross (2018) emphasize that, as an approach to qualitative study, interviews give the most direct and uncomplicated means of gathering comprehensive and in-depth data about a certain topic.

Research interviews can be classified into individual face-to-face interviews, and group face-to-face interviews (Muhammad, 2016). In order to conduct the interview, the researcher can choose to use an appropriate device (phone or any other electronic device, such as a computer) to reach the respondent and contact an interviewee. A researcher can decide to organise the interview as structured, semi-structured or unstructured (Giudice et al., 2019). The difference between the three is that a structured interview requires the researcher to develop a pre-set range of questions. In contrast, unstructured interviews give the researcher complete control over the content of questions. In this case, no questions are prepared beforehand (Kumar, 2014). Whether employing structured or unstructured interviews, the researcher must ensure that the questions are well-thought-out and accurately worded so that data acquired from it would help achieve the research objectives.

For this study, the researcher used semi-structured interviews. The researcher devised two distinct interview schedules, which were used to collect information from the five University Librarians and the five Research/Reference Librarians. The researcher conducted the interviews at the interviewees' workplaces. The respondent and the researcher both agreed on a suitable date and time for the interview. Using open-ended questions allowed the interviewer to elicit more information. As Kumar (2014) pointed out, using an interview will allow the researcher to obtain detailed information by probing and explaining to respondents some of the data literacy concepts that appear to be complex to grasp.

4.8   Data analysis

Analysis of data according to Bergin (2018:23) is a systematic application of a number of procedures and techniques (statistical and/or logical techniques) for the purpose of describing, summarizing and evaluating data. He adds that it "entails de-synthesising of data, information, or fact in order to respond to research questions". As one of the most important stages towards achieving the purpose of the study, its "goal is to reveal the underlying patterns, trends, and relationships of a study's contextual situation" (Albers, 2017:215). Analysis of the collected data depends on the research approach and research design adopted by the researcher. Ary et al. (2019:524) state that, because mixed methods research incorporates both qualitative and quantitative approaches, the fundamental analysis tools used in those approaches are still valid in mixed methods research. The fundamental distinction is in data integration.

Since the researcher adopted a mixed method research approach and a triangulation research design for the study, the analysis of the collected data was cognisant of the two sets of data (quantitative and qualitative) collected using questionnaires and interview techniques respectively. The researcher applied both quantitative and qualitative techniques in the analysis. Triangulation research design models influenced the process of analysis. However, Onwuegbuzie and Combs (2011), warn that analysing data from a mixed method study is challenging since it calls for combining of data from several tools in order to achieve meaningful results. Being aware of these challenges, in the analysis process, the researcher employed a side-by-side technique where quantitative data was analysed first, then qualitative data. Thereafter, the merging of findings from the two sets of approaches followed.

4.8.1   Quantitative data analysis

Quantitative data analysis entails "procedures and rules used to reduce large amounts of data into more manageable forms that allow one to draw conclusions and insights about patterns in

123

the data" (Scherbaum & Shockley, 2015). Closed-ended questions in the questionnaire generated quantitative data. The IBM SPSS statistical package version 21 was used to analyse quantitative data. This entailed creating data files, validating and processing the data, interpreting and analysing the data, and writing reports.

The data from the study was presented as descriptive statistics in the form of charts and tables, which show the various categorizations and correlations between the various data sets. Given the measurement scales in the data sets in the study, the researcher chose to use charts and tables accompanied by text to explain the findings.

4.8.2   Qualitative data analysis

Mihas (2019:99) concisely defines qualitative data analysis as "an array of tools of inquiry, including coding, memo writing, and diagrams". Ary et al. (2019:456) provide a much more and comprehensive description by pointing out that qualitative analysis "involves attempts to comprehend the phenomenon under study, synthesise information and explain relationships, theorise about how and why the relationship appear as they do and reconnect the new knowledge with what is already known".

In this study, some questions from the questionnaire and the interviews generated qualitative data which were analysed qualitatively using content analysis. Hsieh and Shannon (2005:1278) describes content analysis as a "a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns". Kumar (2014) defines it as a qualitative analysis technique that identifies emerging major themes from collected data based on participant responses to questions. Flick (2011:76) recommends its application in qualitative data analysis because it aids textual analysis as it "aims at classifying the content of texts by allocating statements, sentences, or words to a system of categories."

In the application of content analysis, the researcher identified particular content to be analysed. The researcher pointed out specific characteristics to be studied in relation to the themes and questions that were included in the open-ended questions and interview guide. This was followed by the interpretation of data by scrutinising it against the set themes. Subsequently, premised on the study's objectives, a discussion of the findings took place while incorporating quantitative results.

124

4.9 Validity and Reliability

The possibility of an error occurring in measurements in research is there. However, through validity and reliability, the researcher can identify the extent of the error and provide measures to minimize it. Ary et al. (2019:91) define validity as the degree to which test scores allow for "meaningful and appropriate interpretation" while reliability is the consistency with which results obtained from an instrument or a test measure what they are supposed to be measuring. Leedy and Ormrod (2020) describe validity and reliability by stating that, in research, it is the degree to which an instrument measures what it is supposed to measure, whereas reliability is the degree to which a measuring instrument yields the same results when the entity remains unchanged. Therefore, the score of validity and reliability of instruments employed in a study must be an area of concern for every researcher. In general, by opting for a mixed research method in this study, where various types of data collection procedures will be used, the researcher hopes to improve the study's validity and reliability. According to Zohrabi (2013), the use of different data collection methods and obtaining information from various sources can improve the validity and reliability of the data as well as their analysis.

4.9.1 Validity

Validity is an issue of trustworthiness, thus it's an important factor for assessing research's quality and acceptance (Creswell & Clark, 2018). It is concerned with whether the study is credible and accurate, as well as if it is testing what it claims to be testing. According to Zohrabi (2013) every researcher must ensure validity must be built into various phases of research. He states that this should be the responsibility of every researcher as well as the participants.

Researchers collect data using various instruments. It is critical to ensure the quality of the instruments used as a way of achieving validity because "the conclusions researchers draw are based on the information they obtain using these instruments" (Fraenkel & Wallen, 2012:158). As a result, it is critical that both the data and the instruments be validated. Kumar (2014) observes that validating of instruments can be done by demonstrating that the questions used are relevant to the study's objectives. The validity of the tools that the researcher used to gather information from the respondents were enhanced by generating open-ended questions from the stated objectives.

The researcher observed internal validity. Zohrabi (2013) explains that, internal validity is concerned with the correspondence of research findings to reality. This is also the extent to which the researcher examines and measures what should be measured. To achieve internal

125

validity, the researcher used triangulation, where data collection was carried out using a variety of techniques (questionnaires and interviews), thereby confirming the findings.

The study also achieved external validity which is the assurance of the findings' applicability in other contexts or with different subjects (Zohrabi, 2013). The focus of external reliability is on how generalizable the study is beyond the subjects under investigation. The researcher used a multiple-case study to accomplish this. The study, as previously stated, included five private chartered universities in Kenya. The study's findings were intended to be applicable to all Kenyan private chartered universities.

4.9.2 Reliability

As previously stated, the extent to which a research methodology produces consistent and accurate result is referred to as reliability. When a test is applied to the same element of measurement repetitively and consistently produces the same outcomes, it is said to be reliable (Ary et al., 2019). In a study, there are different ways of achieving reliability. One of the ways is what Zohrabi (2013) terms as the investigators position. He writes, "in order to increase the reliability of the research, the investigator needs to explain explicitly the different processes and phases of the inquiry" (Zohrabi, 2013:259). This can be achieved by elaborating at every stage of the study procedure and thoroughly describing the study's purpose, design, and participants. For this study, the researcher will ensure that all the stages are well elaborated in line with the rationale of the study.

4.10 Ethical considerations

It is significantly important to pay attention to any ethical issues in research for the purpose of ensuring integrity of the investigator and adherence and protection of the rights of the participants. Researchers must ensure that they employ a variety of ways in protecting the participant. Ethical implications in the process of research necessitate adherence to ethical norms in the study's planning, data collection and analysis procedures and use of the findings. Today, there is a greater emphasis on ethical conduct in research, reflecting current societal expectation of increased transparency and accountability (Pantzos et al., 2020; Zegwaard, Campbell, & Pretti, 2017).

Mertler (2019:46) points out "protection from harm, voluntary and informed participation, right to privacy, and honesty" as the four main categories of research ethics. He emphasizes that these ethical concerns apply to all research investigations, whether they use qualitative,

126

quantitative, or mixed research methods. This study, will be cognizant of the four categories of ethics before, during, and after data collection. Each of these concerns are discussed below.

4.10.1 Protection from harm

The researcher has the responsibility to ensure that all subjects in the research study do not suffer any physical or psychological harm. Loss of resource, including time, and reputation should be considered as facing eminent danger of harm during research (Fleming & Zegwaard, 2018). Fleming and Zegwaard (2018:211) advice that before choosing a research design for the study, the researcher should "consider the potential of harm to the participants, the researcher, the wider community, and the institution". To be noted is that, if a study is to include a community of vulnerable populations such as the elders, disabled and children, then the researcher must guarantee that they are given extra attention and be protected from any form of harm (Creswell & Creswell, 2018).

While evaluating potential harm, the goal should be to eliminate, isolate, and minimise the extent in descending order. Study participants must be fully informed about potential risks (Fleming & Zegwaard, 2018). Potential harm should be no greater than that experienced in daily life. Participants should not be exposed to dangers that are greater than or additional to those encountered in daily life.

This study did not include any vulnerable person. The participants who were included in the study were of adult age. Furthermore, in order to adhere to this consideration, the researcher avoided including sensitive and potentially traumatizing questions in the data collection instruments.

4.10.2 Informed consent

Informed consent is when participants agree to take part in a study after learning about its essence as well as what their participation will entail (Mertler, 2019). Denzin and Lincoln, (2011), term it as the cornerstone of an ethical research. The concept of informed consent can be divided into two for easy comprehension, that is, 'informed' and 'consent'. Fleming and Zegwaard (2018:210) stress that "participants must provide explicit, active, signed consent to taking part in the research, including understanding their rights to access to their information and the right to withdraw at any point. The informed consent process can be seen as the contract between the researcher and the participants".

In this context, "informed" means that the participants must be fully clear about what is expected of them, how the information they will give will be used, and what (if any)

implications may result. On the other hand, the term "consent" implies that participants must be given an expressive, active, signed agreement to participate in the study. The signed consent should acknowledge their rights to see their data and should allow them to freely withdraw from being a participant in the study at any point if they choose to (Creswell, 2014).

The researcher is always responsible for ensuring that participants are made aware of that the study entails and can independently choose whether to take part in the study or not. Potential subjects must need adequate time to ask questions about the study for them to make an appropriate decision (Creswell, 2014).

The researcher ensured respondents' consent in this study by providing enough information about the study to the participants. The consent form was included as the first item in the questionnaire and respondents had to fill it first before proceeding. They were also be given the opportunity to examine all options, including consenting to participate in the study or opting out before, during, or even after the study. In order to provide clarification, the researcher provided contact detail in case participants were interested in finding out more detail about the study. In this scenario, the researcher endeavoured to guarantee that the participants understood the information before proceeding. Finally, the researcher ensured that the participants have given their permission to take part in the study.

4.10.3 Anonymity and confidentiality

In research, the privacy of a participant must be respected and it is the responsibility of the researcher to adhere to it. According to Clark-Kazak (2017:13), "researchers have a duty to protect respondents' personal information and not disclose any identifying characteristics that would compromise anonymity…". This is why one of the ethical considerations the researcher had to adhere to was to ensure the confidentiality and anonymity of participants.

Anonymity in as an ethical consideration in research is when the researcher is oblivious of subjects' identities. In this case the identity of the participant is unknown to the researcher. When a participant's identification cannot be associated to the personal responses recorded in the collected data, then the participant's anonymity is safeguarded. According to Creswell (2014:138) in order to maintain or achieve anonymity, "in survey research, investigators disassociate names from responses during the coding and recording process. In qualitative research, inquirers use aliases or pseudonyms for individuals and places, to protect the identities of participants".

In the wake of the researcher not being able to guarantee anonymity, they must address confidentiality. Confidentiality is described as a situation whereby a researcher is fully aware of who the participant in the study is, however, they take precautionary measures hide their identity. It should be the researcher's handling of personal details in order to preserve the participants' privacy. According to Creswell (2014), one thing to keep in mind about confidentiality is that some respondents may wish to keep their identity hidden. By allowing this, the researcher lets the participants to retain full of their voices and exercise their decision-making independence. Maintaining confidentiality is therefore a crucial step to ensure the security of ones' information. Researchers use a different method to conceal the identities of their respondents.

To begin, when exchanging information over the internet, they use encryption key files to keep their records secure. They do not document information in such a way that subject responses are linked to personal information. This is typically accomplished by employing a code known only to them. Because participants may not even be identified by their names, but by other identifiers or combinations of information about them, researchers will only release aggregate findings to the public, rather than personal data.

In this study, the researcher ensured that the participants' identities were kept confidential and anonymous. The assurance extended beyond the confidentiality of the participants' names. Instead, it prohibited the use of self-identifying statements and information. For example, as shown in Tables 4.1 and 4.2, the researcher avoided identifying each institution that was included in the study alongside the specific population and sample. Instead, the researcher identified private universities using the alphabetical codes A-F. Furthermore, no data collection instrument included a component requiring participants to record their names or any other identifying information.

4.10.4 Honesty and compliance

It is the responsibility of the researchers to ensure a level of honesty whenever they are presenting their work. They must disclose any potential conflicts of interest. The reporting of data, results, methods, and procedures of the study should demonstrate honesty in research. Every researcher is expected not to make up, falsify, or misrepresent data. In this study, the researcher made certain that all sources were properly acknowledged. The researcher also made certain that the findings were accurately reported, with no tampering or addition of unreported data.

129

4.10.5 Ethics clearance process

Conducting research without the necessary clearance is considered an offence. It is a requirement by the University of Pretoria that a researcher must receive a research ethics clearance before conducting any research. The same is a requirement within the Kenyan context where a researcher is expected to seek clearance from the National Commission for Science, Technology & Innovation (NACOSTI) as well as the respective institutions where data is to be collected.

For this reason, before collecting data, the research applied for research ethics clearance from the University of Pretoria. Once the clearance had been given (see Appendix IV) the researcher applied for a research permit from the NACOSTI (see Appendix V). It is after receiving both the UP ethical clearance and the NACOSTI permit the researcher now wrote to the five institutions (Catholic University of Eastern Africa, United States International University, African International University, Adventist University of Africa, and Daystar University) requesting to collect data from participants in their respective institutions (see Appendices VI-XI).

4.11 Summary

This chapter described the methodologies proposed by the researcher for use in conducting the study. The pragmatic paradigm was identified as the study's philosophical worldview in the chapter. This discussion was followed by the selection of mixed methods research as the study's proposed research approach. A mixed methods research approach, according to the researcher, was ideal for achieving the study's objectives. Furthermore, the mixed methods approach was thought to be consistent with the pragmatic paradigm proposed as the study's philosophical worldview. The researcher chose the triangulation research design, which is a mixed method research design, during the discussion of possible research designs.

The study took place in the Nairobi metropolitan area and included five chartered private universities. Researchers (PhD students and faculty members at these universities), heads of university libraries in the selected libraries, and research or reference librarians in the libraries comprised the study's sample. The researcher employed two non-probability sampling methods: purposive and total population sampling. The researcher used a questionnaire to collect predominantly quantitative data and an interview schedule to collect qualitative data in this mixed methods study.

In addition to raising awareness of potential errors, the chapter outlined some of the measures the researcher needed to take in order to achieve validity and reliability. Concerning data analysis, the chapter stated that the researcher would consider the two sets of data that were to be collected; thus, the triangulation method of data analysis was used. Finally, the chapter outlined the ethical guidelines that the researcher was expected to follow prior to, during, and after the study.

CHAPTER 5

5. DATA ANALYSIS AND PRESENTATION – QUESTIONNAIRES

After detailing the study's procedures and methods in Chapter 4, Chapter 5 is dedicated to presenting the analysed data, with a specific focus on the information gathered through questionnaires.

5.1 Introduction

In this chapter, the researcher presents the findings from the questionnaire. Data collected were predominantly quantitative. However, the questionnaire included a few questions which were open ended. These questions resulted in qualitative data. Using statistical methods, the researcher examined the quantitative data to identify patterns and relationships among the variables. The goal of this analysis was to achieve specific objectives and answer research questions as identified in sections 1.5 and 1.6.

5.2 Preliminaries

A total of 546 responses were collected from the questionnaires, and each response was assigned a unique data ID from 1 to 546 (R1 to R546). Additionally, a timestamp was recorded for each response, indicating the moment at which the participant began filling out the online questionnaire. Participants were asked to indicate whether they provided informed consent for their data to be used in the research. Any data from participants who did not explicitly provide informed consent were considered ethically unsuitable for analysis and therefore were removed during the data cleaning stage prior to commencing with the data analysis.

The following five (5) data records were deleted for not giving consent:

Table 5-1: Records deleted - no consent provided

Data ID	Time stamp
329	11/26/2022 19:14:25
429	11/31/2022 20:25:07
446	12/5/2022 18:23:14
495	11/31/2022 18:49:04
515	12/5/2022 10:45:31

The following two (2) data records were deleted for leaving informed consent blank though proceeded to fill the questionnaire:

Table 5-2: Records deleted for leaving informed consent blank

Data ID	Time stamp
240	11/24/2022 15:02:15
391	12/1/2022 8:43:34

As a result of participants not providing informed consent, a total of 7 records were deleted. Following the deletion of these records, the analysis was conducted using a total of 539 records. Each record denotes each participant's response and was abbreviated as R, which stands for Researcher. This represents a return rate of 88%.

5.3 Section one: Biographical information

The participants were asked to indicate if they were PhD students/candidates or faculty members.
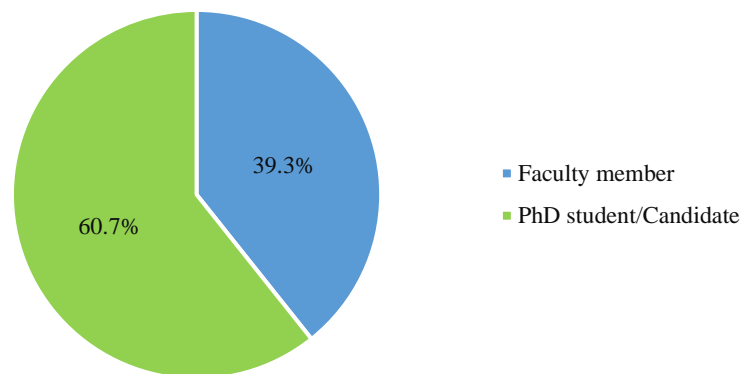
Figure 5- 1: Participant distribution

Out of the 539 participants in the study, 327 (60.7%) were PhD students or candidates, and 212 (39.3%) were full-time faculty members holding a PhD degree. The strong representation of both groups in the sample population is noteworthy, but since the majority of respondents are PhD students/candidates, it is acknowledged that the research findings may be biased towards their views.

5.4 Section Two: Research Data Management (RDM)

In this section the phases in the research data management cycle were presented to the respondents. Research Data Management (RDM) involves the organization, storage,

133

preservation, and sharing of research data throughout the research lifecycle. It includes practices for data documentation, data security, and data sharing, as well as adhering to ethical and legal considerations.

5.4.1 Data creation

First, the participants were asked to specify the various tasks where the library has provided assistance in the process of data creation. Data creation encompasses tasks such as developing a data management plan, selecting appropriate data formats, obtaining consent for data sharing, searching for pre-existing data, collecting data, and capturing data.
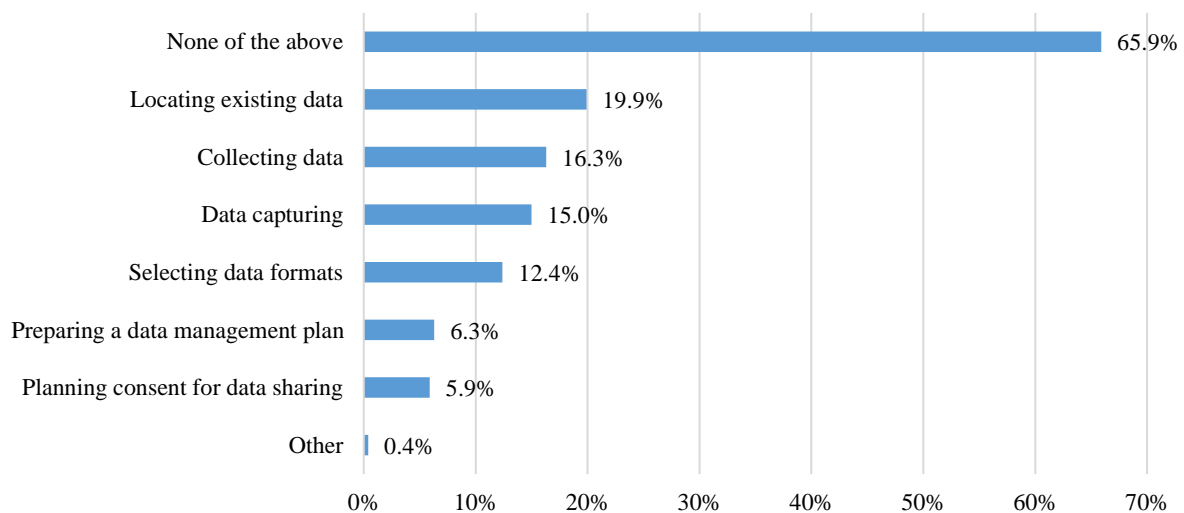
Figure 5- 2: Data creation

The majority of participants (65.9% or 355 respondents) felt that the library did not assist in creating data, as they chose the "none of the above" option. This suggests that respondents did not believe the libraries are able or willing to provide any of the data creation services that were listed in the questionnaire. Fewer than 20% of respondents were offered help, with data creation tasks, by the libraries. A small number of respondents (0.4% or two (2) respondents) mentioned other services such as storing data and providing online access to library services during the COVID-19 pandemic, which are not considered data creation services. The top services offered by libraries for data creation were locating existing data (19.9% or 107 respondents), collecting data (16.3% or 88 respondents), data capturing (15% or 81 respondents), and selecting data formats (12.4% or 67 respondents). All other data creation services were identified by less than 10% of the responses.

5.4.2 Data processing

Data processing involves several tasks such as data entry, data translation, data transcription, data validation, data cleaning, data anonymization, and data description. Participants in the survey were asked to indicate the type of support and assistance their library offered with regards to data processing.
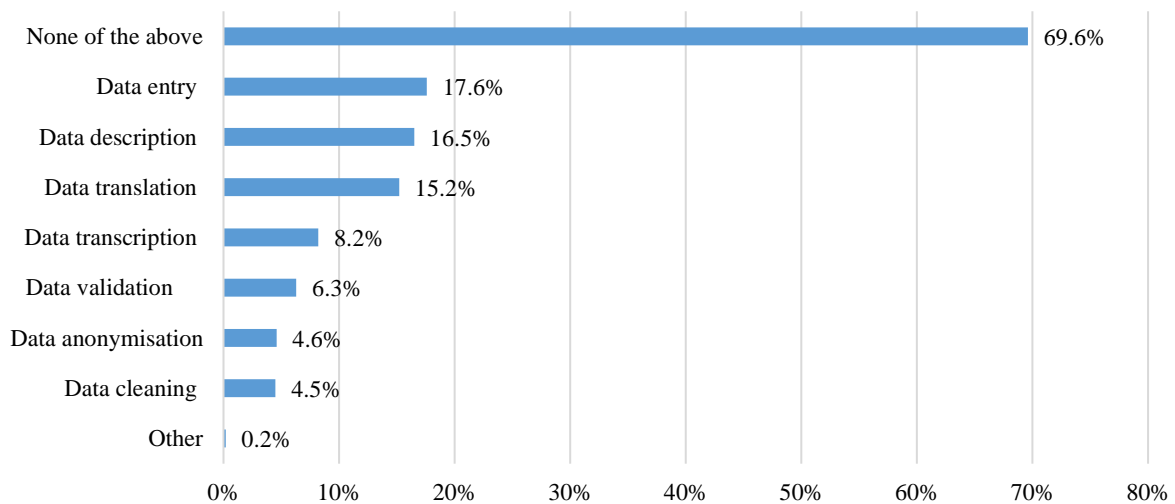
Figure 5- 3: Data processing tasks

The respondents, who made up the majority at 69.6% (375), felt that the libraries did not provide the listed data processing services. Among the services offered, those selected by over 10% of the respondents were data entry (17.6% or 95 respondents), data description (16.5% or 89 respondents), and data translation (15.2% or 82 respondents). For the remaining services listed, less than 10% of the respondents felt that libraries provided them. Only one respondent (0.2%) selected "other" and stated that they processed the data themselves without any assistance from the libraries.

5.4.3 Data analysis

Data analysis services encompass various tasks such as data interpretation, research output generation, author publications, data visualization, and data documentation preparation. As a result, the participants were requested to indicate the type of support and assistance their library offered concerning data analysis.
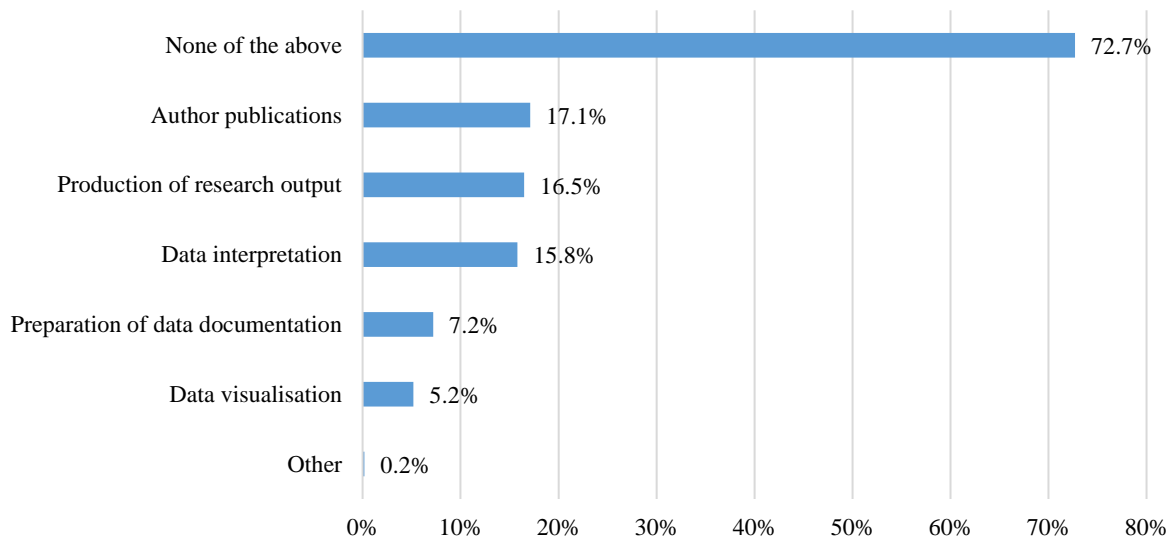
Figure 5- 4: Data analysis services

The survey revealed that a significant proportion of the respondents, (72.7% (392), believed that the libraries did not provide any of the data analysis services that were mentioned. From the services that were available, only author publications, at 17.1% (92), production of research output, at 16.5% (89), and data interpretation, at 15.8% (85), were offered to more than 10% of the participants. One respondent, representing a negligible percentage of 0.2%, reported receiving other types of data analysis services, but did not specify the nature of the services.

Based on the findings it appears a majority of the respondents believe that the libraries do not provide any of the data analysis services that were mentioned in the survey, this suggests that there may be a gap in the services being offered by libraries and the perceived needs of their users.

5.4.4 Preservation of data

Preserving data involves various activities, such as migrating data to suitable formats, backing up and storing data, creating metadata, archiving data in an open access repository, publishing data in the university's institutional repository, and publishing data in disciplinary repositories.
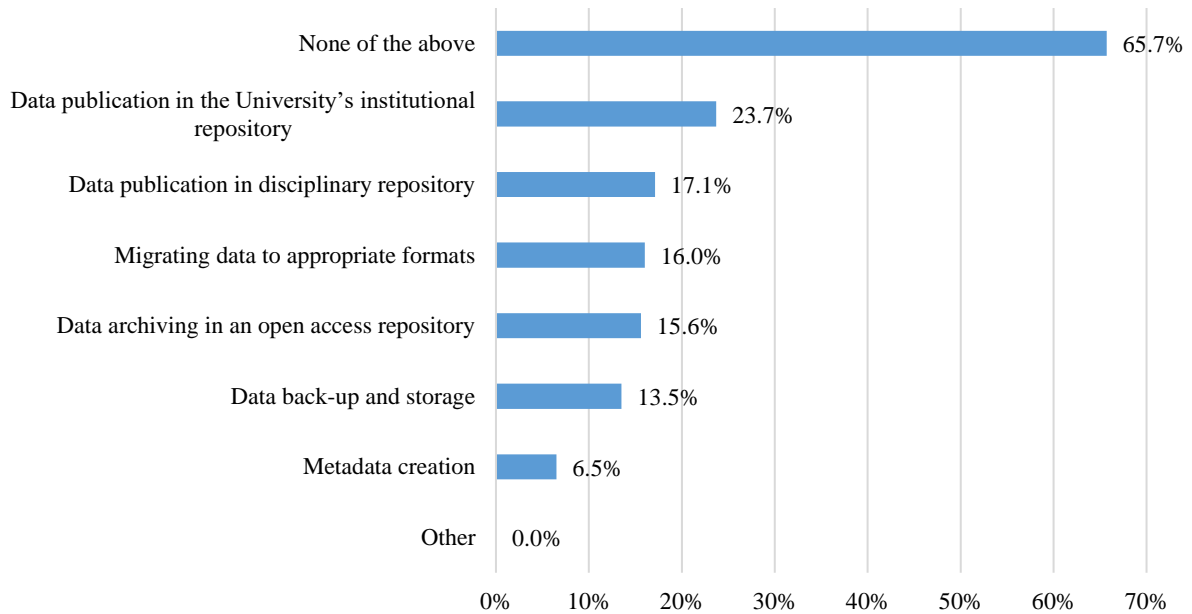
Figure 5- 5: Preservation of Data

The participants were requested to identify the type of support and assistance their library provided concerning data preservation.

Data preservation was one of the most widely acknowledged services provided by libraries, as all the services, except for metadata creation, were recognized by more than 10% of the respondents. Data publication in the university's institutional repository was the most recognized service, with a recognition rate of 23.7% (128). However, a substantial proportion of the participants, accounting for 65.7% (354), believed that libraries did not offer the preservation services that were listed.

5.4.5   Enhancing access to data

Enhancing access to data encompasses various activities, such as data sharing, data control, establishing copyright, promoting data, and creating a reference for citing data. The participants were requested to indicate the type of support and assistance their library offered concerning enhancing access to data.
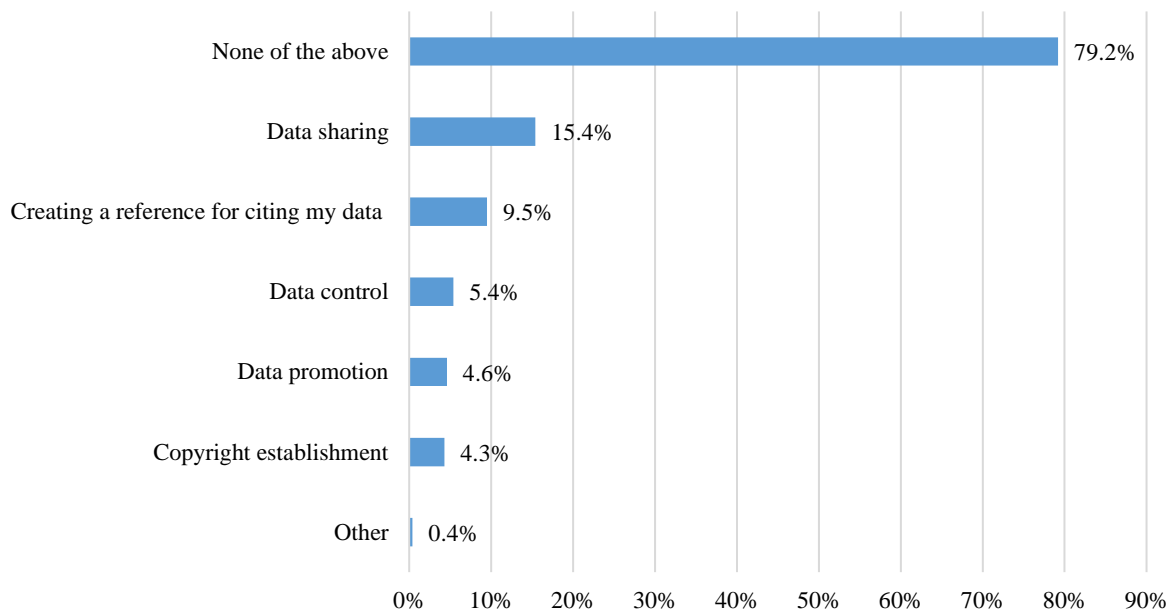
Figure 5- 6: Enhancing access to data

The data indicates that enhancing access to data was not considered a prominent service provided by libraries, with only data sharing being recognized by more than 10% of the respondents at 15.4% (83). A considerable majority, comprising 79.2% (427) of the participants, believed that libraries did not offer services related to enhancing access to data. Only a small number of respondents, representing 0.4% (1) reported receiving other services.

The one respondent mentioned receiving a plagiarism check which is not a related service in the context of enhancing access to data.

5.4.6   Created a Data Management Plan (DMP)

A data management plan (DMP) is a document that outlines the steps taken to manage the data generated, collected, or used during a research project. DMP includes information on the types of data to be collected or generated, how the data will be organized and stored, how data will be backed up and preserved, and how data will be shared and made accessible to others (Gajbe, Tiwari & Singh, 2021). The DMP also covers issues such as data privacy and security, ethical considerations, and legal compliance with relevant regulations.
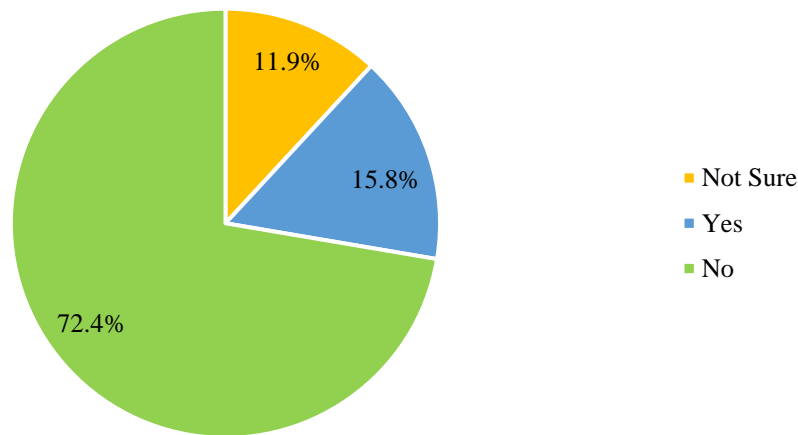
Figure 5- 7: Creating DMPs

In this section of the questionnaire, participants were asked to indicate if they had ever created a data management plan.

Out of the total respondents, only 15.8% (85) indicated that they have ever created a data management plan, while 72.4% (390) were certain that they had not created one. Notably, 11.9% (64) of the respondents were uncertain whether they had created a data management plan or not. This uncertainty might suggest that some respondents lack knowledge about the concept of a data management plan, or they are unsure if their data management practices meet the criteria for a data management plan.

There was a follow-up question asking respondents to comment on their choices regarding the creation of data management plans. The analysis of responses identified some key themes as to why participants had or had never created data management plans. These themes were, lack of awareness about data management plans, DMP being a requirement for research that they were carrying out, lack of personal experience, lack of knowledge about DPMs, lack of institutional support, lack of understanding about the importance of a DMP and future plans of creating DMPs.

The majority of the respondents, 312, who had indicated that they had never created a DMP gave the reason that they lacked awareness about data management plans. All the respondents, 15.8% (85) who had indicated that they had ever created a DMP gave the reason that DMP was requirement for research that they were carrying out. The requirement was either by the research funder or the publisher. Other respondents indicated that they had no personal experience in creating DMP (125), they lack knowledge about DPMs (162), they lacked

139

institutional support in creating DMPs (96) and they never understood the importance of a DMP (89). Lastly, some respondents (89) planned to learn and create a DMP in the future, highlighting the importance of training and education on data management practices.

In summary, the analysis of respondents' comments revealed that various factors, including the level of knowledge, awareness, experience, and institutional support, influence the uptake of DMPs. Furthermore, the findings suggest a need for more training and education to promote the creation and adoption of DMPs in research projects.

5.4.7 Created any metadata

Metadata are data that describe other data. It provides a structured and standardized way of describing data and its characteristics, including the format, location, size, creator, date of creation, and other details relevant to the data's content and context.

The participants were asked to indicate if they had ever created metadata.
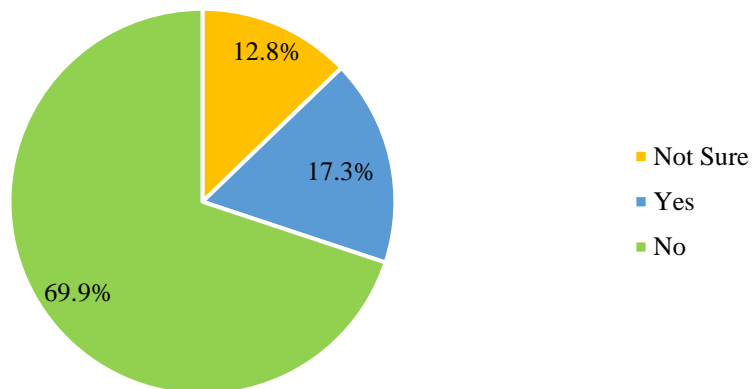
Figure 5- 8: Figure 5- 9:Creation of a metadata

Out of the total number of respondents, only 17.3% (93) were confident that they had created metadata, while 69.9% (377) were sure that they had not. It is possible that the 12.8% (69), who were uncertain about whether they had created metadata or not, may have felt unsure because of a lack of understanding of what metadata is in the context of information and data literacy.

The low percentage (17.3%) of respondents who confirmed they had created metadata suggests that there may be a lack of awareness or understanding of the importance of metadata in data management. This could have implications for the discoverability, reuse, and sharing of data, particularly in research settings where data are often shared and reused by others.

There was a follow-up question which required respondents to share their comments on the choices they made regarding metadata. This second part of the question was analysed qualitatively and some underlying themes emerged from their responses. One of the primary themes was awareness, with the majority of respondents (321) indicating not having a good understanding of metadata and its significance in their research work.

Regarding the creation of metadata, some respondents (93) reported having created metadata during their research. Those who created indicated that the creation of metadata was a requirement for publishing their research, both internationally and in journals.

The lack of knowledge emerged as another significant theme, with majority of the respondents (352) reporting not having encountered metadata in their research work or not knowing how to create it. This lack of knowledge could have implications for the consistency and accuracy of data, particularly if metadata is not created and managed correctly.

5.4.8   Where research data is stored

The participants were asked to indicate the medium in which they stored their research data.
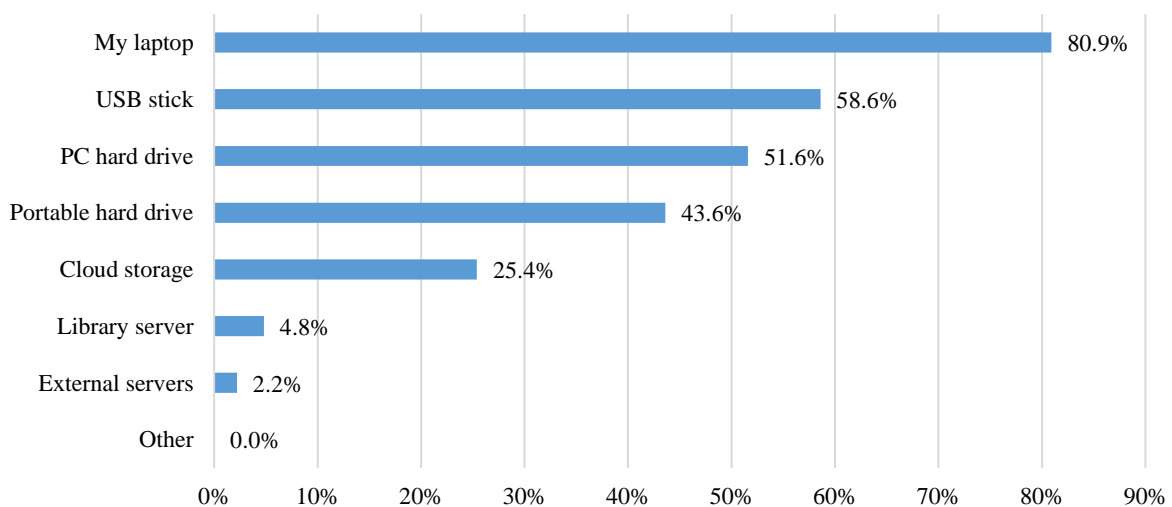
Figure 5- 10: Research data storage

Respondents had the opportunity to choose more than one medium.

According to the survey results, the primary storage medium used for research data is a personal laptop, with a significant majority of 80.9% (436) of respondents using this method. Following this, USB sticks were the second most popular choice, with 58.6% (316) of respondents using them, followed by PC hard drives at 51.6% (278) and portable hard drives at 43.6% (235). The

141

data shows that respondents prefer physical storage devices that they have control over, as evidenced by the popularity of USB sticks and hard drives. Despite this, virtual storage is a popular choice, with cloud storage being the most popular virtual option at 25.4% (137).

However, it is important to note that cloud storage usage lags far behind physical storage usage. Library servers were a poorly rated choice for storage, with only 4.8% (26) of respondents choosing this method. External servers were also not popular, with only 2.2% (12) of respondents selecting them.

Given the clear preference for physical storage over virtual, it would be beneficial to investigate the reasons behind this trend. Further analysis could provide insight into why respondents favour physical storage devices, and could potentially inform strategies for improving the uptake of virtual storage options considering that virtual storage makes it for data sharing.

5.5   Section Three: Data management competence

This section provides the results of a question that aimed to assess the perceived competence of the respondents in various areas of research data management. The survey asked respondents to rate their knowledge, skills, and abilities in the following areas: data planning, data collection, data processing, data analysis, data preservation, data sharing/publishing, data re-use, FAIR data, ethical collection of data. It is important to note that research data management is an integral part of the research process. Effective management of research data ensures that data is properly organized, stored, preserved, and made accessible for future use. Furthermore, adherence to ethical data collection practices is crucial in ensuring that data is collected in a responsible and trustworthy manner.

The following figure displays the results of the survey, showcasing the percentage of respondents who rated their competence in each area as "very competent," "somewhat competent," " little competence," " no competence" or "not applicable."

	Data planning	Data collection	Data processing	Data analysis	Data preservation	Data sharing/ publishing	Data re-use	FAIR data	Ethical collection of data
■ No Competence	7.4%	0.4%	1.7%	1.9%	7.2%	15.0%	39.3%	57.7%	26.3%
■ Little Competence	46.8%	20.0%	25.6%	27.8%	46.4%	53.4%	39.5%	26.2%	25.2%
■ Somewhat Competent	35.6%	51.8%	51.2%	53.1%	37.5%	26.3%	17.1%	11.1%	32.3%
■ Very Competent	10.2%	27.6%	21.3%	17.3%	8.9%	5.2%	3.0%	2.2%	15.6%
■ Not Applicable	0.0%	0.2%	0.2%	0.0%	0.0%	0.0%	1.1%	2.8%	0.6%

Figure 5- 11: Level of agreement with data related statements

A significant proportion of respondents reported having only limited or somewhat competent skills in data management. Specifically, 82.4% (444) of respondents rated themselves as having little or somewhat competent skills in data planning, 71.8% (387) in data collection, 76.8% (414) in data processing, 80.9% (436) in data analysis, 83.9% (452) in data preservation, 79.7% (430) in data sharing/publishing, 56.6% (305) in data re-use, 37.3% (201) in FAIR data, and 57.5% (310) in ethical collection of data.

143

Notably, the percentage of respondents who rated themselves as having no competence in data re-use, FAIR data, and ethical collection of data were 39.3% (212), 57.7% (311), and 26.3% (142), respectively, which is significantly higher than the proportion of respondents who rated themselves as having no competence in other areas of data management, which did not exceed 15%. Moreover, 57.7% of respondents reported having no competence in FAIR data, with 83.9% (452) of respondents reporting limited or no competence in this area. Similarly, 78.8% (425) of respondents had limited or no competence in data re-use. Only data collection at 27.6% (149) and data processing at 21.3% (115) were the areas where more than 20% of the respondents rated themselves as highly competent.

Based on the findings, it appears that research data management practices in the institutions surveyed may be inadequate or insufficient. Given the importance of proper research data management in ensuring the integrity, quality, and accessibility of research data, these findings highlight the need for institutions to invest more in developing the skills and knowledge of their researchers in the area of research data management.

5.6 Section Four: Research process

Within this section, individuals were presented with a series of statements outlining potential means by which the library can furnish researchers with data literacy-related assistance during the research process. Respondents were tasked with indicating their level of agreement with each statement by selecting one of the following options: "Agree," "Disagree," or "Not sure."

5.6.1 Managing data

Participants taking part in the study were requested to provide their viewpoint on how libraries handle research data. This was accomplished by indicating agreement or disagreement with a series of statements related to the management of research data.

144
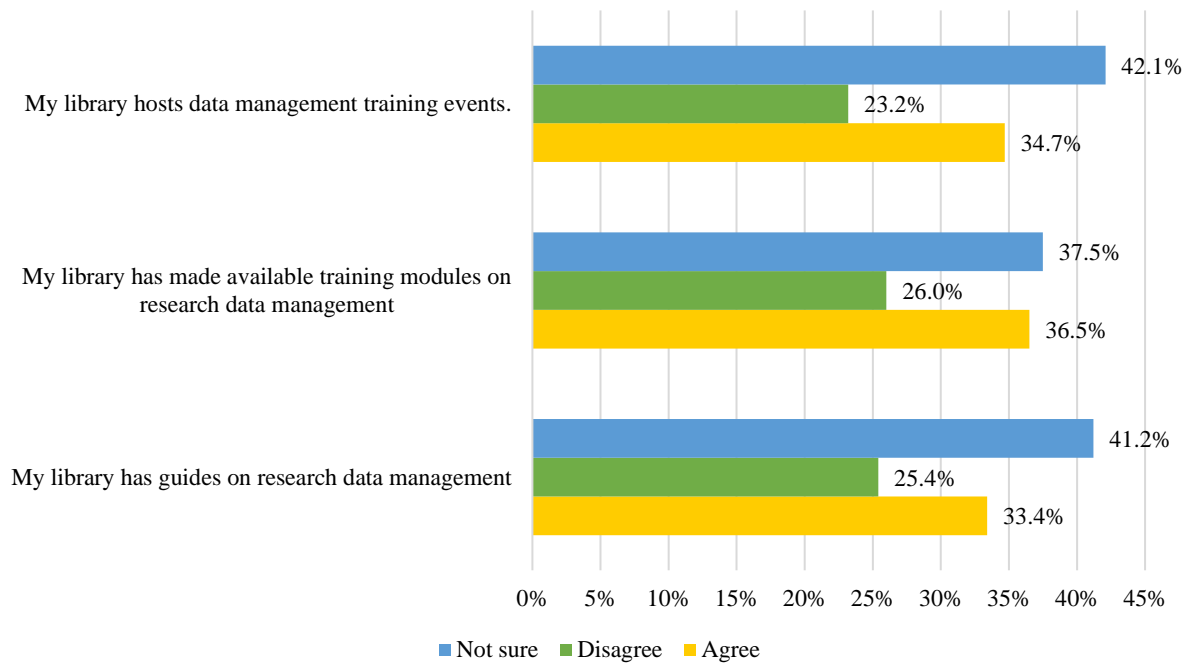
Figure 5- 12: Library in promoting data literacy data

The findings indicate that a majority (66.6% (359) of respondents were dissatisfied with the level of assistance provided by libraries in managing research data. The respondents disagreed or were uncertain about whether their libraries offered guides on research data management, 63.5% (342) disagreed or were uncertain about whether their libraries offered training modules, and 65.3% (352) disagreed or were uncertain about whether their libraries hosted data management training events. However, many respondents (30%) expressed satisfaction with their library's data management services. This aspect should be further explored to determine the underlying reasons and potential ways to improve library services.

5.6.2 Data publishing and sharing

Participants taking part in the study were requested to provide their viewpoints on how libraries handle research data. This was accomplished by indicating agreement or disagreement with a series of statements related to data publishing and sharing.
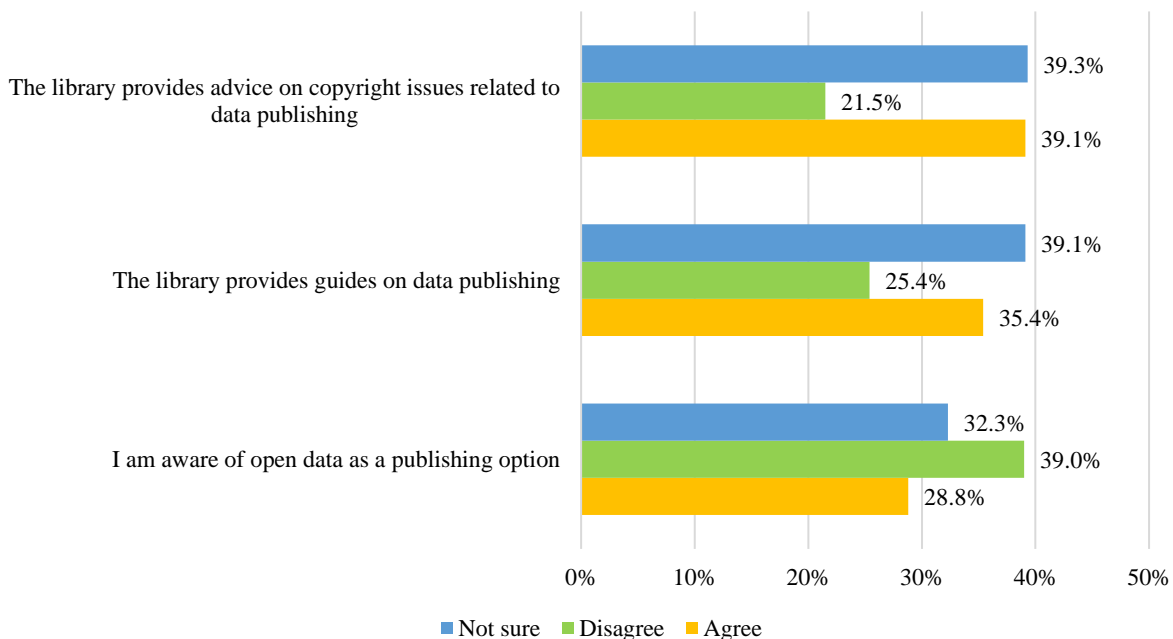
Figure 5- 13: Data publishing and sharing

The findings reveal that a majority of respondents were dissatisfied with the level of assistance provided by libraries in data publishing and sharing. Specifically, 71.3% (384) of the respondents were unaware (disagree and not sure) of open data as a publishing option, 64.5% (348) were unaware (disagree and not sure) of the libraries providing guides on data publishing, and 60.8% (328) were unaware (disagree and not sure) of the libraries providing advice on copyright issues related to data publishing.

However, the survey also found that over 28% of the respondents acknowledged the provision of data publishing and sharing services by libraries. This suggests that libraries have made some progress in this area, but more needs to be done to increase awareness and understanding among library users. The results of the survey reveal a gap in knowledge about data publishing and sharing among the majority of researchers.

5.6.3 Tools

Participants taking part in the study were requested to provide their viewpoint on the tools used by libraries to manage research data. This was accomplished by indicating agreement or disagreement with a series of statements related to data management tools.
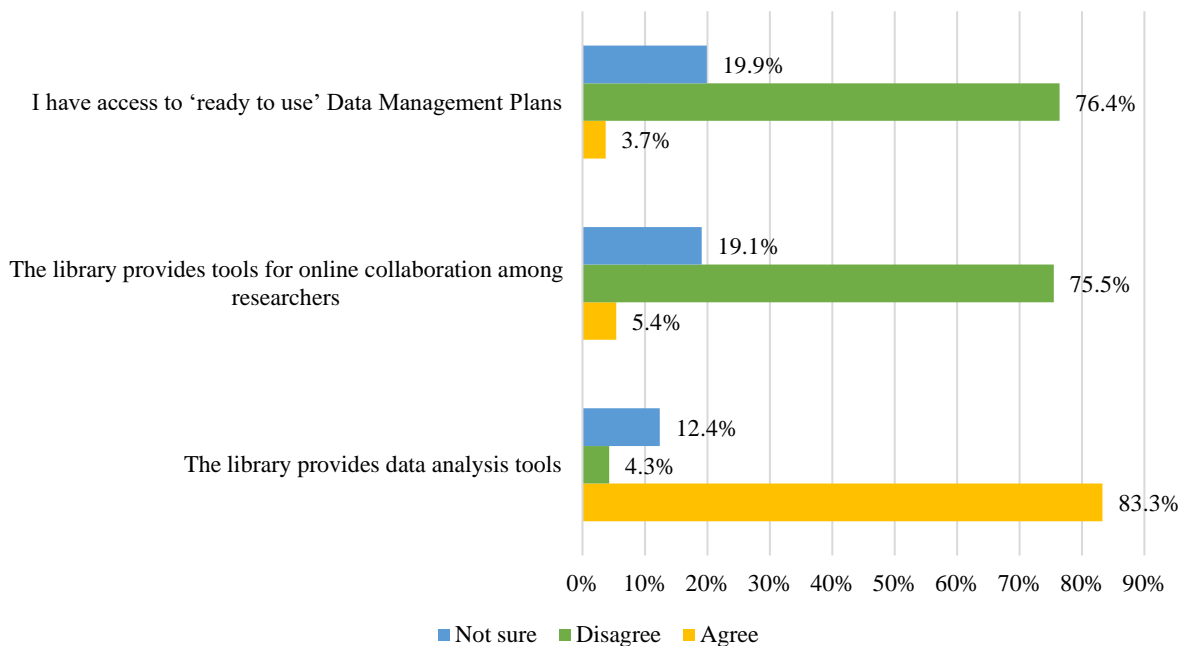
Figure 5- 14: Data management tools

The majority of respondents expressed dissatisfaction with the level of access to certain data management tools offered by libraries. Specifically, 96.3% of respondents indicated that they did not feel that libraries provided adequate access to 'ready to use' data management plans, while 94.6% indicated a lack of access to tools for online collaboration among researchers. In contrast, a significant majority agreed that libraries do provide data analysis tools, with 83.3% indicating their agreement.

This indicates that there may be a need for libraries to make available necessary data analysis tools in order to better meet the needs of their users.

5.7   Section Five: Importance of data literacy

Data literacy refers to the ability to read, analyse, interpret, understand, and communicate data in a meaningful way. It encompasses a broad range of skills and knowledge, including the ability to collect, manage, and process data, as well as the capacity to use data to make informed decisions, draw conclusions, and solve problems. Data literacy also involves an understanding of basic statistical concepts and data visualization techniques, as well as the ability to critically evaluate data and its sources. In today's data-driven world, data literacy is becoming an increasingly important skill for individuals and organizations to possess.

147

In this section, the research participants were asked an open-ended question to express their opinions on the importance of data literacy. Based on their responses, the researcher drew out some related themes.

i. Data literacy is an essential aspect for researchers as it improves data management and promotes research outcomes.

Improves data management in research (R469)
Improves data management (R507)
Promotes research outcomes (R11)

ii. Data literacy enhances the researchers' data competencies and equips them with the necessary skills for effective data management hence reducing or shortening time spend while working with data as well as the cost of working with data.

It saves time for the researcher, and cost that could be incurred in hiring a data management (R3).

It equips researchers with skills necessary to ease their researcher work (R101).

Empowers researchers to improve on their data competency necessary for their research (R459).

iii. Data literacy also builds a researcher's capacity to carry out research, communicate results, and determine valuable materials and instruments for data management.

It enables a researcher to possess the necessary skills required to understand, explore, use and communicate using data (R193).

Data literacy improves our ability to read, create, understand and communicating data (R428)

It equips the researcher with the knowledge of data management (R423)

It helps in carrying out research effectively (R424).

It enhances proper communication between the researchers and expertise (R427).

It promotes data management efficiency (R433).

It prepares researchers for successful research data management (R460).

iv. Data literacy improves efficiency in working with research data, enables effective data analysis, and helps promote data privacy, confidentiality, and reuse.

It enables one to be able to collect data, to analyse data and to interpret data in a more accurate way (R396).

It improves data efficiency (R405).

148

v. Data literacy a necessity for quality research and effective decision making, improving their data analysis skills and enabling meaningful and critical interaction with data.

Enhances effective data management process (R439)

Enhances effective research data management (R13).

Data literacy is a prerequisite for researchers to interact with data insights meaningfully and critically (R272).

5.8 Level of agreement with the statements

Participants were requested to express the level of agreement with a set of statements provided in this section. Respondents had the option to select from a five-point Likert scale that ranged

149

from "Strongly agree" to "Strongly disagree", with intermediate options such as "Agree", "Neither agree nor disagree", and "Disagree".
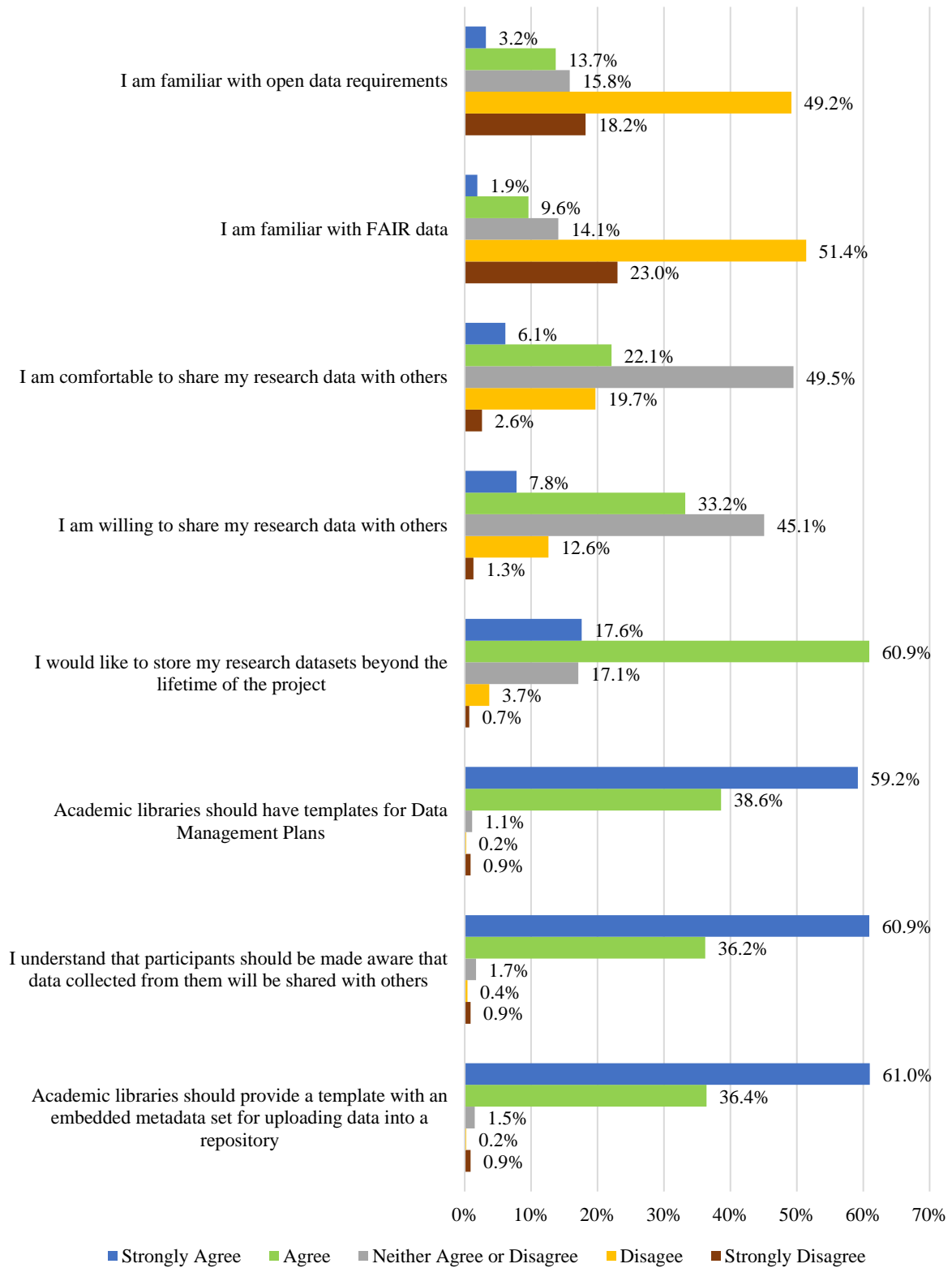
Figure 5- 15: Level of agreement with data related statements

The analysis revealed a significant lack of familiarity with open data and FAIR data, with 67.4% (363) and 74.4% (401) respectively disagreeing or strongly disagreeing that they were familiar with these concepts. Only 16.9% (91) of respondents reported some level of familiarity with open data, and 11.5% (62) with FAIR data.

Regarding the comfort and willingness of respondents to share their research, 49.5% (267) and 45.1% (243) respectively indicated a neutral stance, while only 28.2% (152) were comfortable with sharing their data and 41.0% (221) were willing to share their research data.

A large majority of the respondents, 78.5% (423), expressed a desire to store their research datasets beyond the lifetime of the project, indicating an appreciation for the value of long-term data preservation.

The respondents strongly felt that libraries should provide templates for data management, with 97.8% (527) advocating for templates for data management plans and 97.4% (525) for a template with an embedded metadata set for uploading data into a repository. This suggests that there is a need for clarity and guidance on best practices for data management and metadata development that is tailored to their specific university and research needs.

Additionally, an overwhelming majority of the respondents, 97.1% (523), believed that participants should be made aware that the data collected from them will be shared with others, indicating a strong preference for transparency and ethical data sharing practices.

Overall, the data suggests that there is a need for greater education and support in the area of data management and sharing, as well as a need for greater transparency and communication about the collection and sharing of research data.

5.9 Section Six: Research data management challenges

In this section, the respondents were asked to identify some of the challenges they face when working with data during research.
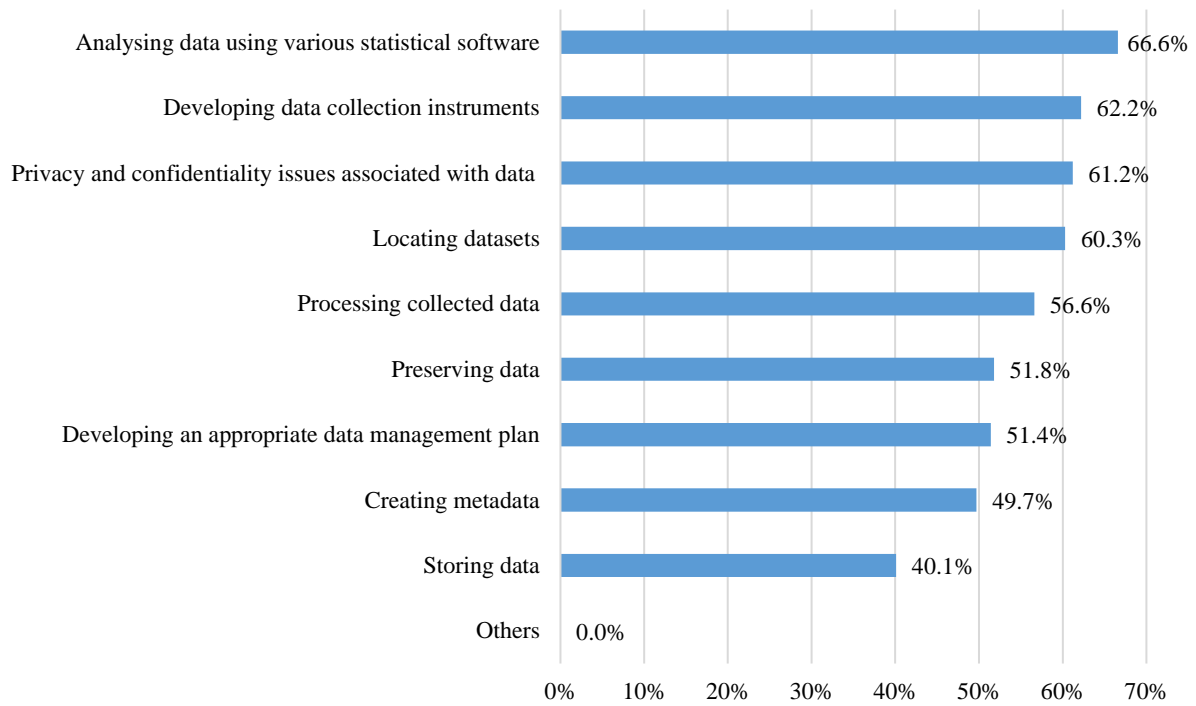
Figure 5- 16: Research data management challenges

In this section, the participants were asked to identify all the challenges they faced while working with research data. The respondents rated the challenges in order of significance, with the most challenging being the analysis of data using various statistical software at 66.6% (357), followed by the development of data collection instruments at 62.2% (355), privacy and confidentiality issues associated with data at 61.2% (330), locating datasets at 60.3% (325), processing collected data at 56.6% (305), preserving data at 51.8% (279), developing an appropriate data management plan at 51.4% (277), creating metadata at 49.7% (268), and storing data at 40.1% (216).

This prioritized list of challenges can guide efforts to address the issues associated with managing research data among researchers.

5.10 Section Seven: Data literacy framework

To answer the questions in this section, respondents were provided with a proposed data literacy framework and were required to refer to it.

The first question asked was if they were familiar with the literacies listed in the framework.
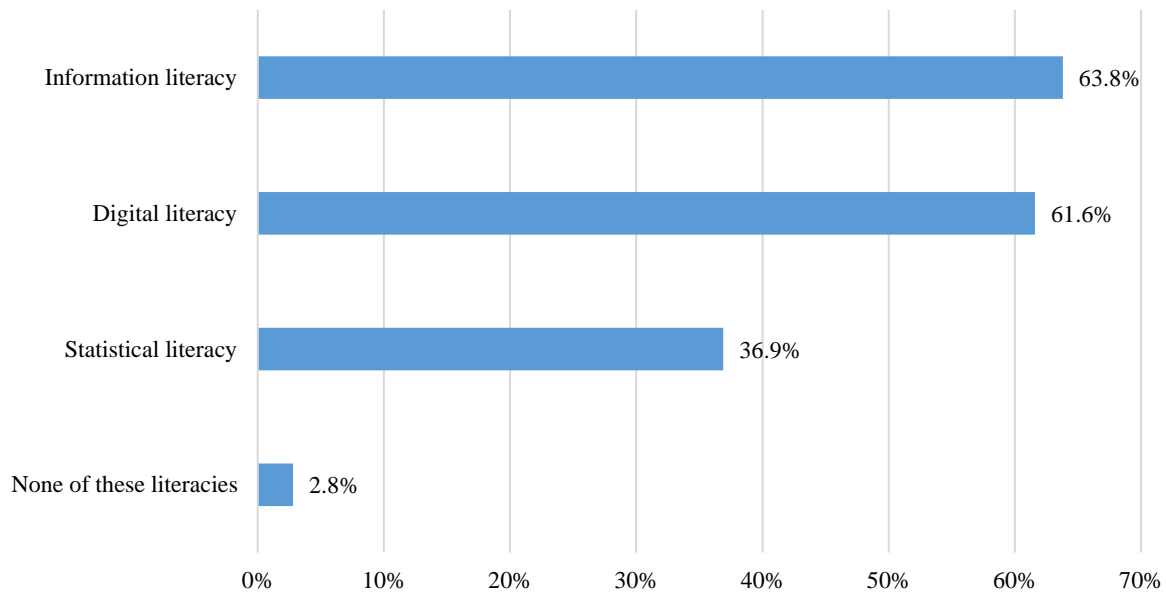
Figure 5- 17: Familiar with other literacies

In terms of familiarity, the respondents were most familiar with information literacy with 63.8% (344) indicating that they were familiar, followed by digital literacy at 61.6% (332). Statistical literacy had the lowest level of familiarity with only 36.9% (199) of respondents indicating that they were familiar with it. About 2.8% (15) of respondents were unfamiliar with all the listed literacies. The disparity between statistical literacy and the other literacies requires further investigation, as does the fact that over 30% of respondents appear to lack proficiency in the various literacies.

5.11   Key areas to be prioritised in the development of a data literacy training program

The respondents were asked to identify key areas in research data management that they believe should be given priority in the development of a data literacy training program. As an open-ended question, a qualitative thematic approach was used to analyse the responses.

Based on the responses, the key areas that should be prioritized in the development of a data literacy training program were data organization and management, data privacy and security, data visualization and presentation, data analysis and statistics, data ethics and governance, and data sharing and collaboration. These themes were identified as a result of the qualitative analysis of the responses. The following is a list of the key areas and sub key areas as analysed

from the responses. The number in the brackets indicate the number of respondents who suggested that as a key area to be prioritized. Respondents could select/indicate as many as they needed to.

1. Data organization and management
   - Data management planning (36 respondents)
   - Data preservation (17 responses)
   - Developing data processing tools and methods (10 responses)
   - Data collection methods (21 responses)
   - Creating metadata (20 responses)
   - Developing appropriate data management plan (58 responses)
   - Data storage (37 responses)
   - Capturing data (11 responses)

2. Data privacy and security
   - Data privacy (10 responses)
   - Ethical data collection (17 responses)
   - Privacy and confidentiality issues associated with data management (10 responses)

3. Data visualization and presentation
   - Data visualization (33 responses)
   - Data presentation (6 responses)
   - Data interpretation (39 responses)
   - Data communication (8 responses)

4. Data analysis and statistics
   - Data processing (68 responses)
   - Data analysis (38 responses)
   - Digital literacy (38 responses)

5. Data ethics and governance
   - Ethical collection of data (27 responses)
   - Establishing copyrights (9 responses)
   - Open science (20 responses)

6. Data sharing and collaboration
   - Data reuse (37 responses)
   - Open data publishing: (24 responses)
   - Data publication (22 responses)

- Data sharing (29 responses)
- Information literacy (6 responses)

It will be important to note that the number of responds in brackets for each category is not indicative of the urgency of each theme when developing a data literacy training program. Instead, the figures represent how frequently each sub-theme was stated by research participants. As a result, a higher number of responses indicates that the sub-theme was cited more frequently by respondents, but it does not necessarily imply that it is more urgent.

It is also worthwhile to point out that data literacy training priorities may differ based on the specific context and demands of the target audience. As a result, before identifying the most pressing areas to prioritize, it is critical to assess the precise aims and objectives of the training program, as well as the target audience's level of data literacy.

5.12 Library's role in promoting data literacy among researchers

Respondents were asked to indicate whether they think the library has a role to play in promoting data literacy among researchers.
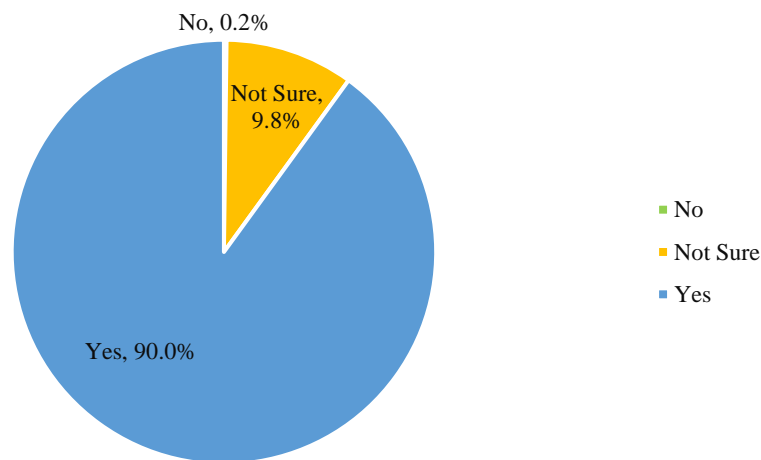
Figure 5- 18: Role of library in promoting data literacy

The findings indicate that an overwhelming majority of the respondents, accounting for 90% (485), acknowledged the role of libraries in promoting data literacy among researchers. The high percentage is indicative of a clear and resounding recognition that libraries serve as key provider in fostering data literacy. This is in contrast to the 0.2% (1) of the respondents who felt that libraries have no role in promoting data literacy, and the 9.8% (53) who were unsure of the role of libraries in data literacy.

The respondents were asked to provide their comments regarding their choice of the library having a role to play in promoting data literacy among researchers. Since this was an open-ended question, the researcher employed thematic techniques for analysis.

From the responses, the following themes were identified:

1. Role of the library: from the responses it was indicative that the library is already playing other key roles that could be related hence deserving a chance to play a role in promoting data literacy among researchers.

   The main goal of a library is promoting research, and there is no research without data. so, the library should ensure the researchers know and understand how to manage their data (R1)

   The library plays a pivotal role in accelerating research productivity, so it is their responsibility to ensure that their patrons have the necessary competencies in data management (R2).

   Data literacy is related to information literacy and knowledge management which are the responsibilities of the library (R6).

   The library offers space for knowledge promotion which should include data literacy (R32).

2. Resources: some of the responses indicated that the library already has necessary resources that could make it play a key role in promoting data literacy among researchers

   The library has the necessary resources and competencies for research data management (R504).

   The library has the right infrastructure and personnel to promote data literacy (R?).

   The library is equipped with the right resources and staff have the right competencies for data literacy (R26).

   The library has ICT infrastructure and personnel with the right data skillset for data literacy trainings (R459).

3. Human Capital: another reason pointed out by respondents as to why the library should play a role in promoting data literacy among researchers is that the library has the necessary human capital (librarians) who can train researchers in data literacy.

It [the library] *has relational capital between the personnel and the researchers (R83).*

Librarians are highly equipped with data management process (R429)

Librarians have the right data skillset and general knowledge for research (R20)

The library staff have the right competencies for data literacy (R394)

4. Research Support: Some of the respondents mentioned the fact that the library is already providing research related services.

The main goal of a library is promoting research, and there is not research without data. so, the library should ensure the researchers know and understanding how to manage their data (R4).

The library supports research through information resources and personnel (R43).

5. Other aspects: From the responses of the research was able to pick out some aspects. In this case the researcher summarises them as most of the were either one or two words mentioned hence not able to make a direct quotation.
    - The library has a variety of resources, both physical and digital.
    - The library can guide researchers on copyright issues and assist in locating and setting up data collection instruments.
    - Librarians have a good understanding of data literacy and can play a key role in training.
    - The library is resourceful for research development and has the intellectual capital.

In summary, the responses indicate that the library has a significant role to play in promoting data literacy among researchers. Many respondents believe that the library has the resources, technology, and personnel needed to support researchers in their work with data. They mention the availability of a variety of resources, high-tech infrastructure, knowledgeable librarians, and easy accessibility as key factors that contribute to the library's ability to promote data literacy. Some respondents also emphasize the importance of the library in preserving data,

promoting data reuse, and improving data discoverability. In conclusion, many respondents believe that the library is well-positioned to promote data literacy among researchers.

5.13   Respondents' opinions on what made sense in the proposed framework

After studying the proposed data literacy framework for successful data literacy training, the respondents were asked to identify what they felt made sense.

Based on the responses given, the majority of the respondents found the proposed data literacy framework to be clear and comprehensive, with no major concerns or issues. Some direct statements from the respondents included:

Everything makes a lot of sense (R4).

The framework is clear and comprehensive enough (R8).

The framework, as a whole, makes sense (R501).

Some respondents, however, mentioned that they did not understand certain concepts, such as H-Index scores, FAIR data, mediating variables, and relationship capital.

I do not understand H-Index score (R13).

I do not understand FAIR data (R22).

What is the meaning of mediating variables? (R82).

There were also a few comments regarding the role of stakeholders and the use of bibliometrics as a motivator for data literacy. Some respondents suggested that the framework could be elaborated on further, or that other stakeholders should be included in the conversation about data literacy. Additionally, a few respondents expressed concerns about the availability of resources and the need for research ethics training. Overall, the responses indicate that the framework is generally well-received, although there may be some areas that could be clarified or expanded upon.

5.14   What respondents liked about the framework

Having had a chance to look at the proposed data literacy framework for successful data literacy training, the respondents were asked to identify what they like about it. Various responses were given by the respondents and the researcher distilled the following themes. Although some themes are supported by direct quotations from participants' responses, others are not due to

158

the fact that some responses contained only a few words. The researcher had to interpret the meaning of these words to create appropriate themes.

1. The framework is comprehensive: More 400 respondents indicated that the framework was comprehensive in that it had captured all that is needed for a data literacy program.

    The framework is all-encompassing and takes into account every aspect and variable related to data literacy, including the interrelationships between them (R64).

    The framework is thorough and covers the entirety of data literacy, including the connections between each aspect and variable (R226).

    Data literacy is fully addressed in the framework, which considers all elements and variables, and the framework highlights their interdependence (R332)

    The framework is comprehensive in its scope of data literacy, encompassing all variables and aspects, and exploring the interplay between them (R492).

2. It is well structured, simple and easy to follow, clear and precise.
3. The framework is relevant to the needs of researchers, including faculty members and students, and takes into account the different skillsets that need to be acquired.

    The framework is tailored to meet the requirements of researchers, including both faculty and students, ensuring that all necessary data literacy skills are addressed (R20).

    The framework recognizes the unique needs of researchers and incorporates various skill levels, providing a comprehensive approach to developing data literacy (R88).

    The framework is designed with researchers in mind, taking into consideration the different skill sets that need to be developed to be data literate (R382).

    By considering the specific needs of researchers, including faculty members and students, the framework offers a relevant and thorough approach to developing data literacy skills (R457).

4. It is inclusive, involving all the stakeholders in the data literacy process.
    Quotes?

159

5. It has a clear focus on the key competencies required for research data management (RDM).

   The framework has a laser-sharp focus on the core competencies necessary for effective research data management, ensuring that researchers have all the necessary skills to work with their data (R516).

   The framework provides a clear roadmap for researchers to develop the key competencies required for research data management, including data organization, data curation, and data dissemination (R13).

6. The framework is well organized and well fabricated, making it easy to navigate through the variables.

7. It highlights the role of the library in the data literacy process.

8. It is detailed, providing guidance on the data literacy process and identifying the relevant competencies involved.

   The framework is extremely detailed, providing a comprehensive understanding of the data literacy process and the competencies required for effective research data management (R488).

   It provides valuable guidance on the data literacy process, highlighting the relevant competencies involved and ensuring that all stakeholders have a clear understanding of the steps involved (R529).

9. The framework is well designed and relevant to research, creating awareness among all stakeholders.

   The framework's design is well thought out and relevant to research, raising awareness among all stakeholders involved in research data management (R96).

In summary, the respondents like the framework for its comprehensiveness, interconnectivity of variables, relevance to current research, clear and easy to follow structure, relevance to researcher's needs, inclusiveness, feasibility, clear distinction of elements, comprehensiveness,

160

interactivity of variables, identification of key competencies, involvement of all stakeholders, comprehensiveness, and simplicity. Furthermore, it is well organized, well fabricated, comprehensive, and easy to navigate. The framework is also seen as meeting the objectives of the study, being well designed, creating awareness to all stakeholders, incorporating all involved in research publication, being detailed and well organized, and having the potential to be applied in university libraries.

5.15   Any possible gaps identified by the respondents in the framework

Respondents were asked to indicate some of the gaps they could identify from the proposed framework. From the responses given, the researcher was able to identify some themes. No direct quotations are provided in this case due to the brevity of the responses, which often only consisted of one or two words. Consequently, the researcher had to interpret the meaning of the responses and formulate appropriate themes based on their understanding.

1. No major gap in the proposed data literacy framework based on respondents' opinions.
2. Minor suggestions were made regarding the roles of key players, place of intellectual capital, and training personnel responsible for RDM training. A few respondents (28) suggested clarifying the roles of key players in the framework and incorporating the importance of intellectual capital. There was also a suggestion to ensure the personnel responsible for RDM training received proper training

Based on the responses, there was no major gap identified by the respondents in the proposed data literacy framework as per their opinions.

5.16   What respondents think should be excluded from the framework

Having interacted with the proposed framework, respondents were asked to identify what they could exclude from it. Due to the shortness of the responses that frequently comprised of only one or two words, the researcher had to interpret the responses to derive suitable meaning. As a result, there are no direct quotations available in this case.

Based on the responses provided, the majority of the respondents (331) did not indicate any item in the proposed data framework that would cause them to exclude it. The responses suggest that they were willing to accept all aspects of the framework. These findings imply that the framework may have been well-designed and is inclusive enough to encompass most, if not all, of the essential elements required for a data literacy program.

However, 93 respondents did mention some other elements that could be excluded, such as "leaving out analytic competencies" and "stakeholders - because they are just for business," It

161

is unclear from the responses whether these elements were included in the proposed framework or if they were hypothetical elements that respondents suggested could be excluded. This aspect, therefore, needs further investigation.

Overall, it is difficult to identify major gaps in the framework based on the responses provided, as most of the responses were either non-specific or indicated that the framework was satisfactory.

5.17 Any other comment about the data literacy framework

As the last question, respondents were requested to give any other comment, about the proposed data literacy framework, they had interacted with. From the responses given, the researcher generated some themes. However, while some themes have direct quotations from participants' responses to support them, others do not. This is because some responses only consist of one or two words, which the researcher had to interpret in order to come up with a suitable theme.

1. Majority of the respondents (more than 50%) highly recommended the framework for adoption, implementation, and use in research, data literacy training, and libraries. They described it as comprehensive, elaborate, suitable, appropriate, and feasible.
   I recommend the framework for any research work since it is comprehensive and elaborate (R2).
   This framework is quite appropriate for data need (R48).
   It would be good if the framework would be adopted by learning institutions (R16).

2. Some respondents (159) expressed a desire to learn more about the framework, its implementation, and its potential benefits. Some asked for more information about information literacy in general.
   I would like to learn more about the framework and how it can be useful for our library (R169).

3. Many respondents (308) expressed positive feedback on the quality and clarity of the framework, stating that it is well done, of good quality, clear, readable, and/or understandable.

4. Some respondents suggested that the framework could be used for data literacy training, be adopted in libraries or institutions, or included in the curriculum for research and postgraduate students.

5. A few respondents suggested that the framework should be included in the curriculum for research and postgraduate students or should be taught in classes rather than acquired in the library.

According to the feedback, the suggested data literacy framework looks to be well-received by respondents. These observations indicate that the suggested data literacy framework has the potential to be a beneficial resource for researchers.

5.18 Summary

The analysis of the data from the study involved 539 participants, consisting of PhD students or candidates and full-time faculty members holding PhD degrees. From the findings, the majority of respondents expressed dissatisfaction with the level of data related services offered by their respective libraries. For instance, most respondents disagreed or were uncertain whether their libraries offered guides on research data management, training modules, or hosted data management training events. The findings reveal that there is a need for education and support in the area of data management and sharing, as well as a need for transparency and communication around the collection and sharing of research data.

The findings also highlighted several challenges associated with managing research data. Most respondents were unaware of open data as a publishing option or that libraries provide guides on data publishing and advice on copyright issues related to data publishing. Although over 28% of respondents acknowledged that libraries provided data publishing and sharing services, the majority expressed dissatisfaction with the level of access to certain data management tools offered by libraries. Respondents indicated that libraries should provide templates for data management and metadata development.

The findings from the crosstabulation tables provide valuable insights into the association between participant category and the creation of a data management plan (DMP) and metadata for research, the relationship between the category of participants and their beliefs regarding the role of the library in promoting data literacy, and the correlation between participant category and familiarity with information literacy, statistical literacy, and digital literacy.

The findings suggest a gap between the services being offered by libraries and the perceived needs of their users. The prioritized list of challenges associated with managing research data can guide efforts to address these issues. Respondents considered data literacy an essential aspect for researchers, as it improves data management and promotes research outcomes. It also enhances data competencies, equips researchers with necessary skills for effective data management, and enables effective data analysis. These findings reveal the importance of training and education on data management practices and the need for libraries to enhance their services to meet the changing needs of their users.

164

CHAPTER 6

6.   PRESENTATION OF FINDINGS – INTERVIEWS

Another data collection method employed by the researcher involved conducting interviews with a subset of the sampled participants. This chapter delves into the analysed data derived from these interviews.

6.1   Introduction

This chapter presents findings from the interviews which were predominantly qualitative. However, some of the questions (2, and 3 for university librarians and 2, 4, 6 and 7 for research/reference librarians) from the interview were closed ended hence the researcher analysed them quantitatively as descriptive statistics. The findings are based on data collected from university librarians and research/reference librarians from each of the participating institutions. The chapter presents a comprehensive analysis of the data collected and provides an understanding of the experiences, perspectives, and opinions of the participants. The results are presented in a descriptive manner, and themes and patterns that emerged from the data are highlighted and discussed in detail. The aim of this chapter is to provide a comprehensive and in-depth understanding of the research topic.

To ensure that the interviewees were well prepared and had a clear understanding of the interview process, they were sent the interview questions and a consent form at least one week prior to the scheduled interview. The interviewees had the choice of selecting the location for the interview, with most of the interviews being conducted at their place of work, except for one interviewee who chose to conduct the interview at a hotel.

To maximize efficiency and minimize disruption, the interviewer carried out interviews with both staff members, based at a specific university, on the same day. The interviews lasted between 25-30 minutes, and prior to the start of the interview, the interviewer sought permission from the interviewee to record the session. The purpose of recording the interviews was to ensure that the data collected was accurate and could be used for later analysis.

6.2   Thematic analysis of responses

The researcher presented the results of qualitative data analysis using thematic analysis. The collected data from the interviews were analysed to identify common themes and patterns that emerged from the responses of the participants. The goal of this analysis was to gain a deep understanding of the experiences, perceptions, and perspectives of the participants related to

the feasibility of offering standardised data literacy services at selected private university libraries in Kenya and to identify key themes that emerged from the data.

While using thematic techniques in analysing the qualitative data, the researcher started by transcribing the interviews from audio recordings into written text. The researcher familiarised himself with the data, reading through the transcripts several times in order to identify initial impressions, observations, and ideas. Initial codes were identified. This was done by breaking down the data into smaller chunks, or segments, and labelling them with initial codes that capture the key topics, ideas, or themes that were emerging from the data. The researcher then grouped the codes together into broader themes and sub-themes, while reviewing the data to ensure that all relevant segments were included in each theme. The findings of the interviews, starting with the university librarians are presented in the order in which the questions were asked.

Although most of the themes were backed by direct quotations, there were a few themes that the research could not support because some responses were only one or two words. As a result, the researcher had to interpret and generate a relevant theme.

6.3   Feedback from the University Librarians (UL)

The university librarian is the most senior manager and leads the university library strategically.

6.3.1   Roles and experience

To start of the interview, the participants were queried regarding their role as a University Librarian at the university and their tenure in the position.

> i.   Role as University Librarian: All respondents indicated that they are university librarians at the university, and that their roles involve overseeing the library and its activities. Some of the verbatim responses in relation to the question are presented below.
>
> *I am in charge of the library overseeing all activity in it (UL1 interviewed on 7th November 2022)*
> *My roles include coordination of all Library services. I am the contact between University Library and other arms of the University (UL3 interviewed on 9th December 2022)*

My role generally is overall supervision of the functions of the library (UL5 interviewed on 12th December 2022)

ii.   Duration in the position: Respondents were also required to indicate duration of the in their positions. According to the responses the duration ranges from 6 months to 11 years.

iii.   Responsibilities: The responsibilities of the university librarians as indicated by the respondents include policy development, staff management, coordination of library services, communication with other parts of the university and participation in university management. Some of the respondents indicate that they also ensure optimal access to information resources, position the library as a centre of learning, provide overall library leadership, and supervise the functions of the library.

iv.   Focus on enhancing learning and research: As university librarians, they also indicated that in their duties they aim to enhance the university's strategic goals by providing vision, strategic direction, and leadership for the library. They also advance the research agenda and enhance the research profile and ranking of the university by ensuring students, faculty, researchers, and staff have access to an international wealth of information resources. One of the librarians indicated that;

I ensure that students, faculty, researchers and staff have optimal access to an international wealth of information resources, matched to the university's mission thus furthering the research agenda and enhancing the research profile and ranking of the university (UL2 interviewed on 17th November 2022).

v.   Advising on library services: From the response there is an indication that university librarians also provide advice on issues related to library services.
I provide advice on the issues related to library services (UL5 interviewed 12th December 2022).

6.3.2   Research data services provided by the libraries

The interviewees were presented with a range of research data services and were asked to indicate which of these services were being offered by their libraries.
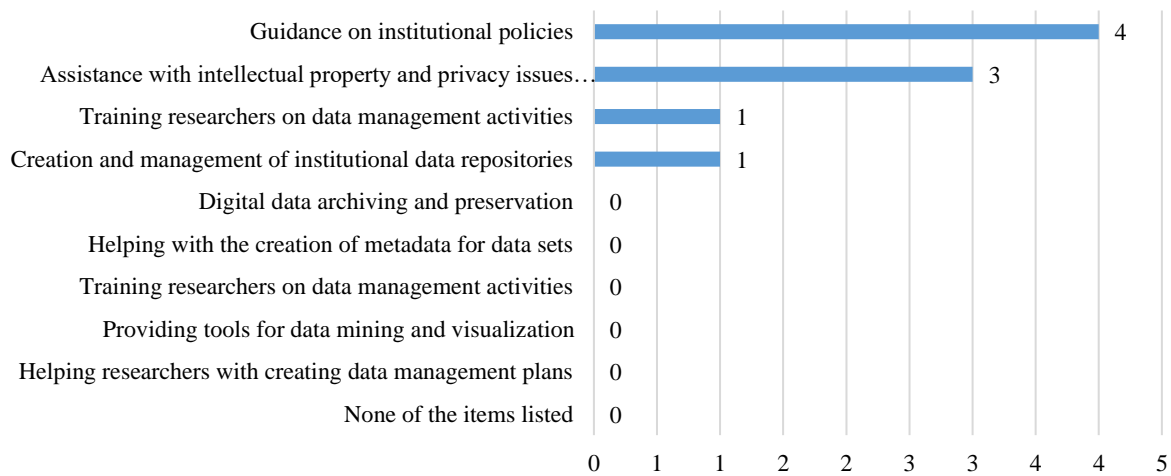
Figure 6- 1: Services being offered by their libraries

From the findings, the results showed that the majority of the pre-selected services were not being offered by the libraries.

None of the respondents indicated that their libraries were helping researchers with creating data management plans and providing tools for data mining and visualization. This already indicated a gap in support for researchers who may need guidance in managing their data effectively and could be seen as a missed opportunity for libraries to support researchers in making the most of their data. Similarly, none of the respondents indicated that their libraries were helping with the creation of metadata for data sets, digital data archiving and preservation, or offering no data services at all.

Only one respondent indicated that their library was providing creation and management of institutional data repositories. This suggests that there is room for improvement in terms of providing infrastructure for long-term storage and access to research data.

Another respondent indicated that their library was offering some level of training for researchers on data management activities. This suggests that libraries are taking steps to support researchers in managing their data more effectively.

A significant number of respondents, four in total, indicated that their libraries were providing guidance on institutional policies. This suggests that libraries are playing an important role in

helping researchers navigate the various policies and regulations related to research and therefore also research data. Some of the policies that are available in these universities are captured in figure 6.2.

Finally, three respondents indicated that their libraries were providing assistance with intellectual property and privacy issues related to research data. This suggests that libraries are taking steps to support researchers in navigating the legal and ethical issues surrounding their data.

Overall, the findings suggest that while libraries are playing a role in supporting researchers with their data management needs, there is still room for improvement in terms of offering a wider range of services. This finding corroborates findings in Figure 5.11 where the majority of respondents expressed dissatisfaction with the level of data related services offered by their respective libraries. For instance, most respondents disagreed or were uncertain whether their libraries offered guides on research data management, training modules, or hosted data management training events. The findings reveal that there is a need for education and support in the area of data management and sharing, as well as a need for transparency and communication around the collection and sharing of research data.

6.3.3 Research data management policies

The interviewees were shown various policies supporting research data services and were asked to identify which of these policies were already in effect at their institutions, given that they provide postgraduate research opportunities.
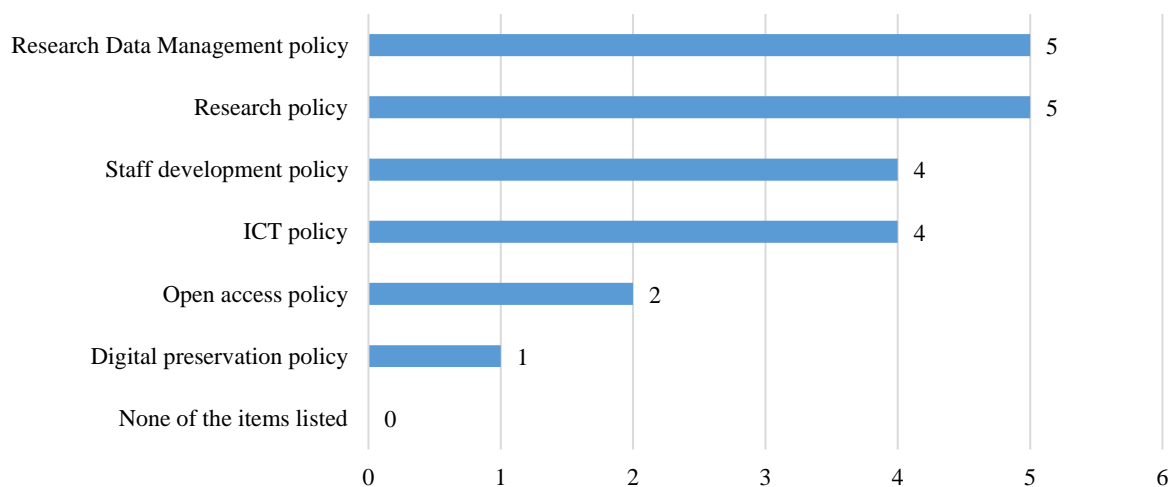
Figure 6- 2: Research data management policies

Results from findings showed that out of the policies considered, research policy and research data management policy were the most prevalent, with five (5) respondents indicating that their institutions had these policies in place. This suggests that institutions are recognizing the importance of having policies that support research and data management, and are taking steps to implement these policies.

In addition to research-related policies, the study found that four respondents indicated that their institutions had ICT policy, which may indicate a recognition of the importance of technology in research data management. Staff development policy was also mentioned by four respondents, suggesting that institutions are taking steps to ensure that staff members are given a chance to further their professional skills.

While open access policy was only mentioned by 2 respondents, this is still an important finding, as it suggests that some institutions are recognizing the importance of open initiatives. Finally, only 1 respondent indicated that their institution had digital preservation policies, which may indicate that there is room for improvement in this area.

Overall, these findings suggest that institutions are taking steps to implement policies that support research data services, but that there is still room for improvement in certain areas.

6.3.4 Structure and setup to support data management and data literacy

During the interview, the interviewer aimed to understand how the overall organizational structure and setup supports data management and data literacy for researchers. To answer this question thoroughly, the researcher asked the four sub-questions.

a. The interviewees were asked whether their institution has various organizational structures in place that are focused on promoting data literacy.

   i. Availability of organizational structures: One of the respondents indicated that there are no organizational structures in place aimed at spearheading data literacy, while others mentioned the availability of relevant structures within the university.

   We have no available structure (UL5 interviewed 12th December 2022).

   ii. Research office and directorate: Some respondents mentioned the existence of a research office or directorate within the university, with a research director whose objective is to promote evidence-based research for decision-making. One of the respondents said that,

170

We have the Institutional Research Office under the University Advancement division. There are relevant personnel within the office and its objective is to promote the use of evidence-based research for decision-making towards achieving the mission of the University through accurate, timely, tailored research/surveys to suit the diverse audiences (UL2 interviewed 17th November 2022).

b. The interviewees were asked about the importance of data management in the organizational structure of their university. From the responses, the researcher noted that:

   i. Data management is considered a key component in the university's organizational structure.

   Yes, it is. We need to think about this critically (UL1 interviewed 7th November 2022).

   ii. According to one of the respondents, the creation of an Institutional Research Office to manage university data is an indication of the university's stance on the importance of data.

   The establishment of an Institutional Research Office to manage the university data is an indication of how the university values and relies on data (UL2 interviewed 17th November 2022)

   iii. Some respondents did not know whether data management was seen as a key component in the university's organisational structure

c. Respondents were asked to identify what they foresee as some of the cost implications of implementing a data literacy programme- in their respective institutions. Two cost-components were mentioned as indicated below

   i. Training costs: This includes what they termed as the cost of training stakeholders, training trainers, and capacity building for librarians and research staff.

   Training of the stakeholders has cost implications (UL1 interviewed on 7th November 2022).

   Cost in terms of training those to implement the program (UL4 interviewed on 13th December 2022).

© University of Pretoria

ii. Infrastructure costs: This includes the cost of improving the technical infrastructure to support the program and the cost of facilities and technology for data storage and management.

Technical infrastructure to support the programme – cost implications if there is need to improve ICT infrastructure (UL2 interviewed on 17th November 2022).

Facilities and proper technology to store data in a manner that is easily retrievable and training on data management (UL3 interviewed on 9th December 2022).

d. The interviewees were asked to identify any other stakeholder that would be involved in data literacy programme activities other than the library.

i. Respondents pointed out various key university departments. They mentioned the involvement of departments such as the Research department, ICT, School of Graduate Studies, Institutional Research Office and the Directorate of Research.

ii. Involvement of Academic staff: The faculty members especially those who are supervising postgraduate students were identified as key players to be involved.

All postgraduate supervisors because they work closely with students who are handling research data (UL3 interviewed on 9th December 2022).

It is not known why this respondent pointed out postgraduate supervisors when the model had faculty members who majorly are postgraduate research supervisors.

iii. Involvement of Leadership: Leadership such as the Deputy Vice Chancellor in charge of Academics was mentioned as a potential participant in the data literacy program.

6.3.5 Level of attention to research and related services

Interviewees were asked about the level of attention their university gives to research and related services. There were sub questions under this questions that interviewees were to respond to.

    a. They were asked how research is captured in the strategic plan of the university.

        i. Importance of Research: One of the interviewees emphasized the importance of research in the university in view of the nature of the university as it is a postgraduate institution and research is highly valued.

        This is a postgraduate institution and we value research, that's why there is a research department with a director heading the research department (UL1 interviewed on 7th November 2022).

        ii. Research as a pillar in the strategic plan: They identified research as one of the pillars within the university's strategic plan

        It is one of the pillars in the strategic plan (UL2 interviewed on 17th November 2022).

        iii. Research as a prominent component: According to some of the interviewees, research is included as one of the main aspects, sub-themes or components of the university's plan and is given prominence.

        It is one of the sub themes and so the importance is it's among the component of research and consultancy theme (UL3 interviewed on 9th December 2022).

        It is one of the main aspects covered. It is at the top. It is given prominence (UL4 interviewed on 13th December 2022)

        iv. Clear inclusion in the university plan: One of the interviewees said that research is included in the university's plan and stands out clearly as one of the pillars of the university.

        It's included in the university plan. It stands out clearly as one of the pillars of the university (UL5 interviewed on 12th December 2022).

b. Interviewees were asked how researchers are recognised in their university. From the responses, there are different ways of recognising researchers.

    i. Annual recognition ceremony during which researchers are awarded.
The university has or holds an annual recognition ceremony where researchers receive monetary awards according to how many papers and book chapters are published. This is the criteria used to award and it's just for faculty and staff (UL1 interviewed on 7th November 2022).

    ii. Research funding was mentioned as a form of recognition where researchers get funded by the university.

    iii. Two of the respondents mentioned that their universities recognise researchers through awards and other forms of incentives.

    iv. Research support: Some of the universities, according to the responses, provide some form of support. However, the interviewee could not detail what it is as they indicated that the provision is detailed by the research office.

    v. Financial compensation: One of the interviews indicated that there is some form of compensation that researchers receive depending on the grant they attract in relation to research.
There is some form of compensation (which is clearly stipulated) for example, if you attract a grant, there is a percentage that remains with you as the researcher (UL3 interviewed on 9th December 2022).

    vi. Another form of recognition that was mention is that researchers are exposed to research seminars, where the cost is covered by the university.

c. The researcher wanted to find out from the interviewees how research activities are funded in their respective universities.

    i. Budgetary allocation: According to one of the respondents, there is a university budgetary allocation through the office of Deputy Vice

Chancellor Academic and Administration that is meant for research activities.

There is a budgetary allocation under the DVC Academic Administration (UL1 interviewed on 7th November 2022).

Another interviewee indicated that there is annual budgetary allocation which is a small percentage of the total university expenditure.

A small percentage of the annual budget is set aside, though it's a small margin (UL3 interviewed on 9th December 2022).

ii. Research activities are also funded through the School of Graduate Studies, Research and Extension or the department that is charged with research activities

Through the School of Graduate Studies, Research and Extension (UL2 interviewed on 17th November 2022).

Funding is done through the research department and there is someone who is in charge (UL5 interviewed on 12th December 2022).

iii. Research Fund: According to one of the interviewees, the university has got a research fund which is used to support research activities.

There is a research fund. University provides a theme and anyone carrying out research around that is funded (UL4 interviewed on 13th December 2022).

iv. Conference Funding: In their response, one of the interviewees include conference attending as one of the research activities and indicated that researchers are funded to attend conferences outside the country.

They are funded to attend conferences outside the country (UL5 interviewed on 12th December 2022).

6.3.6   ICT infrastructure provided to researchers

The researcher requested the interviewees to discuss the ICT infrastructure that is available to researchers to enable data collection, data manipulation, and data sharing.

i. Availability of data analysis software: According to two of the respondents their universities have made provision for data analysis software such as SPSS and

175

NVIVO, which are available in the computer lab and researchers have access to it.

The university has provided data analysis software e.g., SPSS for researchers which is available in the computer lab and researchers have access to it (UL4 interviewed on 13th December 2022).

We have data analysis software such as SPSS and NVivo which researchers have access to (UL2 interviewed on 17th November 2022).

ii.   No dedicated ICT structure: In some of the universities, according to the responses given by the interviewees, there is no dedicated ICT structure provided by the university for data collection, manipulation, and sharing.
In view of data collection, data manipulation and data sharing I would say that there is no ICT infrastructure provided (UL3 interviewed on 9th December 2022).

I would say that there is no ICT structure dedicated to data collection, manipulation and data sharing (UL5 interviewed on 12th December 2022).

iii.   Individual management of data: One of the interviewees said that researchers were managing their data at an individual level. Whether this was because the university did not have ICT infrastructure provide or it was a choice by the researchers even though ICT infrastructure had been provided, is a line of thought should be investigated in another research.

6.3.7   Respondents' views on the library spearheading the data literacy program

The interviewees were asked about their views on the library spearheading the data literacy program in their university. The following themes were generated from their responses.

i.   Positive views: From the responses it was clears that respondents had positive views towards library spearheading data literacy program.
I am very open to the idea that the library spearheads the data literacy program (UL2 interviewed on 17th November 2022).
It's an interesting initiative and the library will be interested to spearhead or be involved in it (UL1 interviewed on 7th November 2022).

Will be happy and would want the library to be on the forefront in its implementation (UL5 interviewed on 12th December 2022).

ii.   Involvement of relevant stakeholders: One of the respondents was of the view that even if the library was to spearhead, there is need to involve other stakeholders within the university.
It would however require the involvement of other relevant stakeholders within the university (UL2 interviewed on 17th November 2022).

iii.   Library's competitive advantage: One of the reasons that stood out in the response was that librarians have the necessary skills to rain researchers hence making the library ideal in spearheading data literacy program. Other than data related skills, respondents mentioned the fact that, librarians have some ICT related background that would give the library an added advantage in spearheading the program.

The library needs to be on the frontline due to the skills librarians have in research and data manipulation and management (UL3 interviewed on 9th December 2022).

Majority of librarians also have a background in ICT and can easily integrate infrastructural requirements and aspects of dissemination (UL3 interviewed on 9th December 2022).

iv.   Confidence in the library: Some respondents were of the view that the library stands out to spearhead the program since there has been a long-term relation and confidence in the library by the students and also since the library has a history of implementing information literacy.

The library can play a key role because students have so much confidence in the librarians as they consult them regularly. Furthermore, the library has experience with information literacy which it has implemented and it is charged with running it within the university (UL4 interviewed on 13th December 2022).

177

The reason I am saying this is because, so far, the library spearheads the Information Literacy program, teaches students, faculty and staff Information literacy skills and research skills. That gives the library an advantage to do it (UL1 interviewed on 7th November 2022).

6.3.8 Addressing data literacy training needs and services

Respondents were asked to provide information on how they are addressing data literacy training needs and services within their libraries.

i. Through information literacy program: one of the respondents indicated that they address issues to do with data literacy through information literacy.
It's within the Information Literacy program. We do not teach data literacy as an entity but we do it through the Information Literacy Program. We have an intensive Information Literacy Program (UL1 interviewed on 7th November 2022).

ii. Limited activity. Some responses indicated that some libraries had a limited focus on data literacy training.
I am not sure we are doing that at the moment. We haven't conducted any training on data literacy as a library (UL2 interviewed on 17th November 2022).

iii. Potential for relevant activity: Even though some respondents indicated that they not doing anything about data literacy, they exuded confidence that there is potential for data literacy training to be incorporated in library seminars.
So far, I don't think we are doing it. However, we have a platform that provides a place to engage because every semester we do library seminars. For that reason, data literacy could be one of the training areas during these seminars (UL3 interviewed on 9th December 2022).

iv. Referral: Some responses indicated that the library staff may offer some level of assistance to users in data literacy related queries but often refer them to other departments for further support.
I would say that some staff who have skills in analysis do help the users who come in asking for assistance. When users come to us in the library for such

related assistance, we direct them to ICT department to install SPSS. It is the ICT that will also train them on how to use it and how to analyse data. As a library we are limited in the whole area of data (UL4 interviewed on 13th December 2022).

6.3.9 Research data skills levels among librarians

The researcher asked for the interviewees' thoughts on the research data skills levels among librarians at their universities. This question divided into sub questions as follows:

a. The researcher asked the interviewee to share an opinion on the extent to which the current set of librarians possess the necessary skills and knowledge to provide data literacy training, and requested an explanation for the perspective.

  i. The respondents believe that the current set of librarians under-equipped and lack sufficient skills and knowledge in data literacy.

  I would say very minimally equipped. We have concentrated on Information Literacy and not data skills or research data skills (UL1 interviewed on 7th November 2022)

  We are not very competent. Out of five we rate them at three (3) (UL5 interviewed on 12th December 2022).

  ii. While concurring that there is lack of skills among librarians, one of the respondents indicated that this need has been recognized and must be addressed by various libraries.

  It's something the library can think about to develop and implement (UL1 interviewed on 7th November 2022).

  ...regardless of the skills set that we currently have, there is definitely a need for data literacy training for librarians (UL2 interviewed on 17th November 2022).

  iii. The respondents acknowledge the need for data literacy training for librarians and recognize its significance in the library.

  It is therefore important to establish where we are at and move on from there (UL2 interviewed on 17th November 2022).

b. The researcher asked what measures can be taken to improve the skills and knowledge of librarians in data literacy so that they can assume the role of data literacy experts, in case they are not currently prepared to train researchers.

    i.   Training: what came out clearly from the all the respondents is that librarians need training in research data and data literacy to understand its meaning and to become data literacy experts. They indicated that this training can be initiated by the library, through the staff development policy, capacity building, mentorship by respective departments, and professional development programs.

They need training in research data to understand what data literacy means. The training to be initiated by the library could take the advantage of staff development policy since there is provision for further training of staff members (UL1 interviewed on 7th November 2022).

We need to have capacity building and mentorship by respective departments in order to equip librarians. We can also Take advantage of free online courses and allow librarians to enrol and get trained in matters data literacy (UL2 interviewed on 17th November 2022).

They can be equipped with skills through professional development programs (UL3 interviewed on 9th December 2022).

They can be trained via staff development initiative (UL4 interviewed on 13th December 2022).

    ii.   Capacity building: it was mentioned that librarians can be equipped with data literacy skills through capacity building, mentorship, and professional development programs.

We need to have capacity building and mentorship by respective departments in order to equip librarians (UL5 interviewed on 12th December 2022).

iii. Partnerships: One of the respondents was of the opinion that the library can partner with other universities or external specialists and alumni, to impart knowledge related to data literacy among librarians.

To enhance training, we can also partner with other universities or external specialists and alumni. to impart knowledge related to data literacy among librarians (UL2 interviewed on 17th November 2022).

6.3.10 Data literacy framework

As was previously mentioned the respondents were provided with the proposed data literacy framework, which they needed to review in order to answer the following questions.

a. The researcher asked if there was anything in the data literacy framework that did not make sense to the respondent.

From the responses given, the following themes were generated by the researcher.

i. Positive perception: From the responses, there was a general positive perception indicating all that was included in the framework makes sense. Some of the respondents said that the framework is needed

Everything makes sense and we need it (UL5 interviewed on 12th December 2022).

ii. No objections to the framework: None of the respondents raised objections to the proposed data literacy framework, which suggests that it is well-received and meets their expectations.

iii. Identified gap: One of the respondents expressed doubt regarding how others can reuse stored data.

Am yet to understand from the framework how one can use my data if I have stored it somewhere (UL1 interviewed on 7th November 2022).

b. Respondents were asked to indicate what they like about the framework.

From the responses, the researcher was able to pick out the following:

i. Clarity: Some respondents indicated that they like the clarity of framework. They said that they liked how clear and systematic

181

presentation of independent and dependent variables had been presented.

I like the way you have captured the connection between independent and dependent variables. To me this is simple and clear and done systematically (UL1 interviewed on 7th November 2022).

ii. Comprehensive coverage: One of the respondents indicated that they liked the comprehensive coverage of aspects related to data literacy implementation

The framework covers most aspects as far as implementation of data literacy is concerned hence, I like it as it is (UL2 interviewed on 17th November 2022).

Another respondent (UL4 interviewed on 13th December 2022) indicated:

I like the fact that it has captured most of the items, for example the current environment that researchers are operating in, what is required, stakeholders, among others.

c. Respondents were asked to identify the gaps in the framework.?
   i. From the response, it was clear that majority of the respondents did not identify any major gap in the framework.
   ii. Non-faculty university staff: One of the respondents felt that there was inadequate inclusion of non-faculty university staff in the framework.
   I feel that the general university staff have not been captured in the framework. The focus seems to be only on faculty and postgraduate students. There are staff that do research and their articles are published. This framework should recognize staff (UL1 interviewed on 7th November 2022).
   iii. Faculty members: Another responded felt that there is need to include faculty members as stakeholders
   Add 'faculty members' as a stakeholder. They are well equipped in data management (UL4 interviewed on 13th December 2022).

d. The respondents were also asked about what they would like to remove. According to the responses given by the respondents, there was absence of objections to the proposed data literacy framework and the perception was that there was nothing to be removed.

6.3.11 Request to collaborate in the development of a single data literacy curriculum

In view of the framework, the respondents were asked what their response would be if they were requested to collaborate in the development of a single data literacy curriculum that would provide standardized data literacy training to all Kenyan researchers, similar to what was done with information literacy.

All respondents indicated positive willingness to collaborate in the development of a single data literacy curriculum for standardized training for Kenyan researchers.

There were other follow-up questions once the respondents had indicated willingness to collaborate.

a. In view of the proposed framework, the respondents were asked what some of the key components that should be included in the data literacy curriculum were. From their responses, the researcher was able to identify the following:

i. Clarity: One of the respondents (UL1) indicated there were areas that would need some clarity especially on the appropriate courses/topics.
I would suggest that we need to have clarity on of the appropriate courses/topics that could form part of the curriculum (UL1 interviewed on 7th November 2022).

ii. Key components: One of the respondents went as far as highlight key components that they felt should be included.
I would suggest that Introduction to data literacy, data management concepts, exploring and describing data, analytical skills, data literacy tools and building a data culture as some of the key areas that could be considered as components (UL2 interviewed on 17th November 2022).

© University of Pretoria

Another respondent was of the view that there should be a focus on digital literacy and statistics from basic levels.

Some of the key components that should be included are digital literacy and statistics from basic levels (UL4 interviewed on 13th December 2022).

There was also the mention of including Data awareness, technical competence, and data communication competence.

I would suggest the inclusion of Data awareness, technical competence especially those that are ICT related as well as data communication competence (UL5 interviewed on 12th December 2022).

b. Respondents were asked to identify some of the roles they think the library should play in the development and implementation of data literacy.
The following themes were generated from the responses given by the respondents:

   i. Facilitation and advocacy role: The library should provide facilitators, space, and equipment and play a leadership role in advocating for data literacy.
   The library should provide facilitators, provide space and equipment and be on the forefront of advocacy of data literacy (UL1 interviewed on 7th November 2022).

   ii. Initiation and coordination role: The library should initiate the data literacy process and coordinate with other departments.
   The library should initiate the process then invite the other departments to join the conversation, the same way we did with information literacy (UL2 interviewed on 17th November 2022).

   iii. Curriculum development and training role: The library should develop the data literacy curriculum and provide training to researchers.

The library should be the one to develop the curriculum and should be the one to offer training to the researchers (UL4 interviewed on 13th December 2022).

iv. Policy development role: The library should develop a data literacy policy to guide the implementation process.
The library should come up with something like Literacy policy which will guide the whole process of implementation (UL5 interviewed on 12th December 2022).

v. Competency development for library staff: The library staff should be competent in data literacy before the implementation process.
It is good that before the implementation the library staff be competent (UL5 interviewed on 12th December 2022).

c. Respondents were asked who, other than the library, would play a key role in the development and implementation of a data literacy program and what their role would be.
Based on the responses given, the following are the common themes identified, arranged in order of frequency:

i. ICT department: This department was mentioned in almost all the responses as a stakeholder that could provide technical support and infrastructure for the data literacy program.
In the case of our university, ... the ICT department which I feel could provide the infrastructure (UL1 interviewed on 7th November 2022).

The ICT department should be considered as a stakeholder…. (UL3 interviewed on 9th December 2022).

ii. Research department: This department was also frequently mentioned as a stakeholder that would play a key role in the development and implementation of the data literacy program.

185

In the case of our university, we have the research department, ... (UL1 interviewed on 7th November 2022).

...the research department in the university and the DVCs in charge of research and innovation would be considered as stakeholders who ought to be involved in the implementation of the program (UL5 interviewed on 12th December 2022).

iii.   University management: The university management, including the Management Board, Vice Chancellor, and DVCs, were mentioned as a stakeholder grouping that should be involved in the budgetary approvals and allocation for the program.

Other stakeholders could be the Management Board or Vice Chancellor, this is for the purpose of budgetary approvals and allocation (UL2 interviewed on 17th November 2022).

iv.   Quality Assurance and Educational Effectiveness: The Quality Assurance office was mentioned as stakeholder that could offer guidance on how the program could be integrated into existing programs.

v.   Other departments and units: Other departments such as the department responsible for online programs and the faculty were also mentioned as stakeholders in the implementation of the program.

vi.   External bodies: External bodies such as the Commission for University Education and the National Commission for Science, Technology and Innovation (NACOSTI), as well as the Library Association and consortium, were mentioned as stakeholders that could be considered in the implementation of a collaborative program.

The Commission for University Education and NACOSTI though they are external, could be stakeholders to be considered in the implementation (UL1 interviewed on 7th November 2022).

It will be good to also consider the Library Association and consortium as a stakeholder (UL3 interviewed on 9th December 2022).

6.3.12 Any other information

Lastly, respondents were asked if they had any other information, related to data literacy, that would be of use to the study.

According to the responses that were given by the interviewees, the researcher drew out the following:

The respondents stated that they had no further information to offer on data literacy beyond what they had already provided. However, they found the study thought-provoking and believed it could serve as a foundational guide for a library's engagement with researchers.

The respondents emphasized the importance of data literacy and suggested that the study should clearly outline its rationale, significance, and benefits, which would help boost its implementation. They also acknowledged that the initiative or program would greatly benefit researchers, but they cautioned that it would require highly skilled trainers and careful planning before implementation.

In conclusion, the respondents believe that the study's comprehensive coverage of data literacy is adequate unless further research is needed.

6.3.13 Summary of feedback from university librarians

From the feedback of the university librarians, the researcher notes that they have a variety of roles, including overseeing the library and its activities, developing policies, managing staff, coordinating library services, communicating with other university departments, and participating in university management. They are responsible for ensuring optimal access to information resources, positioning the library as a centre of learning, providing overall library

187

leadership, supervising library functions, and advising on library services. The findings reveal that there is room for improvement in offering a wider range of research data services. Some librarians indicated the availability of relevant structures within the university to promote data literacy, while others mentioned the absence of such structures. The importance of data management was mentioned as a key component in the university's organizational structure.

The university librarians recognized the need for data literacy training and suggested training, capacity building, and partnerships to improve the skills and knowledge of librarians. There was a positive perception of the proposed data literacy framework, with no objections and a liking for its clarity, comprehensive coverage, and capturing of the researcher's environment, requirements, and stakeholders. The respondents suggested the inclusion of non-technical aspects and faculty members as stakeholders and expressed a willingness to collaborate in the development of a single data literacy curriculum for standardized training for Kenyan researchers. They were of the opinion that data literacy curriculum should include key components such as digital literacy, statistics, data awareness, technical competence, and data communication competence. There was an emphasis that the library should play a facilitation and advocacy role in the development and implementation of data literacy in Kenya.

6.4   Feedback from the Research Librarians (RL)

In this study, research librarians were identified as a significant group of interviewees, chosen based on their involvement in working with researchers. While some universities may not have had an explicit title for research librarians, reference librarians were responsible for providing research services in the library, and therefore, they were closely involved with researchers. A total of five research librarians were targeted for this study, and they were interviewed on the same day as university librarians to reduce the number of trips required by the interviewer to conduct the interviews. All five of them were interviewed at their offices, and like the university librarians, the interview questions and consent form were sent to them in advance, and an appointment was scheduled.

6.4.1   Demographics

The interviewees were requested to provide an introduction that included information about their present role and the duration of their tenure as research/reference librarians at their respective universities.

   i.   Length of service: From the responses, the research was able to identify that the length of service of the respondents varied from 4 years to 9 years.

ii. Background: The respondents have a background in Library and Information Science, Education, and Masters of Philosophy and Information Science. Some of them are pursuing further education.

iii. Role: The respondents mostly described their current role as research or reference librarians in the university. Some of them hold the title of Library Assistant or Librarian in Charge of User and Referral Services. However, they were in one way or another working closely with researchers, either post graduate students or faculty members.

iv. Responsibilities: The responsibilities of the respondents include providing services to students, staff and faculty in different ways including research related services, teaching information literacy and referencing styles, answering reference questions, assisting with research, uploading institutional repositories, providing circulation services, and making resources available to students.

Teach students on how to access resources; print and electronic based resources (RL1 interviewed on 7th November 2022).

My work mostly entails helping postgraduate students especially at research stage (RL2 interviewed on 17th November 2022).

I mostly assist students with their research; literature searches and carry out referral services (RL5 interviewed on 12th December 2022).

6.4.2 Data services on offer

During the interviews, the participants were presented with a range of research data services and asked to indicate which ones were currently offered by their own library.
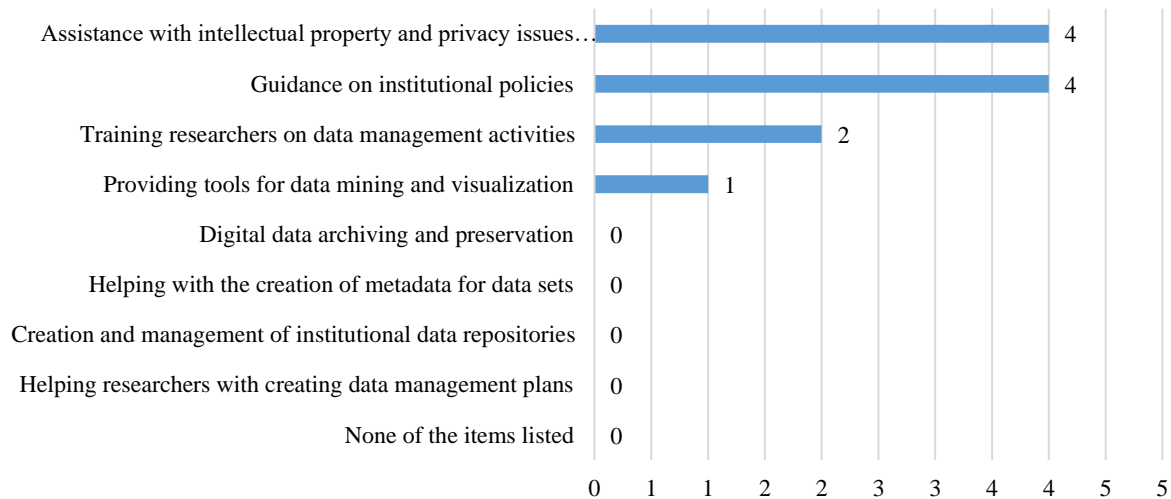
Figure 6- 3: Data services on offer

According to the findings of the interviews, none of the respondents reported that their library currently offers assistance in creating data management plans or creating and managing institutional data repositories. Similarly, none of the respondents reported that their library offers assistance with creating metadata for data sets or digital data archiving and preservation. These services are important for ensuring that research data is organized and preserved for future use. This could indicate a gap in the services provided by libraries, as these are important aspects of managing research data.

One respondent (RL2) indicated that their library provides tools for data mining and visualization, while two respondents reported that their library provides training to researchers on data management activities. These services can help researchers better understand and manage their data, which can lead to more effective and efficient research.

Four respondents (RL1, RL3, RL4, and RL5), reported that their library provides guidance on institutional policies related to research data management, as well as assistance with intellectual property and privacy issues. These services can help researchers navigate complex legal and ethical issues related to their research.

The findings suggest that there may be room for libraries to expand their services related to research data management, particularly in areas such as data archiving and metadata creation. However, it is encouraging to see that many libraries are already providing important services in areas such as policy guidance and intellectual property assistance.

Since the university librarians had been asked same question, the research acknowledges some points of similarity between the two groups' responses. From the findings, none of the respondents indicated that their libraries were helping researchers with creating data management plans. None of the respondents indicated that their libraries were providing tools for data mining and visualization. None of the respondents indicated that their libraries were helping with the creation of metadata for data sets and lastly none of the respondents indicated that their libraries were offering digital data archiving and preservation.

6.4.3 Data services required

Respondents were asked to identify some of the most common research data services sought by researchers in their university.

The researcher was able to pick the following themes from the responses given;

    i.   SPSS for data analysis: Most of the respondents pointed out that the most sought service was related to data analysis especially requesting assistance in using SPSS.

        Most of the researchers come in asking for assistance on how to use SPSS for data analysis (RL1 interviewed on 7th November 2022).

    ii.   Training: One librarian mentioned 'Evaluation and collection of data' as a need that is expressed by researchers.

        They also request to be trained in how to evaluate and collect data (RL1 interviewed on 7th November 2022).

    iii.   No data support needs expressed: Searching for information for research work; one of the respondents indicated that they are not aware of any specific data related service request other than search for information for their research work

        None that I am aware of specifically other than search for information for their research work (RL2 interviewed on 17th November 2022).

6.4.4 Skills and competency

As part of the study, the respondents were asked to reflect on their skills and competency as research/reference librarians in the provided areas related to research data management.
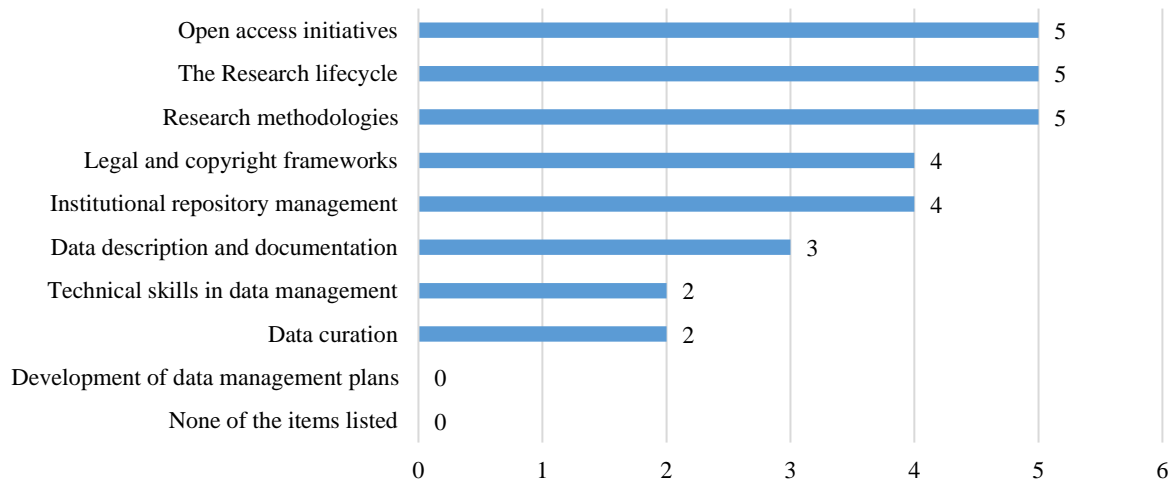
Figure 6- 4: Skills and competency as research/reference librarians

The findings revealed that none of the respondents indicated having skills in development of data management plans, while two respondents indicated they had skills in data curation and technical skills in data management. Five respondents indicated their proficiency in the research lifecycle, whereas three respondents reported their competency in data description and documentation. Four respondents said that they had skills in institutional repository management, open access initiatives, and legal and copyright frameworks, respectively.

Interestingly, none of the respondents indicated that they lacked skills in any of the listed items. Finally, five respondents reported that they had skills in research methodologies.

Overall, the findings suggest that the research/reference librarians surveyed in this study generally possess varying levels of skills and knowledge related to research data management. While some areas, such as the research lifecycle and research methodologies, were areas of strength for a significant number of respondents, others, such as development of data management plans, require more attention and training. Additionally, the results indicate that librarians may benefit from additional training and support in areas such as data curation, technical skills in data management, data description and documentation, institutional repository management, open access initiatives, and legal and copyright frameworks.

6.4.5   Gaps in knowledge base and skills in relation to the current roles and duties

Respondents were asked to identify gaps in their knowledge base and skills, both technical and non-technical, in relation to their current roles and duties of serving researchers as research/reference librarians.

i.   ICT skills: Lack of competency ICT skills was also mentioned as a gap in knowledge and skills of research/reference librarians.
I can say data literacy seems to be a wide field with many technical aspects which require competency in ICT skills (RL1 interviewed on 7th November 2022).

ii.   Tools: The ability to use tools and teach researchers how to manage data effectively is another gap.
Also, I think the use of tools and the ability to teach the researchers on how to how to manage data in terms of data reuse and data publishing (RL1 interviewed on 7th November 2022).

iii.   In-depth expertise required: The respondents feel that the support they provide is too general and superficial.
I would say that the support we give is too general and superficial (RL2 interviewed on 17th November 2022).

iv.   Research cycle knowledge: The respondents feel that they lack knowledge of the entire research cycle and cannot work with researchers throughout the entire process.

Personally, I am not competent enough because I cannot work with the research throughout the whole cycle of the research as there are many things I do not know (RL2 interviewed on 17th November 2022).

v.   Data management planning: The respondents see a gap in their knowledge of data management planning and related areas.

For sure the whole realm of Data management planning and all that is related to it (RL4 interviewed on 13th December 2022).

vi. Upgrade personal education: One of the respondents felt that they the need to upgrade their education to PhD level due to the type of students they serve.

Due to the type of students, I serve for instance, the Doctoral students, I need to upgrade to a PhD (RL5 interviewed on 12th December 2022).

The researcher compared the findings from section 6.4.4 and the themes deduced from the responses in section 6.4.5 and noticed some areas of alignment and some potential contradictions. For instance, there was an agreement on data management planning. In section 6.4.4, none of the respondents indicated having skills in the development of data management plans. In section 6.4.5 the respondents also identified a lack of knowledge of data management planning and related areas. This suggests that there is a need for more training and support in this area. In section 6.4.4, two respondents indicated having skills in data curation and technical skills in data management. In section 6.4.5, the respondents mentioned the need for more competency in ICT skills and the ability to use tools to manage data effectively. This suggests that there is a need for more technical training and support.

One area of contradiction was when five respondents in section 6.4.4 indicated that they had skills in research methodologies. However, in the second question, none of the respondents mentioned research methodologies as a gap in their knowledge base and skills. This suggests that the respondents may feel more confident in their abilities in this area and do not see it as a gap.

In summary, there is a need for more training and support in certain areas related to research data management, such as data management planning and technical skills. Additionally, librarians may benefit from more in-depth training to work with researchers throughout the entire research cycle. Overall, the findings suggest that there is a need for continued offering professional training to ensure that librarians have the necessary skills and knowledge to serve researchers effectively.

6.4.6 Policies in place that support research data services

Respondents were asked to indicate, from the provided list, which of the policies are in place at their institution to support research data services.
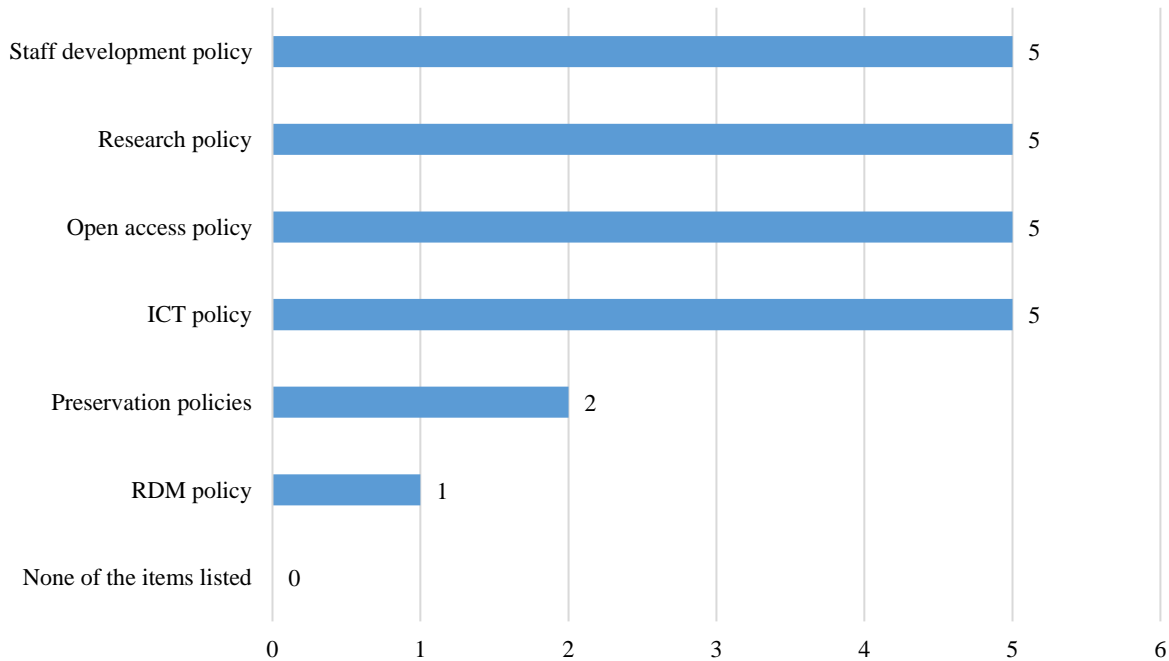
Figure 6- 5: Policies in place that support research data services

The findings showed that out of the seven policies listed, five institutions had research policies in place.

One respondent indicated that they had a research data management (RDM) policy, showing that some institutions have recognized the importance of specific policies for data management. All five respondents indicated that they had ICT policies, highlighting the need for technology infrastructure to support research data services.

Additionally, staff development policies were also prevalent, with five institutions indicating their presence. This demonstrates that institutions are aware of the need to provide adequate further professional training and development opportunities to their staff. This could be an avenue for university librarians to develop their library staff in areas such as research data management in order to help researchers.

All the five respondents indicated that open access policy was available in their respective institutions. However, only two respondents indicated that they had preservation policies in place, suggesting that there may be room for improvement in this area.

Overall, the results indicate that institutions are aware of the importance of policies to support research data services, with the majority of institutions having policies in place. However, there may be a need for more specific policies around research data management and preservation to ensure that research data is effectively managed and preserved for future use.

6.4.7   Training interventions offered to researchers by the library

Respondents were provided with a list of training interventions and were asked to identify which of them were offered to researchers by the library.
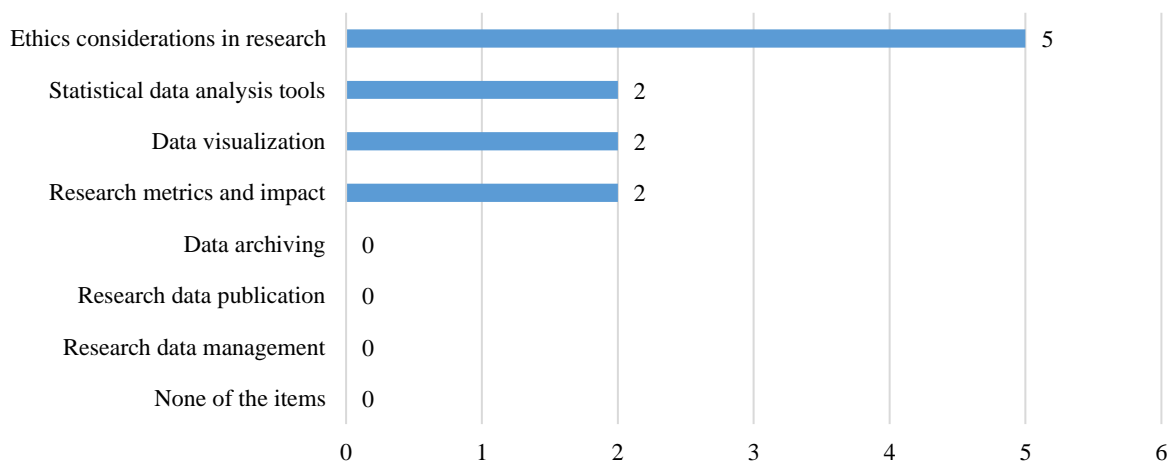
Figure 6- 6: Training interventions offered to researchers by the library

According to the findings, none of the respondents indicated that the library offers research data management or data publication training interventions. However, five respondents reported that the library offers training on ethics considerations in research. Two respondents indicated that the library offers statistical data analysis tools, research metrics and impact, as well as data visualization training interventions.

Interestingly, none of the respondents reported that the library offers training on data archiving, and none of them chose the option "none of the items listed."

These results suggest that while the library provides some training interventions for researchers, there may be room for improvement in offering more diverse training options, particularly in areas such as data management and archiving. It may also be worth considering ways to increase awareness of the training interventions that are currently available to researchers, to ensure that they are taking advantage of the resources that are on offer.

These findings were compared with the findings in section 5.4 (specifically sections 5.4.1, 5.4.2, 5.4.3, 5.4.4, 5.4.5 and 5.4.6). The results in this section show ed that the majority of researchers did not believe that libraries offered assistance in the creation, processing, analysis, and access to research data. Regarding training interventions offered to researchers by libraries, as reported in this section (6.4.7) none of the research/reference librarians who indicated that their library offers research data management or data publication training interventions. This suggests a need for libraries to improve their support for research data management and to offer more training interventions to researchers in this area.

6.4.8   Perspectives on library leading university's data literacy program

The respondents were asked to share their views on the library spearheading the data literacy program in their university.

i. Importance of library: The respondents view the library as a crucial aspect of the data literacy program in their university as it is the custodian of information and the main contact for researchers.

The library is the backbone of any academic institution and it is where students and researchers go to for information. It makes sense for the library to spearhead the data literacy program (RL3 interviewed on 9th December 2022).

The library is the first place that people think of when they need information, so it's important for the library to be at the forefront of data literacy initiatives (RL5 interviewed on 12th December 2022).

The library is the hub of knowledge at our university, so it's only fitting that the library takes the lead in ensuring that our students and researchers are data literate (RL2 interviewed on 17th November 2022).

ii. Research conducted in the library: The respondents see the library as a central location for research, making it an ideal place to spearhead the data literacy program.

The library is the central location for research at our university, so it makes sense for the data literacy program to be based there (RL1 interviewed on 7th November 2022).

The library is the first stop for researchers, so it's important that the library provides them with the data literacy skills they need to be successful (RL3 interviewed on 9th December 2022).

iii. Librarians' expertise: The respondents believe that librarians have the expertise to run the data literacy program.

It's the best thing to Librarians. Librarians know how to develop metadata and if given instructions and dedication they will be able to run the program (RL4 interviewed on 13th December 2022).

6.4.9 The extent librarians work with researchers

Respondents were asked to what extent librarians already worked with researchers at their university. The researcher was able to draw out the following themes from the responses given.

i. Training and assistance: Librarians provide training and individualised assistance to researchers, including installation of data analysis tools and training on reference tools.
Librarians provide trainings to researchers and at times offer individualized form of assistance to the researchers. They assist in Installation of data analysis tools, for example, SPSS, provide and train users on reference tools like Zotero (RL1 interviewed on 7th November 2022).

ii. Access to resources: Librarians assist researchers in accessing e-resources and provide training on how to use them.

iii. Limited assistance: Some respondents felt that that the assistance offered by librarians is limited and focused more on collections rather than research related service.

The help offered is not to that very big extent. I always feel as a library and librarians, we have focused more on collections. It would be good to start thinking of researcher's needs (RL2 interviewed on 17th November 2022).

iv. Consultation: During the interviews, all the five research librarians indicated that there was some sort of consultation made by researcher to librarians for various needs related to their research.

6.4.10 Challenges faced while providing research data-related services to researchers

Respondents were asked to point out some of the challenges they faced when providing research data-related services to researchers. The researcher was able to identify some challenges from the responses given. They include:

i.   Researcher skill gaps: Many researchers lack the relevant skills and ICT knowledge, which makes training time-consuming.
     Most researchers lack relevant skills (ICT tools). It then becomes time-consuming training them because you have to start with the basics in ICT (RL1 interviewed on 7th November 2022).

ii. Researcher needs: Respondents find it challenging to identify and fulfil researcher needs. The mentioned the fact that some of the researchers have a specific need and do not listen to alternatives.
     Sometimes it's a challenge to identify exactly what a researcher needs and satisfy them. Some researchers don't listen to alternatives as they have a very specific need (RL1 interviewed on 7th November 2022).

iii. Perception regarding librarian skills: Researchers may have a perception that librarians are not capable of providing data-related services.

     Researchers' perception of the librarians that they can't do (RL2 interviewed on 17th November 2022).

iv. Lack of skills and manpower: Respondents mention one of the main challenges is lack of relevant skills and manpower. These are skills related to data literacy

v.   Understanding of researcher needs: Respondents face challenges in encountering researchers who do not understand their own needs.
     Lack of adequate ICT literacy and encountering researchers who have no understanding of what they want (RL3 interviewed on 9th December 2022).

vi. Older generation of researchers: Some respondents mention difficulty in working with an older generation of researchers who are not familiar with technology.

I do sometimes encounter an older generation of researchers who take time to appreciate available technology (RL4 interviewed on 13th December 2022).

vii. Internet downtime: Respondents mention internet downtime as a challenge in providing data-related services to researchers.

6.4.11 Data literacy framework

Respondents (interviewees) were provided with a copy of the proposed data literacy framework. They were asked to respond to a set of question in view of the proposed data literacy framework.

a. Respondents were asked to mention anything that they felt did not make sense in the proposed data literacy framework.

  i. Understanding: The majority of the respondents seem to understand the proposed data literacy framework and agree with its goals as they indicated that all that was captured in the framework makes sense.

  According to the explanation you have given, I have understood what you intend to achieve. Everything makes sense to me. (RL1 interviewed on 7th November 2022).

  ii. Inclusiveness: One respondent expressed concern about the inclusion of all types of students in the training for data literacy.
  Everything makes sense; however, I am concerned whether all type of students will be included in the training for them to be data literate (RL5 interviewed on 12th December 2022).

b. Respondents were asked to share what they liked about the proposed data literacy framework.

  i. Library as a primary stakeholder: One respondent liked the fact that the library has been identified as the main stakeholder in the proposed data literacy framework.

The fact that the library has been put as the main stakeholder in the proposed data literacy framework (RL1 interviewed on 7th November 2022).

    ii.   Inclusiveness: Respondents liked the inclusion of information literacy and digital literacy in the framework.
…….the inclusion of information literacy and digital literacy in the framework (RL1 interviewed on 7th November 2022).

    iii.   Clarity: Respondents liked the clarity and comprehensiveness of the proposed data literacy framework and how it covers various aspects of data management.
The clarity of data literacy, you can see how to fulfil the need and who is supposed to fill the need and how it's going to be filled to address the challenge of awareness of data management. It's very clear (RL4 interviewed on 13th December 2022).

    iv.   Basics in data management: One respondent appreciated that the framework introduces students the basics in data management.

c. Respondents were asked to point out some gaps they could identify in the framework.

    i.   Technology: One respondent indicated that they can identify a gap in the framework regarding technical skills and technology.
Technical aspects-skills, there is a lot to be done as far as technology is concerned (RL1 interviewed on 7th November 2022).

    ii.   Stakeholders: One respondent suggested the framework did not have graduate research and extension school/department as stakeholders, as they play a big role and therefore it advised that they be included.
Include graduate research and extension school/ department among the stakeholders as they play a big role (RL2 interviewed on 17th November 2022).

iii.   Satisfaction: All the respondents felt that there were no gaps and they were satisfied with the comprehensiveness of the proposed data literacy framework and felt that it was able to address their needs.

d.   Respondents were asked to point out some elements they would like to remove from the proposed data literacy framework.

From the responses given by the interviewees, the researcher noted that there were no changes desired. None of the respondents expressed the desire to remove anything from the proposed data literacy framework.

6.4.12 Request to collaborate in the development of a single data literacy curriculum

In view of the data literacy framework, respondents were asked to indicate what their response would be if they were requested to collaborate in the development of a single data literacy curriculum that would provide standardised data literacy training to all researchers in Kenya, similar to what was done with information literacy.

This question received positive response from all the respondents. Respondents were willing to collaborate in the development of a single data literacy curriculum. They showed a willingness to support and participate in the project. Some of the responses were:

As a Librarian, I would gladly accept the offer to collaborate in the development of a shared curriculum (RL1 interviewed on 7th November 2022).
I would be delighted" (RL2 interviewed on 17th November 2022).
It is an educative venture; I would support it and also be part of it (RL3 interviewed on 9th December 2022).
I would gladly accept the opportunity to learn and also to expose what I know (RL4 interviewed on 13th December 2022).

a.   Key components that should be included in a data literacy curriculum
After indicating that they would be willing to accept the request to collaborate in the development of a single data literacy curriculum that would provide standardised data literacy training to all researchers in Kenya, respondents were asked to point out some of the key components that they think should be included in the data literacy curriculum. They pointed out the following:

i. Technology: One respondent emphasized the importance of incorporating ICT skills and technology.

ICT- supportive technology, ICT skills of the of the staff (RL1 interviewed on 7th November 2022).

ii. Defining data: One of the respondents (RL2) suggested including a scope of what data is and what parts of data need to be archived.

iii. Metadata and discoverability: One respondent (Rl2) suggested including the concepts of metadata and discoverability in the curriculum.

iv. Data archiving justification: One respondent suggested including a justification for archiving data.

v. Data ethics: One respondent suggested including data ethics or ethics involved in the management of data.

vi. Discipline-specific considerations: One respondent suggested that the framework should consider the different disciplines of study.

vii. Communication: One respondent emphasized the importance of data communication

viii. Visualization: A respondent indicated presentation and visualization as key areas to be included.

b. Respondents were asked to identify roles they think the library should play in the development and implementation of data literacy

Respondents identified several roles that they believe libraries should play in the development and implementation of data literacy. According to one respondent, "Libraries have a critical role in this area because they have a tradition of being responsible for creating, managing and preserving knowledge and data," (RL2 interviewed on 17th November 2022)

i. Training and instruction: The first role identified was providing training and instruction in data literacy, as librarians have necessary skills derived from information literacy. As one respondent noted,

Librarians have a lot of experience in training people in information literacy, and it makes sense to leverage that experience in developing training programs for data literacy (RL1 interviewed on 7th November 2022).

ii. Awareness creation: Creating awareness was identified as another role, with respondents believing that libraries should be very much informed as they are the custodians of information.

The library should be the place where people can get information on what data is available and how it can be used, (RL3 interviewed on 9th December 2022).

iii. Collaboration: Collaboration with other stakeholders was also highlighted as important, with libraries expected to spearhead data literacy implementation while incorporating other stakeholders. As one respondent put it,

The library should play a leading role in promoting data literacy, but this needs to be done in collaboration with other key stakeholders, such as IT services and research support offices (RL3 interviewed on 9th December 2022).

iv. Preservation: Finally, respondents highlighted the active role that libraries should take in preservation and providing access to research data, given their custodianship of information, metadata, visualization, and archives. As one respondent explained,

The library has a responsibility to preserve research data and make it accessible to future generations of researchers. This means that the

204

library should be actively involved in the creation of metadata, visualization, and archives for research data (RL4 interviewed on 13th December 2022).

c. The interviewees were asked to identify some of the other stakeholders that would play a key role in the development and implementation of a data literacy program and what that role would be.

The interviewees identified several stakeholders that would play a crucial role in the development and implementation of a data literacy program, as follows:

   i. University research office or Research department- responsible for making data accessible and making recommendations on data collection and use.

   ii. ICT department - responsible for institutional technology and software recommendations.

   iii. Faculty - responsible for student data and recommendations on data use.

   iv. Management or Administration – responsible for funding and cost implications and decision-making.

   v. Consultants - responsible for advising on the process.

   vi. Researchers. Add detail

   vii. Methodology trainers: Faculty teaching research methods, who should integrate data literacy as a component of research methods.

   viii. External stakeholders, such as NACOSTI, the statistics bureau and organizations funding research.

These stakeholders would have different roles to play, such as making data accessible, recommending appropriate technology and software, providing advice on the process, recommending on data use, and providing funding and decision-making.

d. The interviewees were asked to mention some of the factors within their organisation that could be enablers for the emergence of this role in data literacy.

The researcher was able to generate the following from the interviewees' responses on factors that could enable the emergence of a role in data literacy within their organization:

i. Expert trainers: Well-trained researchers and staff development, including comprehensive training for librarians to impart knowledge on others.

ii. Funding as a key enabler: Supportive administration and budget allocation, including goodwill from the administration for funding and acquiring necessary ICT infrastructure such as a data repository.

iii. Collaboration.

iv. Infrastructure: Adequate ICT infrastructure and a data repository.

6.4.13 Any additional information related to data literacy

As the final question, the respondents were asked whether they had any additional information related to data literacy that they believed would be useful for the study.

i. Need for data literacy training: The importance of training on data literacy, especially in handling and management of data:
There is need for providing training to users in handling and managing data. It suggests that data literacy skills are crucial for users to effectively work with data and ensure its accuracy and integrity. Universities should invest in providing training programs to enhance researcher's data literacy skills.

ii. Positive outlook on the study and its potential impact on universities: There is positive response to the study and its potential impact on universities. Respondents believe that this study can bring about a new dimension in data literacy education that some universities may not be using yet. The study can help universities improve their data literacy education programs.

iii. Potential conflicts: Concern about potential conflicts between different departments over ownership and management of data literacy program: The respondents raised concerns about potential conflicts between different departments over the ownership and management of the data literacy program. They highlight the importance of collaboration between different departments to ensure a successful implementation of the program. There would be need for

206

clear understanding of the roles and responsibilities of different departments in managing the data literacy program to avoid any conflicts.

   iv.   Need for clarity on the library to spearhead the program

Respondents suggest that the library can be a suitable department to take the lead in managing the data literacy program, given their expertise in managing and organizing information.

6.4.14 Summary of feedback from Research librarians

From the findings, it can be noted that research librarians had diverse educational backgrounds, including Library and Information Science, Education, and Masters of Philosophy and Information Science. They worked closely with post-graduate students and faculty members, providing research-related services, assisting with research, and making resources available to students.

However, the findings reveal gaps in the data services provided by libraries, particularly in creating data management plans, creating and managing institutional data repositories, creating metadata for data sets, and digital data archiving and preservation. It was noted that the most sought-after research data services by researchers were SPSS for data analysis, evaluation, and collection of data.

The findings also show that research librarians possessed varying levels of skills and knowledge related to research data management. Some areas, such as the research lifecycle and research methodologies, were areas of strength for a significant number of respondents, while others, such as development of data management plans, require more attention and training. The gaps in their knowledge base and skills were mainly in data literacy and ICT skills, which librarians needed to support researchers better.

From their response, it was noted that, research librarians are of the view that libraries should collaborate with stakeholders such as IT departments, researchers, and research administrators to enhance their research data management services. The critical role played by research librarians in data literacy programs and the need for collaboration between researchers and research librarians was highlighted in their responses.

6.5   Alignment between University and Research librarians' feedback

There were some points of agreement as well as divergence between the university librarians and research librarians. As far as the offering data management services is concerned, both

groups of respondents indicated a lack of support for researchers in terms of data management, with a particular gap in providing infrastructure for long-term storage and access to research data. However, research/reference librarians seem to have a slightly better understanding of the importance of providing tools for data mining and visualization, while university librarians seem to have a slightly better understanding of the importance of policy guidance.

From the feedback by the two categories of respondents, it can be seen that both university librarians and research librarians recognize the importance of data literacy and the need for improvement in research data services. The university librarians, in particular, highlighted the need for capacity building and partnerships to improve the skills and knowledge of librarians. The research librarians, on the other hand, pointed out gaps in their knowledge base and skills, particularly in data literacy and ICT skills, which they needed to support researchers better.

Both groups of librarians expressed a willingness to collaborate with stakeholders to enhance research data management services. The university librarians emphasized the library's facilitation and advocacy role in the development and implementation of data literacy, while the research librarians highlighted the critical role, they play in data literacy programs and the need for collaboration between researchers and research librarians.

Overall, the responses from both groups of librarians suggest a recognition of the importance of data literacy and the need for improvement in research data services, as well as a willingness to collaborate to achieve these goals. The study's proposed data literacy framework was well received, with both groups of librarians offering suggestions for its improvement, indicating a readiness to adopt standardized data literacy services.

6.6   Summary

As mentioned in the introduction of this chapter, the chapter presented findings generated from the interviews. These were findings of the data collected from university librarians and research/reference librarians from the participating institutions. The chapter employs thematic analysis to identify common themes and patterns that emerged from the responses of the participants. The interviews lasted between 25-30 minutes, and the data was collected through audio recordings and later transcribed into written text for analysis.

The findings reveal gaps in the data services provided by libraries, particularly in creating data management plans, creating and managing institutional data repositories, creating metadata for data sets, and digital data archiving and preservation. Both university and research librarians recognize the importance of data literacy and the need for improvement in research data

208

services. University librarians highlighted the need for capacity building and partnerships to improve the skills and knowledge of librarians, while research librarians pointed out gaps in their knowledge base and skills, particularly in data literacy and ICT skills, which they needed to support researchers better.

Both groups of librarians expressed a willingness to collaborate with stakeholders to enhance research data management services. The university librarians emphasized the library's facilitation and advocacy role in the development and implementation of data literacy, while the research librarians highlighted the critical role, they play in data literacy programs and the need for collaboration between researchers and research librarians. The proposed data literacy framework was well received, with both groups of librarians offering suggestions for its improvement, indicating a readiness to adopt standardized data literacy services. Overall, the study's findings indicate a recognition of the importance of data literacy and a willingness to collaborate to achieve standardized data literacy services.

The next chapter will provide an in-depth discussion of the findings in view of the existing literature.

CHAPTER 7

7. DISCUSSION, TRIANGULATION, CONVERGENCE AND CONCLUSIONS

With the study's analysed findings presented in Chapters 5 and 6, this chapter aims to discuss these findings while integrating them with the earlier literature presented in Chapters 1, 2, 3, and 4.

7.1 Introduction

In this chapter, the focus is on discussing results and on the triangulation and convergence of outcomes of the study's findings. The research findings were presented separately in quantitative and qualitative formats, as described in the previous two chapters (chapters 5 and 6). Given the mixed methodology used in this study, the triangulation of the findings was considered. This involved comparing and contrasting the results to arrive at a more comprehensive understanding of the research outcomes.

The discussion of the findings was guided by the research objectives which are to:

1. To review the relevance and appropriateness of the proposed comprehensive framework encompassing essential services and components that would facilitate the successful implementation of a data literacy initiative in selected private universities in Kenya (see Section 7.2)
2. Assess faculty and postgraduate student's data literacy needs in selected private universities in Kenya (see Section 7.3)
3. Assess the organisational infrastructure at selected private university libraries in Kenya for the feasibility to offer data literacy (see Section 7.4)
4. Assess the technical infrastructural readiness in the selected private university libraries in Kenya for the implementation of data literacy training (see Section 7.5)

The discussion pertaining to each objective draws upon responses obtained through individual research questions, which were answered either through a questionnaire completed by the researchers or through interviews conducted with university librarians and research/reference librarians.

7.2   Towards developing a data literacy framework

The ultimate aim of this study is to develop a standardized data literacy framework to cultivate a community of data literate researchers at Kenyan universities. The aim in this section of the study was to calibrate a proposed framework with the actual requirements, as these were expressed by the research respondents, before finalizing a workable data literacy framework for Kenyan academic libraries. The study participants were asked to assess the suggested framework and provide feedback on its significance and key areas to prioritize in a data literacy training program. The role of libraries in promoting data literacy was also investigated. Input on various components of the framework was received from researchers, university librarians, and research/reference librarians. These inputs were compared to the findings reported in Chapter 2.

In Chapter 2, the researcher, critically reviewed existing literature on data literacy. The purpose of this analysis was to establish the prevailing discourse about data literacy and to identify requirements encompassing essential services and components that would lead to the development of a comprehensive data literacy framework that could successfully be implemented in universities in Kenya. An initial framework based on the reviewed literature was developed and presented in Chapter 3 (see figure 3.1).

In order to incorporate study participants views so as to develop a final data literacy framework, participants were asked to review the proposed framework. Their insights were discussed with the intension to incorporate feasible options into the finalized framework. A comprehensive description of the data literacy framework is presented in Chapter 8 of this study.

7.2.1   The importance of data literacy as depicted in the framework

This section's findings and analyses emphasize the critical role that data literacy plays in fostering effective data management, research outcomes, and evidence-based decision-making for researchers. Previous studies (see Section 5.7) have underlined the relevance of data literacy, identifying its role in promoting, for example: best practice, collaboration, data privacy, confidentiality, and reuse.

In today's data-driven world, as it was discussed in Section 2.2, data literacy is an essential skill for researchers. This section discusses the importance of data literacy from the point of view of the research participants. The importance of data literacy is highlighted in an earlier study by Palsdottir (2021), who points out that data literacy enables researchers to access, evaluate, and use data in their research, promoting evidence-based decision-making. It builds the

researcher's capacity to conduct research, communicate results, and determine valuable materials and instruments for data management.

Research participants acknowledge data literacy as a crucial skill for researchers (see Section 5.6). Data literacy improves data management and promotes research outcomes by enhancing researchers' data competencies and equipping them with the necessary skills for effective data management, thus reducing time and cost. Furthermore, data literacy improves efficiency in working with research data, enables effective data analysis, and promotes data privacy, confidentiality, and reuse. Data privacy and confidentiality are essential aspects of data literacy as they ensure that researchers handle data ethically and securely. The reuse of research data is an essential requirement for efficiency in research, and data literacy enables researchers to share their data, facilitating collaboration and promoting data-driven decision-making (Raffaghelli & Manca, 2023).

Some studies have also highlighted the importance of data literacy in research. For example, the study by Palsdottir (2021), found that data literacy is critical for researchers to effectively manage their data, collaborate with others, and ensure data quality. Additionally, data literacy is vital for effective decision-making, as it enables researchers to identify patterns and trends in data, facilitating evidence-based decision-making (Raffaghelli & Manca, 2023). Refer to section 2.2 for more detail.

7.2.2 The relationship between data literacy and other literacies

According to the findings reported in section 5.9.1, more than 30% of participants also lack proficiency in the listed literacies. A majority of the respondents confirmed that they have not yet had training. This highlights a need to think holistically about research literacy. It would therefore mean incorporating these literacies as framework actors.

Information literacy is defined as the ability to locate, evaluate, and use information effectively and ethically, while digital literacy is the ability to use digital tools and technologies to find, evaluate, create, and communicate information (Deja, Rak & Bell, 2021; Sample, 2020). These skills are essential for researchers as they need to be able to navigate the vast amounts of data available to them, as well as to effectively communicate their findings to various stakeholders. A study by Olakunle and Olanrewaju (2019) found that information literacy was positively associated with research productivity, suggesting that it is a critical skill for researchers to have (see Section 2.3.1 for a more detailed discussion).

On the other hand, statistical literacy is defined as the ability to interpret, analyse, and communicate statistical information. Researchers need to be able to understand statistical concepts and methods to analyse data accurately and effectively communicate their findings. However, the low level of familiarity with statistical literacy as presented in the findings, suggests that there may be a lack of emphasis on statistical training in research education.

There is need for more effective education in both statistical and digital literacy for researchers. This is in line with the literature that was presented in section 2.3 where Schield (2005) argues that it is hard to separate statistical literacy, information literacy, and data literacy since they all deal with similar issues. This argument was supported by Ridsdale et al. (2015) when they confirmed in their study that data literacy is based on the same theoretical principles as statistical and information literacy.

Conclusion: Therefore, from the analysed literature, the three literacies are interrelated and are essential in dealing with the challenges that researchers face in research.

7.2.3 Priorities in the development of a data literacy training program

This study recognises that data literacy is becoming increasingly important as organizations of all sizes collect and use data to inform decision-making processes (Giese et al., 2020; Pothier & Condon, 2020; Wolff, 2016). However, not everyone is adept in interpreting and working with data, so developing data literacy training programs must be prioritized. In this regard, based on the response of researchers, the key areas that should be prioritised in the development of a data literacy training program have been identified and presented in section 5.10.

From the findings, one of the key areas that respondents indicated that should be prioritized in the development of a data literacy training programs is data organization and management. According to the responses of 58 researchers, developing an appropriate data management plan is essential in ensuring that data is efficiently organized and managed. In the study by Mannheimer (2018), it is argued that proper data management planning is critical in reducing the risk of data loss or corruption and promoting data sharing and collaboration. There is need for efficient data processing tools and methods to ensure that data is appropriately managed. In the inclusion of data organization and management in the implementation a data literacy program, data collection methods, data storage, and data preservation should be considered as essential subareas.

213

Data analysis and statistics are critical areas that should be prioritized in data literacy programs, as noted by 68 respondents. According to Islam (2020), data analysis is an essential process that enables researchers to extract meaningful insights from data sets. Data literacy training programs should incorporate digital literacy skills, data processing techniques to enable researchers to understand and use data effectively.

Data visualisation and presentation are also an essential area to be prioritized in data literacy training programs, as noted by 33 respondents. Data visualisation and presentation enable researchers to communicate complex data sets in a more understandable and meaningful way. Some studies (Park et al., 2022; Rahlf, 2019) have highlighted the importance of data visualisation in decision-making processes, arguing that effective data visualisation improves communication and understanding of data.

Data sharing and collaboration is another crucial area that should be prioritized in data literacy training programs, as noted by 29 respondents. Researchers emphasized the importance of open data publishing, data publication, data reuse, and data sharing. In their study, Hey et al. (2009) highlighted the benefits of data sharing and collaboration, including increased visibility, impact, and innovation.

According to 27 respondents, data ethics and governance should be prioritized in data literacy training programs. Ethical data collection, open science, and establishing copyrights are critical in ensuring that data is used responsibly and ethically. Janssen et al. (2020) stressed the importance of data governance policies in regulating data collection, storage, and sharing in their study.

Finally, data privacy and security is another crucial area to prioritize in data literacy training programs, as indicated by 10 respondents. According to Williams and Pigeot (2017), data privacy and confidentiality issues associated with data management are a significant concern that can affect public trust and perception. Therefore, data literacy training programs should incorporate privacy and confidentiality issues associated with data management, ethical data collection, and data protection measures.

Conclusion: The study's findings in this section emphasize the importance of prioritizing the development of a data literacy training program to provide individuals with the required skills to work effectively with data. Researchers have recommended including certain elements in the development of a data literacy course. Interestingly, these recommendations bear a striking resemblance to the content covered in a research data management course that was taught as

an interdisciplinary course at Bielefeld University (Wiljes & Cimiano, 2019) and which was presented in section 2.9.7 of this study.

In summary, a comprehensive data literacy framework that captures the key areas mentioned by respondents is critical in providing guidance in equipping individuals with the necessary skills to work with data effectively and responsibly.

7.2.4 Role of the library in spearheading data literacy

The three groups of respondents, namely Researchers, University librarians, and Research/reference librarians, were asked about the role of libraries in promoting data literacy among researchers. From the responses there is alignment among all respondents. As is clear in Figure 5.17, the majority of the researchers (90%) acknowledged the role of libraries in promoting data literacy, while all the university librarians (see Section 6.3.7) were positive towards the library spearheading the data literacy program in their universities. Research/reference librarians (see Section 6.4.8) highlighted that they have already been playing an active role in promoting data literacy among researchers and therefore their role remains to be vital in spearheading the data literacy program.

Furthermore, the researchers mentioned the current role the library plays in meeting its users' needs, library resources, the human capital in the library, and the current research support they give to researchers as key reasons why libraries should be allowed to spearhead and promote data literacy (see Section 5.12). As the direct recipients of these services, researchers noted that libraries play a pivotal role in promoting research, and since there is no research without data, libraries should be empowered to help researchers understand how to manage their data. The researchers also agreed that libraries have the resources, infrastructure, and expertise necessary to enhance data literacy. Finally, some researchers pointed out that libraries already offer research-related services, emphasizing their potential for fostering data literacy.

University librarians expressed support for the library driving the data literacy program, pointing out that librarians have the requisite skills and knowledge to do so. However, one respondent emphasized the significance of incorporating other key university stakeholders in the program.

Research/reference librarians highlighted their active role in promoting data literacy among researchers. They noted that they already provide research related services to researchers. They pointed out that the libraries have been in position to implement information literacy which has

been successful and for that reason, libraries were equal to the task of spearheading data literacy.

Overall, responses from the three groups of respondents emphasize the significance of libraries in promoting data literacy among researchers. Researchers and university librarians agree that libraries have the resources, infrastructure, and personnel to foster data literacy, while research/reference librarians emphasize the active role they already play in promoting data literacy among researchers. According to the responses, libraries are well-positioned to shepherd data literacy programs and play an important role in enabling data literacy among researchers- in universities.

Taking into account the discussion in section 2.11.1 the author wishes to reiterate the following three studies:

1. Guss (2016) who acknowledged the place of the library in supporting researchers in confronting data-intensive research challenges. The study affirms that academic libraries have always dealt with the management of data and they are better placed as coordinating centres in universities.

2. Koltay (2017) who argued that libraries are uniquely positioned to provide training and support for researchers in managing and using research data, given their expertise in information management and research support.

3. Pothier and Condon (2020) who emphasized the important role of librarians in promoting data literacy competencies among business students and the broader workforce. They argue that the ability to effectively manage and use data is becoming increasingly important in the business world, and that librarians are well-positioned to provide training and support in this area. They suggested that librarians can play a critical role in promoting data literacy by collaborating with faculty to develop data-related curricula, providing training and support in areas such as data management and analysis, and integrating data-related skills into existing information literacy programs. They finally indicated that librarians can contribute to the development of data literacy competencies by working with industry stakeholders to identify the specific skills and knowledge that are needed in the workforce. They suggested that librarians can play a key role in bridging the gap between academia and industry, and in ensuring that students are prepared with the necessary data literacy skills to succeed in their careers.

Conclusion: The role that libraries and librarians could play in working with data has been acknowledged over a relatively long period of time. The author notes that the increasing availability and complexity of research data requires that researchers have the necessary data literacy skills to effectively manage and use data in their research. Libraries can play a critical role in promoting data literacy by offering training and support in areas such as data management, data visualisation, and data analysis, and it appears to be unanimously accepted that such a role does belong in libraries.

7.2.5 Feedback on various aspects of the proposed framework

The three groups of respondents, namely Researchers, University librarians, and Research/reference librarians, were given the chance to engage with the proposed data literacy framework and provide their feedback on various aspects of the framework. This included their opinions on what made sense in the proposed framework, what they liked about the framework, any gaps they identified in the framework, and what they would exclude from the framework. The following is the discussion of their responses.

7.2.5.1 Opinions on what made sense in the proposed framework

The responses from the researchers (see Section 5:12a), university librarians (see Section 6.3.10a), and research/reference librarians (see Section 6.4.11a) indicated that the proposed data literacy framework is generally well-received and perceived as comprehensive. Most of the respondents did not have any objections to the framework and found it to be clear and useful. However, some concerns were raised with regard to clarification of some concepts, the role of stakeholders and the inclusiveness of the training.

One common theme across all three groups of respondents was the need for clarity on certain concepts. For example, some researchers did not understand the concept of FAIR data. It may therefore be useful to exclude such jargon terms from the finalised framework.

With regard to stakeholder component of the framework the request was that the framework should consider being explicit about the role of key stakeholders in data literacy and the need of involving other stakeholders who were absent from the framework. This is in line with previous studies that have emphasized the importance of collaboration and partnerships in promoting data literacy in universities with the need of including key stakeholders during its implementation (Burress, Mann & Neville, 2020). According to Carlson and Kneale (2011) collaborations between faculty and librarians are highly recommended as the best practice for teaching data literacy skills (see Section 2.11.3).

Conclusion: Overall, the responses from the researchers, university librarians, and research/reference librarians indicate that the proposed data literacy framework is a valuable resource for understanding data literacy. However, there may be areas for improvement, such as the inclusion of clear definitions and explanations of technical terms, and the need for collaboration and partnerships.

7.2.5.2 What respondents liked about the framework

From the responses given, the proposed data literacy framework received positive feedback from researchers, university librarians, and research/reference librarians as reported under sections 5.12b, 6.3.10b and 6.4.11b respectively. The framework was perceived by the respondents as being comprehensive, well-structured, clear, and relevant to the needs of researchers, including faculty members and students. Additionally, it was seen as being inclusive, involving relevant stakeholders in the data literacy process, and well-designed, providing guidance on the data literacy process and identifying the relevant competencies involved. For instance, University librarians appreciated the clarity and comprehensive coverage of aspects related to data literacy implementation, while research/reference librarians appreciated the inclusiveness of the framework, particularly the inclusion of information literacy and digital literacy.

Conclusion: Overall, all three groups of respondents agreed that the proposed data literacy framework has the potential to be implemented in university libraries, achieving the study's objectives and creating awareness among all research stakeholders. However, from the feedback received, it was necessary to make a number of minor modifications, which were added to the finalised version of the framework (see Figure 8.1).

7.2.5.3 Gaps in the framework

The three groups of respondents (Researchers, University Librarians, and Research/Reference Librarians) had the chance to review the proposed framework and were asked to point out any gap they felt was evident. Their responses were reported in sections 5.12b, 6.3.10b and 6.4.11b respectively. The majority of the respondents did not identify any major gaps in the framework, however, from the responses, the researcher was able to identify some suggestions that were made.

Researchers provided feedback on the roles of key players and the importance of intellectual capital. They suggested clarifying the roles of key players in the framework and incorporating the importance of intellectual capital. Their view on intellectual capital was to emphasise the

218

need for equipping librarians with the relevant skills in data related services in order to empower researchers too. They recommended ensuring that personnel responsible for RDM training received proper training. This is in line with a recommendation made by (Tenopir et al, 2017), that the provision of research data services in libraries requires skilled professionals and therefore if librarians are to play a key role, then they ought to be well skilled.

The researchers proposed minor alterations to the framework, such as clarifying the roles of key players, justifying the relevance of intellectual capital, and ensuring that individuals responsible for RDM training received sufficient training. Some of these ideas have already been mentioned in previous sections. According to research/reference librarians, technical skills are required for librarians who will be participating in data literacy instruction.

University librarians noted that the wider university community and staff, were not reflected in the framework, since the focus seemed to be mainly on a small group of faculty members and postgraduate students. One respondent proposed that faculty members involved in teaching research methods must be included as stakeholders because they interact with students who are conducting research. Published resources (Johnson & Steeves, 2019; Piracha & Ameen, 2018; Carlson et al., 2011) also highlighted the significance of involving faculty members in research data management activities, as they can play an important role in promoting data literacy among students and researchers (see Section 2.2.2).

Research/Reference Librarians identified gaps in the framework regarding technical skills and technology among those to be involved in the implementation of data literacy and the recipients of the training. However, the study, as captured in the framework provides a solution to lack of technical skills where researchers will be given some training in digital literacy. Given that data literacy and digital literacy are complementary concepts that researchers must recognize and strive for, as described in Section 2.3.3, it is important to note that a researcher who is proficient in data literacy must also be proficient in digital literacy in order to efficiently handle digital technologies for data manipulation. As a result, digital literacy is an essential component of data literacy. The implication is that researchers must be able to use new technologies to advance their research.

Conclusion: The responses from the three groups of respondents provided useful information to improve the data literacy framework. It is necessary to make provision for a clarified role for stakeholders, simplifying some of the description and by focusing on the importance of developing intellectual capital also for the data librarians.

7.2.5.4 What to exclude from the framework

The results of this question, as documented in the study's Sections 5.12c, 6.3.10b, and 6.4.11c, indicate that the framework was well embraced by the participants. The majority of respondents indicated that it is not necessary to exclude anything from the proposed data literacy framework. This conclusion implies that the framework was comprehensive enough to include most, if not all, of the necessary elements to consider when collaboratively developing a data literacy program. Instead of excluding anything, as was documented above, there were some suggestions on what to include or clarify.

Conclusion: Participants generally provided encouraging input on different aspects of the proposed framework. Nonetheless, there were concerns from a few of the respondents about the clarity of certain areas in the framework as well as the inclusion and role of some stakeholders in the framework. More education and training on research data management and associated ideas were stressed. Participants also stressed the need of explicit consideration of stakeholders' roles and involving a broader range of stakeholders in data literacy programs.

The absence of objections or desired major adjustments in the framework by respondents implies that the proposed framework is comprehensive and appropriate to address their needs. According to the findings, the suggested data literacy framework is well-designed and comprehensive, with only minor gaps reported by respondents.

7.2.6   Collaboration in the development of a single data literacy curriculum

Based on the responses provided by university librarians and research/reference librarians, reported in Sections 6.3.11 and 6.4.12, it is clear that they are willing to collaborate in the development of a single data literacy curriculum for standardized training for Kenyan researchers. All the respondents indicated their willingness to be part of the development single data literacy curriculum if requested to do so.

Based on the comments of the participants, a number of significant topics for consideration in the development of a data literacy training program were highlighted. The respondents suggested that key components of a data literacy curriculum must include:

- An introduction to data literacy,
- Data management concepts,
- Exploring and describing data,
- Analytical skills,

220

- Data literacy tools,

- Building a data culture,

- Digital literacy/technical competence,

- Basic level statistical skills,

- Bata awareness, and

- Data communication competence.

The suggestion made by university librarians and research/reference librarians were fully aligned with the key areas to be prioritised in the development of a data literacy training program as suggested by researchers in Section 5.10.

Respondents recommended that the library should play a facilitation and advocacy role, as well as an initiating and coordination function, in the development and implementation of data literacy. They also stated that libraries will play an important part in the establishment of a data literacy curriculum as well as the execution or implementation of researcher training. Furthermore, libraries would have to take part in developing policies as well as staff competency development.

The respondents identified other stakeholders that would play a key role in the development and implementation of a data literacy program in Kenyan universities. They suggested the inclusion of the ICT department, research department, university management and quality assurance department. They were of the opinion that it would be good to include external bodies such as the Commission for University Education and the National Commission for Science, Technology, and Innovation (NACOSTI). The suggestion of including external stakeholders such as the NACOSTI is based on the consideration that they have an interest in research conducted within the country (see Section 3.3.2). As a government agency meant to provide policy and guidelines related to the whole process of research it therefore has a role to play in the implementation of data literacy framework.

In conclusion, the responses provided by university librarians and research/reference librarians indicate that there is a willingness to collaborate in the development of a single data literacy curriculum for standardized training for Kenyan researchers. These respondents suggested key components that should be included in the curriculum and recommended that the library should play a facilitation, advocacy, initiating, and coordination role in the development and implementation of data literacy. They also identified other stakeholders such as the ICT department, research department, university management, Quality Assurance department, and

221

external bodies like the Commission for University Education and the National Commission for Science, Technology, and Innovation (NACOSTI) that would play a key role in the development and implementation of the program. These recommendations could be aligned with the key areas suggested by researchers in section 5.10, and they highlight the importance of collaboration among various stakeholders in the implementation of a comprehensive data literacy framework for researchers in Kenya.

7.2.7 Final comments

Participants stressed the significance of data literacy for researchers. Researchers reported that data literacy training would improve their data skills, boost research outcomes, and facilitate evidence-based decisions. To them, data literacy increases efficiency while working with research data, enables better data analysis, and ensures data privacy, confidentiality, and reuse.

The participants emphasized the interdependence of data literacy, information literacy, and statistical literacy, underlining the importance of successful education in all three domains.

Participants highlighted the importance of libraries in advancing data literacy. Researchers, university librarians, and research/reference librarians all realized that libraries have resources, infrastructure, and expertise that may help promote data literacy among researchers. There was a strong consensus that libraries are well-positioned to provide training and assistance in areas such as data management, data visualization, and data analysis.

Conclusion: Overall, the study highlights the importance of creating a comprehensive data literacy framework which values the role of librarians in encouraging data literacy among academics. The findings emphasize the critical areas to prioritize in data literacy training programs and provided significant insights into the development of an effective framework that equips researchers with the required abilities to work effectively and responsibly with data.

7.3 Assessment of data literacy needs

Several findings were reported in Chapters 5 and 6 in relation to researchers' (faculty and postgraduate student) data literacy needs. This researcher cross referenced the findings from both chapters while discussing the findings. Different questions from the questionnaires and the interviews provided response to this objective. These findings are corroborated by literature reports, and as such, the subsequent eight points are emphasized.

7.3.1 RDM services offered by libraries

Data literacy is becoming increasingly important for researchers across a wide range of disciplines. Researchers must conduct high-quality research and make significant contributions to their respective fields. From the literature it was evident that data collection, data management, data analysis, data visualisation, data ethics, and data communication are all key elements of data literacy training (Goben & Griffin, 2019; Koltay, 2017). Previous studies (Cox et al., 2017. Boateng, 2015; Tenopir et al., 2015) have emphasised the critical role of libraries can play in supporting research data management, including providing training, infrastructure, and guidance on best practices for data management.

Based on the findings of the study, the following can be stated as data literacy needs for researchers:

7.3.1.1 Assistance in data creation

The process of generating or producing data is referred to as data creation. Capturing raw data from original sources or producing new data through experimental or analytical methods are examples of data creation. It entails collecting data, generating, or acquiring by various approaches such as surveys, experiments, observations, simulations, or data mining (Chigwada, Chiparausha & Kasiroori, 2017). According to the findings, as was captured in Figure 5.2, respondents expressed that librarians did not provide support in creating data, despite the availability of other data-related services as mentioned by Research librarians (see Figure 6.3). There is a definite need for libraries to assist researchers in data creation processes and they appear not to be doing so currently.

7.3.1.2 Data processing support

In data management, data processing refers to the activities and processes conducted on data in order to change the data into a useable and relevant form. It entails a set of methods and techniques for organizing, manipulating, and analysing data in order to extract important insights and information (Rudo, 2013). A significant proportion of respondents felt that libraries did not offer various data processing services, except for a few such as data entry, data description, and data translation (see Figure 5.3). Lack of these services was confirmed by Research librarians (see Figure 6.3). Researchers require more comprehensive support in data processing tasks.

7.3.1.3 Data analysis services

In data management, data analysis services relate to the processes and activities involved in reviewing, interpreting, and drawing relevant insights from data. These services are aimed at obtaining important information and knowledge from data sets in order to aid in decision-making, problem-solving, and strategic planning. Data analysis services primarily entail the exploration, analysis, and interpretation of data using various approaches, methodologies, and technologies (Bhat, 2019). According to the findings, a majority of the respondents believe that libraries did not provide data analysis services, with only a few specific services being accessed by a small percentage of participants (see Figure 5.4). There is a need for libraries to offer more extensive support in data analysis to meet researchers' needs.

7.3.1.4 Data preservation services

Data preservation services are the practices and procedures that ensure the long-term storage, protection, and accessibility of valuable and essential data (Kumar, 2018). These services are concerned with preserving the integrity, authenticity, and usability of data over long periods of time, generally for archive or historical purposes. The findings indicate that researchers recognize the importance of data preservation, particularly data publication in institutional repositories. However, a notable proportion of participants believe that libraries did not offer the listed preservation services (see Figure 5.5) despite this being one of their data literacy needs.

7.3.1.5 Access to research data

Access to research data is an essential component of data management as it facilitates data sharing, collaboration, and the validation of research findings (see Section 2.5.8). A significant percentage of respondents (see Figure 5.6) indicated that enhancing access to research data was not considered a prominent service provided by libraries.

Conclusion: From the discussion above it is concluded that the study demonstrates a mismatch between library services data literacy needs expressed by researchers. It is evident that researchers have specific requirements in data creation, processing, analysis, preservation, and access that libraries are not adequately meeting.

7.3.2   Developing of a Data Management Plan (DMP)

Data management, which includes data gathering, organization, preservation, and sharing, is an essential component of research. Effective data management has been described as requiring

224

the implementation of data management plans (DMPs). However, researchers are still slow to adopt DMPs (Gajbe, Tiwari & Singh, 2021). As discussed in Chapter 2, Section 2.5.1, creation of a DMP is becoming every researcher's need in view of the requirements by funders and even some publishers.

A small percentage of respondents showed a clear grasp of DMPs, according to the results shown in Figure 5.7, while the majority of the respondents were convinced, that they had not created a DMP. Issues, such as a lack of awareness, a lack of institutional support, and a lack of expertise or experience in data management procedures, can be blamed for the low adoption of DMPs among researchers (see Section 2.5.1). These results are in line with earlier research that found comparable obstacles to DMP implementation (Borycz, 2021; Pardo & Siemens, 2014). Through training and education initiatives mean to encourage the use of DMPs, the lack of awareness and inadequate institutional support can be addressed. The results indicate that such programs should highlight the advantages of DMPs, such as their greater effect and visibility, and offer helpful advice on developing DMPs.

According to an earlier study (Tenopir et al., 2015), the limited adoption of DMPs may also be due to the absence of clear criteria and procedures for their creation. It is challenging to compare and assess the efficacy of various plans as a result of variances in the content and format of DMPs due to the lack of a standardized approach (Jones, Pryor & Whyte, 2013). The DMP Tool, which offers guidelines on the content and format of DMPs, is one effort that has been undertaken to create standardized DMP templates.

Conclusion: The study's findings reveal a mismatch between researchers' data literacy requirements and the services provided by libraries. Data creation, comprehensive data processing support, data analysis services, data preservation services, and improved access to research data are all needed by researchers. To close this gap, libraries should increase their help in these areas in order to better fulfil the requirements of researchers and improve data literacy.

7.3.3 Creating metadata

Metadata is information that describes other data. It is information that provides an organized and standardized manner to describe data and its properties. It explains other data. The format, location, size, creator, creation date, and other information pertinent to the data's content and context are all provided in a systematic and standardized manner (Qin & D'ignazio, 2010b).

According to the findings (Figure 5.8), just a small proportion (less than 18% of the respondents) were certain they had created metadata, while the vast majority were certain they had not. This indicates a lack of awareness or comprehension of the significance of metadata in data management. The low percentage of respondents who verified creating metadata has consequences for data discoverability, reuse, and sharing in research settings and should be something to worry about in research considering the importance of metadata in research data management and research in general. A study by Elouataoui and Gahi (2022) emphasise the importance of metadata in enhancing data management practices and enabling data sharing and reuse.

The lack of knowledge emerged as a significant theme, with some respondents reporting not having encountered metadata in their research work or not knowing how to create it. This finding highlights the need for education and training on metadata to improve data management practices. According to Koltay (2017) lack of knowledge in creating metadata indicates that there is a need for training programs to enhance data literacy skills among researchers including training researchers in metadata.

The study findings indicate that there may be a lack of awareness or understanding of the importance of metadata in data management among researchers. Furthermore, the findings indicate that researchers are in need of the skill in view of it being a requirement unfortunately they like the skill of creating it.

Conclusion: The lack of knowledge and low percentage of respondents who confirmed they had created metadata have implications for the discoverability, reuse, and sharing of data.

7.3.4 Where researchers store their research data

The media in which participants saved their research data was one of the areas evaluated to determine researchers' data literacy needs. According to the study findings (Figure 5.9), personal laptops were the leading storage medium for research, followed by USB sticks, PC hard drives, and portable hard drives. Although cloud storage remained the most popular virtual option, physical storage devices were still favoured in general.

This study's findings are consistent with earlier studies indicating that researchers prefer to have control over their data storage devices (Rafiq & Ameen, 2022; Bunkar & Bhatt, 2020; Majid, Foo, & Zhang, 2018). This is due to the fact that physical storage devices enable researchers to access their data fast and readily without relying on an internet connection or

third-party services. However, because physical storage devices are susceptible to loss or theft, this preference for physical storage devices may pose a risk to data security.

The popularity of cloud storage among survey respondents is a positive indicator, indicating that virtual storage choices are becoming more generally accepted. The limited usage of library servers and external servers, on the other hand, underlines the need for additional inquiry into the reasons underlying this trend. It's possible that university libraries don't provide this service, or that respondents are unaware of the benefits of virtual storage choices, or that they are concerned about the security or accessibility of their data on these platforms.

Conclusion: Overall, the findings of this study imply that researchers should be better educated on the benefits and potential risks of various data storage options.

7.3.5   Perceived competence in various areas of research data management

The importance of efficient research data management cannot be overstated. The study assessed the data literacy needs of faculty and postgraduate students as researchers, with a focus on their perceived competence in various areas of research data management such as data planning, data collection, data processing, data analysis, data preservation, data sharing/publishing, data re-use, FAIR data, and ethical data collection. Figure 5.10 depicts the detailed results of their responses.

Conclusion: According to the findings, a substantial proportion of respondents have limited or somewhat competent data management skills, with the lowest levels of competence recorded in FAIR data, data re-use, and ethical data gathering. These findings have implications for academic institutions in Kenya and beyond.

As was reported in section 2.6 of this study, FAIR data principles aim to make research data findable, accessible, interoperable, and reusable, which is crucial in fostering the sharing and reuse of research data (Jacobsen et al., 2020). Similarly, as discussed in section 2.5.7, data re-use is critical for maximizing the value of research data by allowing it to be used in future studies or analyses. Finally, ethical data collection is critical in ensuring that research data collection is responsible and trustworthy, and researchers must follow ethical guidelines to maintain the integrity and validity of research findings.

Second conclusion: the study's findings show that present data management techniques in the institutions assessed may be insufficient or inadequate. This is contrary to what should be the ideal case in institutions conduct research. Tenopir et al. (2015) state that proper research data

management techniques are crucial in guaranteeing the integrity, quality, and accessibility of research data.

Third conclusion: the low levels of competence indicated in data planning, data collection, data processing, data analysis, data preservation, and data sharing/publishing indicate that academic institutions will have to implement more comprehensive data management training programs.

7.3.6 Most common research data services sought by researchers

Previous studies have highlighted the increasing importance of data analysis in various research fields (Hart, 2019; QuestionPro, 2019; Sivarajah et al., 2017). They indicate that researchers require assistance in using data analysis tools to help them make sense of the large amounts of data they collect and analyse.

In order to determine some of the data literacy needs of researchers, research librarians were asked to indicate the research-related services most frequently requested by researchers in their respective institutions. According to the findings in Section 6.4.3, researchers require various data-related services to support their research activities. The most sought-after data service identified by the respondents was related to data analysis. Most researchers are requesting assistance in using SPSS.

Conclusion: According to results presented in Figure 5.4, researchers believe that libraries do not provide any of the data analysis services that were mentioned, suggesting that there may be a gap in the services being offered by libraries and the perceived needs of their users or. researchers may lack an understanding of the services provided by libraries.

Another data-related service that researchers require, according to research librarians, is assistance with the collection and evaluation of data. This finding corroborates the finding under figure 5.10 which revealed that only 27.6% (149) of the researchers felt that they were highly competent at data collection. It could be deduced that a majority of the researchers exemplified the need to know how to carry out data collection. Some studies have highlighted the importance of data collection and evaluation in research activities (Avuglah, 2019; Chigwada, Chiparausha & Kasiroori, 2017). In the case of this study, researchers require assistance in developing effective data collection and evaluation strategies to ensure the quality and reliability of their research findings.

Conclusion: Similarities between the findings from the researchers and the research librarians include the importance of data analysis in research activities, the need for assistance in data

collection and evaluation, and the need for searching for data to support research activities. Both sets of findings highlight the importance of providing support for researchers in managing research data and improving the quality and reliability of research outcomes.

7.3.7 The importance of the data needs

The respondents' data literacy needs were also highlighted when they were provided with data literacy statements and asked to express the level of agreement with the statement. The findings, as shown in Figure 5. 14, demonstrated a substantial lack of awareness of open data and FAIR data.

Furthermore, the majority of respondents expressed a desire to preserve their research datasets after the project was completed, indicating a recognition of the importance of long-term data preservation. However, respondents expressed a lack of confidence and willingness to share their research data. Respondents also expressed a desire for clarity and direction on best practices for data management and metadata generation that are targeted to their individual university and research needs. The findings are congruent with those shown in Figure 5.10, which shows that a large proportion of respondents reported having inadequate or just marginally good data management skills.

Conclusion: Given the vital relevance of efficient data management in modern research, this highlights the need for institutions to expand investment in increasing the skills and knowledge of their researchers in research data management.

This is consistent with earlier research that emphasize the importance of increased education and support in the area of data management and sharing (Carlson et al., 2011). There is a substantial gap between what researchers are capable of doing with data and what they actually accomplish, according to Dallmeier-Tiessen et al. (2014). This is due, in part, to a lack of data management and sharing knowledge and skills. Tenopir et al. (2015) discovered that many researchers have a limited grasp of data sharing and may be unwilling to submit their data owing to fears of data misuse, loss of control, and lack of recognition.

Previous research backs up the study's claim that there is a need for clarity and guidance on best practices for data management and metadata development. Bierer, Crosas, and Pierce (2017), for example, believe that researchers need clear, succinct, and actionable guidance on how they should handle their data. Similarly, Jones, Pryor, and Whyte (2013) found that

developing suitable data management and sharing policies and standards is critical to enabling researchers to manage their data efficiently.

The strong preference for transparency and ethical data sharing techniques shown by respondents is consistent with the increased emphasis on open science and data sharing. The FAIR principles, which were unfamiliar to the majority of respondents, encourage the adoption of open and transparent data practices to improve scientific reproducibility and innovation (Wilkinson et al., 2016). As more researchers recognize the benefits of open science, there is a growing need for education and assistance in implementing open and transparent data policies.

Conclusion: The study's findings indicate the need for further education and support in the field of data management and sharing. The respondents' lack of knowledge of open data and FAIR data highlights the need for increased awareness and training in these areas. The respondents' wish to preserve their research datasets after the study has ended emphasizes the need for long-term data preservation. However, respondents' lack of comfort and willingness to share their research data highlights the need for greater attention to concerns about data misuse, loss of control, and lack of recognition.

7.3.8 Challenges faced by researchers while working with data

The management of research data is a critical component of the research process. It entails gathering, processing, analysing, and storing research data. Because of the large volumes of data created during research projects, research data management has grown increasingly crucial in recent years. However, managing research data presents several challenges that can have an impact on the quality and reliability of research results. In this discussion, we will look at the findings about researchers' data literacy needs, with a particular emphasis on the issues that researchers have when working with research data and how the issues may be addressed.

Figure 5.15 depicts the results, which show that researchers confront considerable obstacles in handling research data. According to the findings, the most difficult aspect of research data management was data analysis using various statistical software, followed by the development of data collection instruments, data privacy and confidentiality issues, locating datasets, processing collected data, preserving data, developing an appropriate data management plan, creating metadata, and storing data. Previous research has identified similar challenges in research data management (Birkbeck, Nagle & Sammon, 2022; Hamad, Al-Fadel, & Al-Soub, 2021). Each of the challenges is discussed in detail below.

7.3.8.1 Analysis of data using various statistical software

The most challenging element of research data management, according to majority of the respondents, is data analysis especially using various statistical software. According to the findings of the two university librarians, as reported in section 6.3.6, it appears that the institutions have made data analysis tools such as SPSS and NVIVO available to researchers in their computer labs. However, it is clear that researchers may face a skill gap that limits their ability to use these tools effectively for data analysis. A lack of knowledge and abilities in statistical analysis software may have a negative impact on the credibility and reliability of study findings.

Data analysis, as stated in Section 2.5.4, comprises tasks such as data interpretation, derivation, and development of research outcomes from data. This finding is consistent with previous research that has recognized data analysis as a major challenge for researchers (Borgman, 2012; Tenopir et al., 2011).

Conclusion: In order to resolve this challenge, researchers must be trained and supported in the use of statistical analysis software.

7.3.8.2 Development of data collection instruments

Another significant challenge mentioned by participants (62.2%) was the development of data collection equipment. This challenge may be attributed to the lack of standardization in data collection instruments, which can affect the comparability of research outcomes. Developing data collection instruments as captured under Section 2.5.2 entails developing tools that will be used to gather data for a specific research topic, such as surveys, questionnaires, or interviews (Chigwada, Chiparausha & Kasiroori, 2017). Earlier studies have emphasized the importance of well-designed data collection tools in creating credible research data (Creswell, 2014; Leedy & Ormrod, 2013) hence calling for training of researchers in developing good instruments for data collection.

Conclusion: It is critical for a researcher to understand how to create data-collecting instruments because the quality of the data obtained is heavily influenced by the equipment used to collect it. A poorly designed instrument can result in erroneous or inadequate data, jeopardizing the validity and trustworthiness of the research findings.

7.3.8.3 Privacy and confidentiality issues associated with data

Privacy and confidentiality issues associated with data were also identified as a significant challenge in research data management by 61.2% of the respondents. This challenge may be

231

attributed to the sensitive nature of some research data, which may require special handling and protection. This finding is consistent with the growing concern in the digital era concerning data privacy and confidentiality (Islam, 2020; Borgman, 2012) hence researchers requiring training on how to navigate and get it right when it comes to issues related to data privacy and confidentiality.

Conclusion: Given the importance of privacy and confidentiality in research data management, it is critical that researchers obtain enough training to handle the issues raised by these concerns.

7.3.8.4 Locating datasets

Locating datasets was cited by 60.3% of respondents as a key issue in research data management (see Figure 5.15 This challenge can be linked to a lack of accessible repositories and metadata standards, which makes it difficult for researchers to identify relevant datasets. The lack of central data repositories was also mentioned by research librarians in Section 6.4.12 when asked to name some of the aspects within their organization that could be enablers in the implementation of a data literacy program.

Conclusion: Finding datasets is one of the notable challenges in research data management. In reality, researchers have a difficult time finding and accessing relevant datasets because there are no accessible repositories and consistent metadata.

7.3.8.5 Preserving data

Preserving data was identified by 51.8% of the respondents as a significant challenge in research data management (see Figure 5.15). Data preservation involves the curation, conservation, and safeguarding of a document, while also ensuring that it remains accessible and usable for future purposes (Kumar, 2018). Even though this was identified as a challenge, a large majority of the respondents, 78.5% (see figure 5. 14), expressed a desire to store their research datasets beyond the lifetime of the project, indicating an appreciation for the value of long-term data preservation.

Conclusion: This implies that, while data preservation may be a challenge, researchers are aware of the need to preserve their data for future use. Effective data preservation helps ensure that significant research data is accessible to future research and can contribute to scientific knowledge advancement. This challenge may be correlated with a lack of consistency in data preservation techniques as well as adequate training of researchers which can have an impact on the longevity and accessibility of research data.

232

7.3.8.6 Developing an appropriate data management plan

Developing an appropriate data management plan was identified by 51.4% of respondents as a challenge in research data management. This challenge may be attributed to a lack of knowledge and skills in developing data management plans, which can affect the organization and documentation of research data as well as funding for projects in future. This research emphasizes the need of offering data management training and education, particularly when building data management plans. Information related to this challenge has previously been presented in Section 7.2.2.

Conclusion: The lack of awareness and understanding of developing data management plan could be attributed to a lack of institutional support for data management planning or a lack of emphasis on data management in research training programs.

7.3.8.7 Creating metadata

According to the findings, creating metadata is a major challenge in research data management, with nearly half of the respondents (49.7%) describing it as such. This finding is supported by an earlier finding (Figure 5.8), which revealed that the vast majority of respondents (69.9%) had never created metadata. The low number of respondents who had developed metadata implies a lack of awareness as to why metadata is important as well as little comprehension of the use of metadata in data management. Metadata are essential for ensuring that data can be found, understood, and reused by others (Beretta et al., 2021; Garnett et al., 2017).

Conclusion: The potential implications of this finding for data discoverability, reuse, and sharing are immense, especially in research contexts where data is frequently shared and reused by others. As a result, the low number of respondents who developed metadata implies that there may be a need for more awareness and training on the use of metadata in data management.

7.3.8.8 Storing data

According to the findings reported in Figure 5.9, respondents are already aware of and engaged in data storage. At 40.1%, storing data was identified as the least challenging component of research data management (see Figure 5.15). Researchers mentioned a personal laptop, USB sticks, PC hard drives, and portable hard drives as their major storage media, However, data storage remains a critical aspect of research data management because the accessibility and usability of research data is dependent on the quality and reliability of research outcomes (Hart et al., 2016).

233

Conclusion: To address the issue of reliable data storage, it is critical to provide secure and dependable data storage solutions that are acceptable to the researchers and that will enable long-term access and usability of research data.

In summary: The assessment of faculty and postgraduate data literacy needs allowed the researcher to identify several data literacy needs in relation to research data services offered by libraries. These needs include:

i.   Assistance in data creation: Researchers expressed a need for libraries to support them in the process of creating data, as many libraries do not currently provide this service.

ii.   Data processing support: Researchers require more comprehensive support in data processing tasks, beyond basic services like data entry, description, and translation.

iii. Data analysis services: Majority of the respondents believed that libraries did not provide adequate data analysis services, highlighting the need for libraries to offer more extensive support in this area.

iv.   Data preservation services: Researchers recognized the importance of data preservation, particularly data publication in institutional repositories. However, many participants felt that libraries did not offer sufficient preservation services.

v.   Access to research data: Enhancing access to research data was not considered a prominent service provided by libraries, according to the respondents.

In addition to these needs, researchers also expressed a lack of awareness and understanding in certain areas:

i.   Developing data management plans (DMPs): Only a small percentage of respondents were familiar with DMPs, indicating a need for more awareness, institutional support, and expertise in data management procedures.

ii.   Creating metadata: Researchers had a limited understanding of the significance of metadata in data management, indicating a need for education and training on metadata creation.

The storage of research data was also explored, with a clear indication that there was limited usage of library servers and cloud servers, suggesting a need for education on the benefits and potential risks of various data storage options.

The discussion revealed that there is limited competence in FAIR data principles with the majority expressing a lack of awareness of open data and FAIR data principles, data re-use, and ethical data collection. There was also a need for more comprehensive data management

training programs in academic institutions, covering areas such as data planning, collection, processing, analysis, preservation, and sharing/publishing.

The most sought-after research data services identified by researchers included data analysis and assistance with data collection and evaluation. There was a mismatch between the services offered by libraries and the perceived needs of researchers.

Their needs were further revealed in the challenges faced by researchers while working with data which include data analysis using various statistical software, development of data collection instruments, data privacy and confidentiality issues, locating datasets, processing collected data, preserving data, developing appropriate data management plans, creating metadata, and storing data.

7.4   Assessment of the organisational infrastructure

The second goal of the study was to identify the organizational infrastructure that fosters data literacy practices within the selected universities. Organizational infrastructure refers to the underlying systems, structures, and procedures that enable an organization to function and deliver services or products. It includes all of the various components that allow an organization to carry out its activities, achieve its goals and objectives, and meet all the requirements of its stakeholders. Physical infrastructure, information infrastructure, human infrastructure, financial infrastructure, and legal and regulatory infrastructure are common components of organizational infrastructure (Chión, Charles & Morales, 2020).

All of these components work together to support the organization's activities and enable it to achieve its goals and objectives. Having a strong organizational infrastructure is essential for the effectiveness and sustainability of any specific function.

7.4.1   Physical infrastructure

Physical infrastructure includes the physical facilities, equipment, and resources that an organization needs in order to carry out its operations. In the context of this study, the library is at the centre of physical infrastructure. The participants were asked various questions to collect relevant information regarding the physical infrastructure of their organisation that could enable data literacy. Libraries form part of the general framework of organisational infrastructure. In the context of reviewing organizational infrastructure that supports the feasibility of enabling data literacy, libraries can be seen as a crucial physical infrastructure

component. Other than providing dedicated spaces for where researchers can access a variety of resources, libraries can play a major role in promoting effective data literacy.

As it was discussed in Section 2.11.1, globally, some libraries have innovated their services to the level of providing separate space for researchers, developed research collections, offering training in database search to researchers, and even providing research data management instruction services. They play a crucial role in promoting data literacy and supporting research data management activities.

Figure 5.17 shows that the vast majority of the Kenyan respondents, recognize the important role that libraries can play in increasing data literacy among researchers. Respondents believe that libraries have the resources, technology, and personnel to assist researchers in their work with data. They cited the availability of a diverse range of resources, cutting-edge technology, knowledgeable librarians, and easy accessibility as key aspects in the library's potential to promote data literacy. Furthermore, some respondents emphasized the significance of libraries in data preservation, reuse, and improving data discoverability.

Respondents generally believe that libraries are well-positioned to foster data literacy among researchers. This finding is consistent with prior studies highlighting the critical role that libraries may play in encouraging research data management (Borgman, 2012; Mooney & Newton, 2012).

In summary, libraries serve as a critical physical infrastructure component within the organizational structure by providing data literacy resources, instruction, and support. Their involvement in promoting data literacy and supporting research data management activities emphasizes the importance of libraries in the broader context of organizational infrastructure that will enable all researchers to be fully data literate.

Conclusion: The recognition of the critical importance of the existing physical infrastructure available from libraries emphasizes the importance of prioritizing data literacy instruction and support in order to allow good data management and promote the benefits of data literacy.

7.4.2 Information infrastructure

Information infrastructure includes the information systems, data management processes, and communication networks that an organization relies on to support its activities. Examples may include computer systems, databases, and communication channels. This section discusses the

236

research data services provided by libraries and the challenges encountered by librarians while offering these services.

Libraries form part of the general framework of organisational infrastructure. In the context of reviewing organizational infrastructure that supports the feasibility of enabling data literacy, libraries can be seen as a crucial physical infrastructure component. Other than providing dedicated spaces for where researchers can access a variety of resources, libraries can play a major role in promoting effective data literacy.

7.4.2.1 Research data services provided by the library

Provision of research data services is singled out in this study as an organisation structure being a process towards achieving data literacy among researchers. In this study, the provision of research data services is singled out as an organizational structure that is a pathway toward establishing data literacy among researchers. According to the responses of university librarians and research/reference librarians (see Figure 6.1 and Figure 6.3 respectively) regarding research data services provided by libraries, while libraries play a role in assisting researchers with their data management needs, there is a need for improvement among the selected universities in terms of offering a broader range of services. None of the respondents stated that their library presently assists in the creation of data management plans or the creation and management of institutional data repositories. Similarly, none of the respondents said their library helps with metadata creation for data sets or digital data archiving and preservation. These services are critical for organizing and preserving research data for future use. This could imply a lack of library services, as these are essential components of managing research data.

Several studies have emphasized the significance of providing research data services in libraries (Safdar, et al., 2023; Perrier & Barnes, 2018; Tenopir et al. 2017; Tenopir et al, 2015; Tenopir et al, 2014). According to a study conducted by Perrier and Barnes (2018), research data management is critical in the research process. As a result, the study calls for university libraries to play an important role in research data management.

According to the study findings, libraries are currently offering essential assistance to researchers by providing information on institutional policies linked to research data management. A significant proportion of respondents in both the university librarian and research/reference librarian categories reported this.

Given the numerous policies and regulations that researchers must navigate in terms of research data management, such guidance is critical. Libraries can help researchers better understand

237

their responsibilities and obligations, as well as guarantee that data is managed and disseminated appropriately, by assisting with institutional policies relating to research data.

The study also indicated that certain libraries provide assistance in navigating intellectual property and privacy issues relating to research data. This assistance can be quite helpful to researchers as they attempt to manage the complicated legal and ethical issues surrounding their research.

Training researchers on data management activities is another service that some the libraries provide. This is significant since many researchers may lack the essential abilities to adequately handle their data. Researchers may lack knowledge of data management methods, and giving data management training and education can help researchers manage their data more effectively (Al-Jaradat, 2021; Ng'eno & Mutula, 2018; Perrier & Barnes, 2018).

According to the study, only one respondent in the university librarian group claimed that their library provided institutional data repositories for the long-term development and maintenance of research data. This conclusion shows that infrastructure for long-term storage and access to research data has to be improved. Infrastructure for data storage and access is critical because it ensures that research data remains preserved and easily accessible to future researchers. As a result, libraries should emphasize the setting up of efficient and long-lasting infrastructure for data storage and access.

It should be noted, however, that comments from university librarians and research/reference librarians at the same institution may differ. For example, whereas the university librarian at one institution mentioned institutional data repositories, the research/reference librarian at the same institution did not. Such disparities may represent differences in roles and duties within libraries and may necessitate further investigation to better understand the reasons behind them.

Finally, the findings of the study highlight the importance of providing effective and sustainable infrastructure for long-term storage and access to research data. Libraries ought to keep exploring ways to improve their infrastructure for data storage and access, while also taking into account potential disparities in responses among librarian groups. Libraries can support researchers manage and preserve their research data in such a manner, ultimately promoting more transparent and reproducible research practices.

While libraries have attempted to put structures in place to assist researchers with their data management needs, there is still room for improvement in terms of providing a broader range

of services. It is critical for effective research data management to provide services such as assistance in developing data management plans, setting up and running institutional data repositories, providing metadata for data sets, and digital data archiving and preservation. However, it is encouraging to see that the majority of libraries are already providing important services such as policy advice and intellectual property assistance, as well as data management training for researchers.

Conclusion: The findings emphasize the significance of research data services provided by libraries in fostering data literacy among researchers. There is still room for improvement in areas such as institutional policy guidance, intellectual property issues, and data management training that libraries now provide. Libraries must broaden their services to cover critical issues such as designing data management plans, establishing institutional data repositories, creating metadata for datasets, and archiving and preserving digital data. The findings highlight the importance of good and long-term infrastructure for long-term storage and access to research data. Disparities in responses among librarian groups show that more research is needed to understand differences in roles and responsibilities within libraries.

7.4.2.2 Data literacy-related support throughout the research process

Data literacy training is an essential part of the research process because it empowers researchers with relevant ability to manage, publish, and share their data successfully. One of the purposes of this study, as previously stated, was to assess the technical infrastructure readiness of the participating university libraries in Kenya for the implementation of data literacy training. The level of support provided by libraries to researchers in managing research data, data publishing and sharing, and access to data management tools was one means of determining technical infrastructure readiness. The study's findings demonstrate gaps in library assistance in several areas, indicating a need for reform.

Managing data: According to the findings, the majority of participants, as shown in Figure 5.11, were dissatisfied with the level of assistance offered by libraries when managing research data. Respondents were particularly dissatisfied with the availability of research data management guidelines, training modules, and data management training events. Inadequate data management practices may emerge from a lack of support in this area, resulting in data loss, duplication, and poor data quality (Borghi & Gulick, 2022). Prior studies have highlighted the importance of libraries assisting with research data management (Borghi & Gulick, 2022;

Frederick, 2019). Such a provision necessitates that the library should have a solid technical infrastructure to support this undertaking.

Data Publishing and Sharing: According to the study, 71.3% of respondents (Figure 5.12) were ignorant of open data as a publishing option, the availability of data publishing guides, and advice on copyright issues associated with data publishing. Previous research has underlined the importance of libraries providing assistance with data publishing and sharing (Cox et al., 2017; Harris, 2012). Lack of support in this area could result in limited access to research data, impeding the growth of research within various fields (Xu et al., 2022).

Data management tools: Findings as captured in Figure 5.13, revealed that the majority of respondents (96.3%) expressed dissatisfaction with the level of access to certain data management tools offered by libraries, specifically, 'ready to use' data management plans and tools for online collaboration among researchers. The need for libraries to provide support in data management tools has been highlighted by previous studies (Semeler, Pinto & Rozados, 2019; Cox et al., 2017). The lack of access to these tools can result in inefficiencies in data management and collaboration among researchers (Borgman, 2012).

In light of the technical services provided by the library, these findings complement conclusions from the interview with research librarians. While the library provides certain training interventions to researchers, research/reference librarians believe there is room for growth in terms of providing more diversified training options in areas such as research data management and archiving. This is consistent with earlier studies that emphasize the importance of libraries providing research data management support (Majid, Foo, & Zhang, 2018; Sewell & Kingsley, 2017; Borgman, 2012).

The lack of training interventions in research data management and data publication is concerning, given the importance of these skills in today's data-driven world. Researchers are expected to handle and manage their data responsibly, which requires them to be proficient in data management practices and ethical considerations (Tenopir et al., 2015). Therefore, libraries have an important role to play in providing researchers with the necessary skills and knowledge to handle their data effectively. Furthermore, the provision of these technical data services is dependent on the availability of technical infrastructure within the library or the institution at large. Nevertheless, it is good to see that several libraries provide training in statistical data analysis tools, research metrics, and data visualization. Such skills are required

for researchers to effectively interpret and present their findings, which can lead to better outcomes for research and higher impact (Xu et al., 2022; Borgman, 2012).

In summary, the study highlights the critical role that libraries play in supporting researchers' data management needs through investment in research-related technical infrastructure and the provision of training interventions. Libraries should invest in technical resources and partnerships with research institutions to create a platform for the dissemination of training materials and resources. This includes offering assistance in managing research data, data publishing and sharing, and providing access to data management tools. Incorporating data literacy training into the research process would facilitate effective data management practices, improve access to research data, and promote collaboration among researchers. Libraries can use digital tools to create interactive training modules and also adopt a proactive approach by promoting what they offer to researchers through targeted marketing campaigns. This study therefore emphasizes the need for libraries to provide diverse training interventions in research data management and archiving to enhance researchers' data literacy skills. By leveraging targeted marketing campaigns and partnerships with academic departments and research groups, libraries can ensure that researchers are aware of the resources available to them. Ultimately, these efforts can enhance researchers' competent research skills.

Conclusion: The study identifies inadequacies in library support for researchers in research data management, data publishing and sharing, and access to data management tools. Participants were dissatisfied with the degree of support provided by libraries in these areas, indicating that there is room for improvement. Libraries play an important role in assisting researchers with their data management requirements, and their investment in research-related technical infrastructure and training interventions is critical. Given the importance of these abilities in today's data-driven society, the absence of training interventions in research data management and data publication raises a concern. Nevertheless, it is encouraging to discover that some libraries offer training in statistical data analysis tools, research metrics, and data visualization. Nevertheless, it is encouraging to discover that some libraries offer training in statistical data analysis tools, research metrics, and data visualization. To improve their support for researchers' data literacy abilities, libraries should invest in technical resources, collaborations with research institutes, and focused marketing initiatives.

7.4.2.3 Libraries addressing data literacy training needs and services

Data literacy has become a critical skill in today's digital age, and libraries play a crucial role in supporting researchers in this area. The responses from the University librarians indicate a varying degree of attention given to addressing data literacy training needs and services in their libraries. However, their responses highlight the need for libraries to recognize and address the importance of data literacy training among their users.

One of the responses (Section 6.3.8), as a way of addressing researchers' data literacy training needs and services indicates that this is done through the Information Literacy program. This approach aligns with the notion that data literacy is a subset of information literacy, which focuses on skills and competencies required to effectively find, evaluate, and use information (Wanner, 2015). By integrating data literacy into information literacy programs, libraries can ensure that researchers have the necessary skills to manage and use data in their research (Calzada & Marzal, 2013).

Another response suggested that there is a lack of focus on data literacy training in their library. This echoes the call to prioritize data literacy in libraries (Fontichiaro & Oehrli, 2016). One of the respondents indicated that library seminars offer a potential platform for addressing data literacy training needs, which aligns with the findings of Dai (2020) on the value of library seminars in engaging with researchers on various topics, including data literacy. Incorporating data literacy training into library seminars can ensure that researchers have access to the necessary skills and knowledge to manage and use data effectively.

Another response revealed that while library staff may offer some level of assistance in data literacy-related queries, researchers are often referred to other departments for further support. However, we need to stress the importance of providing comprehensive data literacy support in libraries in view of the fact that some users consider the library as a one-stop-shop. This involves offering data literacy training and support to ensure that researchers have access to the necessary skills and knowledge to manage and use data effectively.

In summary, the responses highlight the need for libraries to prioritize data literacy training and support for researchers. Incorporating data literacy training into library seminars or workshops can broaden the reach of this training, while providing comprehensive data literacy support can ensure that researchers have access to the necessary skills and knowledge.

Conclusion: Responses received from university librarians regarding data literacy training in libraries show a variety of approaches to meeting the demands of researchers. While some

242

libraries incorporate data literacy into their information literacy programs, others see the importance of focusing more on data literacy training. Library seminars and workshops have been highlighted as suitable training platforms for data literacy. There is, however, a request for libraries to provide full data literacy support rather than referring researchers to other departments for help. Overall, libraries must prioritize data literacy training and support to ensure that researchers have the essential skills and knowledge to handle and use data efficiently.

7.4.2.4 Challenges encountered by librarians

Librarians play a critical role in providing research data-related services to researchers, including training and assistance in using data analysis tools, accessing e-resources, and providing consultations. However, librarians face various challenges when providing these services to researchers. The challenges identified from the interviews as recorded under Section 6.4.10 include researcher skill gaps, difficulty in satisfying researcher needs, perception of librarians, lack of skills and manpower, understanding of researcher needs, working with an older generation of researchers, and internet downtime.

Researchers have identified a skill gap among researchers resulting from a lack of relevant or fundamental abilities, such as the basic knowledge of ICT required to properly use data analysis tools. The challenge has come to light as a consequence of the time-consuming and challenging nature of training in research data management. According to Sabzwari, Bhatti, and Ahmed (2012), researchers must have fundamental ICT skills for both educational and research reasons, emphasizing the need of equipping them with these skills. A comprehensive training program that covers basic ICT skills and data analysis tools should be designed to address this issue. There is need to provide specialized training programs that meet the needs of this group of researchers. Researchers' lack of fundamental ICT skills highlights the value of digital literacy, which should be incorporated in the development of a data literacy training program.

Another challenge that librarians face is identifying and fulfilling researcher needs. According to respondents, some researchers have definite needs and do not listen to alternatives, making it difficult for librarians to meet those needs. To solve this problem, librarians must have good communication skills as well as an understanding of the researcher's research requests. Librarians must collaborate closely with researchers to explore alternate data sources or tools that may be useful.

Researchers' perception of librarians also poses a challenge for librarians in providing data-related services. Researchers may have a perception that librarians are not capable of providing data-related services, which can hinder the collaboration between researchers and librarians. To overcome this challenge, librarians need to demonstrate their expertise in data-related services by publishing articles, conducting research, and attending conferences related to data services (Fuhr, 2022).

Librarians also face a lack of relevant skills and manpower. These are skills related to data literacy, which are crucial for providing data-related services. Librarians need to develop data literacy skills to provide effective data-related services to researchers. This can be achieved by attending training programs, taking online courses, and participating in professional development opportunities (Fuhr, 2022; Tenopir et al., 2015). As indicated in Figure 6.2, librarians should take advantage of existing staff development policy and find capacity building opportunities and equip themselves with relevant skills and knowledge to support research data services.

Understanding the demands of researchers was also cited as a concern for librarians. Respondents mentioned encountering researchers who have no knowledge about what they want, making it difficult to provide appropriate data-related services. To address this issue, librarians need to engage with researchers and undertake needs assessments to better understand their research needs.

Lastly, internet downtime poses a significant challenge for librarians in providing data-related services to researchers. Respondents mentioned that internet downtime can cause delays in providing data-related services, which can negatively impact research projects. To overcome this challenge, librarians need to develop contingency plans to ensure that data-related services can be provided even when there is internet downtime. This is where the leadership of the library needs to present the needs of the library to the general management of the university for the purpose of providing relevant infrastructure that would assist in meeting users' needs.

Librarians play a vital role in providing data-related services to researchers. However, they face various challenges that need to be addressed to ensure effective delivery of services. By addressing these challenges, librarians can provide effective data-related services to researchers and contribute to the success of research projects.

In summary, the discussion made in this section highlights the relationship between information infrastructure and the provision of data literacy services by libraries. The provision

of research data services, such as assisting researchers with data management, creating data management plans, establishing institutional data repositories, metadata creation, and data archiving and preservation, is an essential component of managing research data. These services require a robust information infrastructure, including computer systems, databases, and communication networks, to support data storage, access, and preservation. Furthermore, the challenges faced by librarians, such as researcher skill gaps, meeting researcher needs, researchers' perception of librarians, lack of relevant skills and manpower, understanding researcher needs, and internet downtime, all have implications for the information infrastructure required to support data-related services. For example, addressing skill gaps and providing training programs may involve the development and deployment of appropriate technological infrastructure and resources. Ensuring a reliable and uninterrupted internet connection is also crucial for delivering data-related services. A robust information infrastructure is essential for supporting research data services, addressing researchers' data literacy training needs, and overcoming challenges faced by librarians in delivering these services effectively.

Conclusion: When offering research data-related services to researchers, librarians encounter a number of obstacles, including skill gaps among researchers, difficulty meeting researcher needs, perception issues, a lack of necessary skills and manpower, understanding researcher needs, and internet outages. These challenges highlight the importance of addressing digital literacy and data literacy training for researchers, improving communication and collaboration between librarians and researchers, demonstrating librarians' expertise in data-related services, developing relevant skills and manpower, conducting needs assessments, and ensuring a reliable information infrastructure.

7.4.3   Human infrastructure

Human infrastructure refers to the people who work within an organization, including their skills, knowledge, and experience. It includes elements such as the organizational culture, leadership, and workforce development programs.

7.4.3.1 Availability and ability of university and research/reference librarians

University librarians and research/reference librarians play crucial roles in ensuring effective and efficient provision of library services to students, faculty, and researchers in the university setting. The findings of the study as recorded in Section 6.3.1 reveal the connection between the role of university librarians and the implementation and managing data literacy. According

to the findings, university librarians are responsible for overseeing the library and its activities, policy development, staff management, coordination of library services, communication with other parts of the university, and participation in university management. Their roles are pegged at enhancing the university's strategic goals by providing vision, strategic direction, and leadership for the library, and advancing the research agenda and enhancing the research profile and ranking of the university.

The research/reference librarians (see Section 6.4.1), on the other hand, provide services such as research-related services, teaching information literacy and referencing styles, answering reference questions, assisting with research, uploading institutional repositories, providing circulation services, and making resources available to students. Generally, in their roles, they are offer services to researchers too.

Both university librarians and research/reference librarians play crucial roles in the provision of library services in the university setting. University librarians oversee the library and its activities, policy development, staff management, coordination of library services, communication with other parts of the university, and participation in university management, while research/reference librarians provide direct research support services to postgraduate students and faculty members, including teaching research skills, assisting with literature searches, providing access to relevant information resources, and offering guidance on referencing styles. These roles are critical in ensuring that researchers have access to relevant and up-to-date information resources to support their learning, research, and academic endeavours.

These findings agree with the literature that highlight the significant role libraries and librarians can play in the implementation of data literacy in universities. Some studies indicate that the library plays a key and practical role in providing research services which include research data management services (Brown, et al, 2018; Delaney & Bates, 2018).

Conclusion: University librarians and research/reference librarians play critical responsibilities in assisting with the delivery of library services in the university setting. University librarians are in charge of different areas of the library, such as policy development, staff administration, and service coordination, whereas research/reference librarians give direct research assistance to students and faculty members. Their duties are critical in ensuring that researchers have access to relevant information resources and acquire research skills and referencing help.

7.4.3.2 Organisational structures that support data management and data literacy

The recorded findings in Section 6.3.4a show that four respondents cited the presence of a research office or directorate focused on fostering evidence-based research for decision-making as an organizational structure focused on developing data literacy. However, one respondent stated that their institution lacked such a structure. According to Torres (2019), the importance of organizational structures in supporting research activities is widely acknowledged. As a result, this conclusion emphasizes the importance of universities establishing organizational structures that encourage data literacy in order to support efficient management and usage of research data.

According to the research, data management is a critical component of the university's organizational structure. One of the respondents mentioned that there was a relevant office in their university responsible for obtaining data for university decision-making, suggesting the institution's acknowledgment of the importance of data (see Section 6.3.4b). Nonetheless, the rest of the respondents indicated that they were unsure whether data management was regarded as an important component of the university's organizational structure. This emphasizes the need for enhanced awareness and education about the significance of data management in the academic setting.

The findings in Sections 6.3.4a and 6.3.4b call attention to the critical role of organizational structure in fostering data literacy by establishing appropriate and supportive frameworks. Previous research (see Section 2.1) by Chigwada, Chiparausha, and Kasiroori (2017), Patel (2016), and Liu and Ding (2016) indicated the growing popularity of data support services in institutions. They did, however, identify constraints such as a lack of institutional flexibility and senior management's lack of data literacy awareness. These obstacles may inhibit the effective implementation of data literacy programs and limit their potential impact.

The study discovered that training expenses, infrastructure costs, and facilities and technology costs are key factors when it comes to the financial implications of adopting a data literacy program, as recorded in section 6.3.4c.

The involvement of various key university departments, such as the research department, the ICT department, the school of graduate studies, research and extension, the institutional research office, and the directorate of research, was identified as critical to the success of the data literacy program (see Section 6.3.4d). Respondents described them as potential stakeholders who may play an important role in the adoption of data literacy. Furthermore,

247

faculty members who supervise postgraduate students were identified as important stakeholders.

These study findings align with previous literature on the importance of organizational structures and infrastructure in supporting effective data management and data literacy. According to Koltay (2014), institutions need to have appropriate technical and organizational infrastructures for the success of data literacy implementation. Universities need to establish organizational structures and processes that promote data literacy and management to facilitate effective use of research data. In addition, universities should also invest in the necessary infrastructure, facilities, and technology to support data literacy programs. Finally, collaboration with key university departments and faculty members is crucial in the success of data literacy programs. According to Yang and Li (2020), collaboration among stakeholders is essential in developing data literacy skills among researchers.

Conclusion: It is critical for institutions to build an organizational structure and setup that promotes data management and data literacy. According to the findings, having a research office or directorate focused on evidence-based research for decision-making is important in encouraging data literacy. Some respondents, however, mentioned the lack of a such structure in their institutions, emphasizing the necessity for universities to acknowledge and emphasize the importance of data management. Organizational structures should be built to encourage data management awareness and education, ensuring that it is acknowledged as a critical component of the university's operations.

7.4.3.3 Skills, competencies and knowledge gap in data literacy among librarians

Libraries play a vital role in supporting research, providing access to resources and services. However, as research becomes more data-intensive, librarians must have skills and knowledge in data literacy to effectively support researchers. It is the responsibility of the parent institution to the library to ensure that they have competent staff with up-to-date skills that can help achieve the goal of the institution. Unfortunately, several studies have revealed that many librarians lack this knowledge, posing a significant challenge for libraries seeking to support researchers' data management needs (Moran, 2019; Koltay, 2017).

According to the replies collected in Section 6.3.9, most librarians lacked the essential data literacy skills and knowledge. Respondents recognized the value of data literacy training and the necessity for librarians to become data literacy experts. Respondents proposed a variety of

strategies to increase librarians' data literacy skills, including training, capacity building, mentorship, and collaboration with other universities or external professionals.

The study also found that research/reference librarians have various levels of research data management abilities and expertise (see Figure 6.4). While some areas, such as research methodologies and the research lifecycle, were areas of strength for a significant number of respondents, others, such as data management plan development, require more attention and training. Furthermore, the findings indicate that research/reference librarians may benefit from additional training and support in areas such as data curation, technical data management skills, data description and documentation, institutional repository management, open access initiatives, and legal and copyright frameworks. Gaps in knowledge and skills, such as data literacy and ICT skills, were identified and suggest the need for ongoing training and support to enable librarians to effectively support researchers' evolving needs in the context of data-intensive research.

The view of lack of data literacy competence and the need for retraining for librarians is well captured by Koltay (2017). He argues that to achieve a data literate community of researchers, there is a need to retrain librarians to become data librarians who can offer the necessary training to researchers. This is because librarians are ideally positioned to bridge the gap between researchers and data resources. Koltay suggests that the training should focus on data curation, data management, data sharing, and data analysis. By providing this training to librarians, they can become key players in promoting data literacy among researchers, ultimately leading to a more data-literate community.

Several approaches can be taken to address this challenge, including training, dedicated capacity building, mentorship, and professional development programs. Libraries may work with external experts or universities to provide librarians with tailor-made training opportunities. Furthermore, while research/reference librarians have varying levels of expertise and knowledge in this domain, it is critical to provide them with training and support in areas such as data curation, technical skills in data management, data description and documentation, institutional repository management, open access initiatives, and legal and copyright frameworks in order for them to effectively support researchers.

Conclusion: The study reveals a significant gap in data literacy skills and knowledge among librarians, posing a challenge for libraries in supporting data-intensive research. Respondents recognized the importance of data literacy training and suggested various strategies to enhance

librarians' skills in this area. The findings highlight the need for ongoing training and support, particularly in areas such as data management, curation, and legal frameworks. Librarians, with their unique position to bridge the gap between researchers and data resources, have the potential to play a crucial role in promoting data literacy among researchers and fostering a data-literate community.

To summarize, an organization's human infrastructure is critical to promoting data literacy. University and research/reference librarians are examples of professionals who provide crucial services and assistance to students, faculty, and researchers. University librarians are in charge of the library and its activities, as well as policy creation, staff administration, and service coordination. Research/reference librarians offer direct research assistance, such as teaching research skills.

Data management and data literacy are significantly influenced by organizational setup too. Creating proper structures, such as research offices or directorates, can help to promote evidence-based research and decision-making. Collaboration with important university departments and stakeholders, such as the research department, ICT department, and faculty members, is essential for data literacy initiatives to succeed.

However, librarians have a skills and knowledge gap in data literacy. Many librarians lack the required data literacy skills and knowledge, making it difficult for libraries to serve researchers' data management needs. Training, capacity building, mentorship, and partnership with external specialists are all possible solutions to the challenge. To improve librarians' capabilities in areas like data curation, data management, data description and documentation, institutional repository administration, and legal and copyright frameworks, ongoing training and support are required.

7.4.4   Legal and regulatory infrastructure

An organization's legal and regulatory infrastructure includes policies and procedures, as well as the legal and regulatory framework within which the organization functions. An organization's policies are a set of rules, guidelines, and procedures that it establishes to regulate its operations and activities. These policies frequently address legal and regulatory needs, as well as ethical concerns and organizational norms. They establish the organization's legal and regulatory structure and provide guidelines for employees, researchers, and other stakeholders to follow. These rules can specify the roles and duties of various stakeholders in the company, as well as provide direction on data handling, storage, access, sharing, and reuse.

7.4.4.1   *Availability of institutional policies supporting RDM and data literacy*

Policies, as previously stated, are a component of organizational infrastructure. They are a set of principles, regulations, and procedures which govern individual behaviour and actions inside an organization. Policies can assist an organization in achieving its goals and objectives by establishing a framework for decision-making and directing employee behaviour. They can also aid in ensuring legal and regulatory compliance and promoting uniformity in decision-making across the organization. Policies can thus be viewed as an essential component of an organization's infrastructure. They are a tool that may be used by an organization to guarantee that its operations operate smoothly and efficiently. Effective policies can aid in the development of a positive organisation culture, the development of trust among stakeholders, and the mitigation of risks.

In today's world, when data is becoming increasingly available and vital in research, data literacy is an essential ability for researchers. A number of policies need to be in place for its proper execution. According to Avuglah and Underwood (2019), in order to establish or adopt research data related services, an institutional policy framework must be explicit and thorough.

According to the results of the study in Sections 6.3.3 and 6.4.6, universities lack the majority of policies that encourage data literacy activities. However, most institutions recognize the significance of policies that support research data services. The most common policies were research policy and research data management policy, with the majority of respondents saying that these policies were in place at their institutions. Research data management policies according to Huang, Cox, and Sbaffi (2021), can provide a framework for managing research data and can help to guarantee that data is appropriately documented, preserved, and disseminated. This can aid in promoting research integrity, increasing research visibility, and facilitating collaboration.

This is consistent with earlier research (Llebot & Castillo, 2023; Tenopir et al., 2015; Higman & Pinfield, 2015), which emphasized the need for policies to promote research-related services. These policies serve as the foundation for the creation of a data literacy framework. Though Chigwada, Chiparausha, and Kasiroori (2017), Patel (2016), and Liu and Ding (2016) as presented in Section 2.1, have previously conducted studies that have demonstrate the emergence of data support services some of the findings from the studies have revealed significant challenges related to data literacy. Some of the challenges highlighted include poor policies or no policies regarding Research Data Management (RDM).

251

The study discovered that, in addition to research-related policies, institutions recognize the role of technology in research and data management, with respondents saying that their institutions had ICT policies in place. ICT policies establish a framework for the management and application of technology in research, ensuring that researchers have access to the ICT tools and infrastructure required to efficiently manage their data. This is consistent with prior research that has highlighted the importance of technology in research data management (Chawinga & Zinn, 2021; Ng'eno & Mutula, 2018).

The study discovered that staff development policies were available in all of the institutions surveyed. Having a staff development policy in place is the first step in offering training and development opportunities to upscale their employees in order to ensure that they are able to provide quality research data services. Previous research has demonstrated the significance of staff training and development in research data management (Tang & Hu, 2019). Staff development policies, then, provide a framework for library staff training and development, ensuring that they have the essential skills and expertise to provide excellent research data services.

The presence of digital preservation policies in universities received a very low response, with only one respondent indicating that they have one. The low number of respondents who reported having such a policy shows that related policies regarding research data management and preservation may be required to enable successful research data management and preservation for future use. Preservation policies, as Al-Jaradat (2021) points out, can help to ensure that research data remains accessible and usable throughout time. This is especially significant considering the growing volume of research data produced and its potential worth for future study.

The findings indicate that universities are recognizing the necessity of policies to support research data services, with policies in place at the majority of institutions. All institutions had open access policies in place, demonstrating their commitment to making research data more widely available. This aligns with the global trend toward open science and open research data, which has been propelled in part by the potential benefits of making research data more broadly available (Vickers, 2011). This is an encouragement for the possibility of providing data literacy services since it provides a solid foundation upon which such services may be built.

While the study shows that universities have made progress in developing policies to support research data services, it also highlights areas where more progress is needed. More

specifically, more targeted policies involving research data management and preservation may be required to ensure that data is well managed and preserved for future use. The successful implementation of such policies can provide a solid foundation for the development of data literacy services, which can improve research integrity, increase research visibility, and facilitate collaboration across disciplines and institutions.

Conclusion: Finally, institutions must have policies that enable research data management (RDM) and data literacy. The findings reveal that, while some policies, such as research policy and research data management policy, are in place in many institutions, more comprehensive and targeted policies are still required. These policies establish the foundation for managing research data, ensuring its proper documentation, preservation, and dissemination. The presence of ICT policies also demonstrates the recognition of technology's role in research and data management. Furthermore, staff development policies are critical in preparing library staff members to provide quality research data services. The lack of digital preservation policies, on the other hand, underscores the necessity to develop policies that protect the long-term accessibility and usage of research data. While there has been progress regarding developing policies to promote research data services, further efforts are needed to target specific areas and establish a comprehensive policy framework that facilitates effective data management and data literacy practices.

In summary, this section of the study was aimed at assessing the organizational infrastructure for the feasibility of offering data literacy within universities. As discussed, this encompasses the physical infrastructure, information infrastructure, human infrastructure, and legal and regulatory infrastructure.

Regarding physical infrastructure, libraries were identified as one of the most important components of physical infrastructure. Researchers believe that libraries have the resources, technology, and personnel needed to promote data literacy. They were recognized for providing diverse resources, cutting-edge infrastructure, knowledgeable librarians, and easy accessibility. Libraries were also acknowledged for their role in data preservation, reuse, and improving data discoverability.

Information infrastructure, including information systems, data management processes, and communication networks, was examined. The study found that while libraries assist researchers with their data management needs, there is room for improvement in offering a broader range of services. Majority of the services offered in the selected libraries were not adequately

meeting the data literacy needs of the researchers. Libraries should focus on providing data management plans, institutional data repositories, metadata creation, and digital data archiving and preservation. Training researchers on data management activities was also identified as a valuable service.

The study emphasized the importance of effective and sustainable infrastructure for long-term storage and access to research data. Libraries should invest in technical resources and partnerships with research institutions to provide researchers with the necessary tools and support. Challenges faced by librarians included researcher skill gaps, difficulty in meeting specific researcher needs, researchers' perception of librarians, lack of relevant skills and manpower, understanding researcher needs, and internet downtime.

Human infrastructure, which includes the skills, knowledge, and experience of library staff, was also considered. University and research/reference librarians were highlighted as crucial in implementing and managing data literacy. University librarians oversee library activities and policies, while research/reference librarians play a role in offering training and assistance. The study emphasized the need for librarians to have relevant skills and knowledge in data literacy due to lack of relevant skills among the current librarians. The study suggested capacity-building opportunities for librarians in order to develop the much-needed skills.

The discussion also took note of the legal and regulatory infrastructure necessary for the implementation of data literacy training. It highlighted the importance of policies in providing a framework for data literacy activities and ensuring legal and regulatory compliance. The study revealed that while some universities have policies supporting research data services, there is a lack of policies specifically addressing data literacy. However, most institutions recognize the significance of research and research data management policies. These policies promote research integrity, data documentation, preservation, and dissemination. The study also emphasized the role of technology in research and data management, with ICT policies in place to support researchers. Staff development policies are acknowledged as essential for training library staff in providing quality research data services. On the other hand, the presence of digital preservation policies is limited, indicating the need for further policies to enable successful data management and preservation. The existence of open access policies demonstrates a commitment to making research data widely available. Overall, while progress has been made in developing policies, there are still areas requiring more targeted policies to ensure effective data management, preservation, and support for data literacy services.

© University of Pretoria

7.5 Assessment of the technical infrastructural readiness

Objective three of the study aimed to assess the technical infrastructure readiness of selected private university libraries in Kenya to implement data literacy training. In the current data-driven landscape, possessing data literacy skills is crucial for individuals to manage, interpret, and make informed decisions based on data. Libraries serve as essential repositories of knowledge, providing access to data and information. However, providing data literacy training necessitates adequate technical infrastructure, such as hardware, software, and internet connectivity. Thus, this objective sought to determine the extent to which selected private university libraries in Kenya are technically prepared to provide data literacy training.

According to the responses in Section 6.3.6, the availability of ICT infrastructure to support data collection, manipulation, and sharing varies across universities. While certain universities provide data analysis software such as SPSS and NVIVO, others do not. It should be noted that some researchers may be managing their data on an individual basis, regardless of whether or not ICT infrastructure is available.

Data analysis software were discussed in Section 2.5.4. briefly. The availability of data analysis software such as SPSS and NVIVO can greatly benefit researchers in conducting their studies (Ong & Puteh, 2017: Zamawe, 2015). SPSS is a popular quantitative statistical analysis software that allows researchers to analyse and interpret their data efficiently (Field, 2018). Similarly, NVIVO is a qualitative data analysis software that helps researchers organize and analyse unstructured data (Mortelmans, 2019). These software tools can facilitate the analysis of large datasets, which may be time-consuming and challenging to handle manually.

However, it is also important to consider the availability of other ICT infrastructure, such as data management tools and platforms, that can support data collection, manipulation, and sharing. For example, research data management systems can help researchers to manage their data more efficiently, by facilitating data organisation, storage, and sharing (Goldman, Chen & Palau, 2023; Urbano, Cagnacci and Euromammals Collaborative Initiative, 2021; Briney, Coates, & Goben, 2020). These systems can also ensure data security and compliance with ethical and legal standards.

The lack of dedicated ICT infrastructure, both hardware and software, for data collection, manipulation, and sharing can pose challenges for researchers. Without appropriate infrastructure, researchers may have to rely on manual methods for data collection and organisation, which can be time-consuming and prone to errors. Additionally, the lack of

© University of Pretoria

accessible data management tools and platforms can make it difficult for researchers to share their data with other researchers, which can hinder scientific progress.

Furthermore, most of the universities face a challenge of internet downtime. The research librarians, as captured in Section 6.4.10, mention this as a challenge while providing data-related services to researchers.

In summary, the objective of assessing the technical infrastructural readiness for the implementation of data literacy training in selected private university libraries in Kenya was to determine the extent to which data literacy could be enabled. The availability of ICT infrastructure to support data collection, manipulation, and sharing varies across universities. Some universities provide data analysis software like SPSS and NVIVO, while others do not. However, internet downtime poses a challenge for many universities when providing data-related services. Availability of other ICT infrastructure, such as data management tools and platforms, is important for efficient data management. The lack of dedicated ICT infrastructure can hinder data collection, manipulation, and sharing, leading to reliance on manual methods and hindering data sharing and collaboration among researchers. Overall, the study highlights the need for improved ICT infrastructure, software availability, reliable internet connectivity, and dedicated data management tools to enhance the technical readiness for data literacy training implementation in the selected private university libraries.

Conclusions: Based on the information provided, this study makes some conclusions regarding the technical infrastructural readiness for the implementation of data literacy training as follows:

1. The availability of ICT infrastructure to support data collection, manipulation, and sharing varies across universities. Some universities provide data analysis software such as SPSS and NVIVO, while others do not. This variation in availability may impact the ability of researchers to effectively analyse and interpret their data.

2. Many universities face the challenge of internet downtime, which can hinder the provision of data-related services to researchers. Reliable and uninterrupted internet connectivity is essential for researchers to access online resources, collaborate with peers, and share their research data.

3. The availability of dedicated data management tools and platforms is crucial for researchers to efficiently manage and share their data. These tools can facilitate data organization, storage, and compliance with ethical and legal standards. However, the

study does not explicitly state the extent to which these tools are available in the selected private university libraries.

4. The lack of dedicated ICT infrastructure, both hardware and software, can pose challenges for researchers in terms of data collection, manipulation, and sharing. Relying on manual methods for data management can be time-consuming and prone to errors. Additionally, the absence of centralized data management tools and platforms can hinder data sharing and collaboration among researchers.

Ensuring that all researchers, across all collaborating universities, have accesses to appropriate and equivalent ICT infrastructure may become a stumbling block to be managed carefully.

7.6   Summary

This chapter provided a comprehensive discussion of the findings of the study in view of the topic of the study which is: *The feasibility of offering standardised data literacy services at selected private university libraries in Kenya*. Since the study employed a mixed-methods approach to achieve its objectives, which included assessing the data literacy needs of faculty and postgraduate students, evaluating the organizational and technical infrastructures for data literacy training, and developing a framework to facilitate successful data literacy initiatives in Kenyan universities, a triangulation approach was used in the discussion.

From the discussion, the findings suggest a gap between the services offered by libraries and perceived data literacy needs, low adoption of data management plans (DMPs), and a need for education and training on metadata, data storage options, and researchers' perceived competence in various areas of research data management. The findings highlight the importance of providing comprehensive data literacy support to researchers, including training, capacity building, mentorship, and professional development programs. The findings emphasize the need for universities to prioritize data literacy instruction and support in order to promote transparent and reproducible research practices, improve research quality, and contribute to knowledge advancement in various fields. Ultimately, the study aims to develop a standardized data literacy framework to cultivate a community of data-literate researchers in Kenyan universities, with the crucial role of libraries in promoting data literacy among researchers emphasized throughout the discussion.

General recommendations of the study and the development of the data literacy framework will be presented in the next chapter (Chapter 8).

CHAPTER 8

8.   SUMMARY OF FINDINGS AND RECOMMENDATIONS

Following the in-depth discussion, triangulation, convergence, and conclusions offered in Chapter 7, Chapter 8 serves as the study's completion, providing a summarized depiction of the research's outcomes.

8.1   Introduction

This chapter provides a concise overview of the study's objectives and demonstrates their accomplishment. The chapter also outlines the essential elements of the research design and highlights the major findings pertaining to both the main research question and the sub-questions. Moreover, the chapter includes recommendations derived from the findings and proposes potential avenues for future research. Ultimately, it provides a comprehensive discussion of the developed data literacy framework.

8.2   Summary of study's research design

This section provides a summary of the entire study design employed in the research. The study aimed to investigate the current data literacy needs, assess the organizational and technical infrastructural readiness, and identify the framework requirements for implementing successful data literacy programs among faculty members and postgraduate students in selected private universities in Kenya. By examining these aspects, the study aimed to contribute to the development of effective data literacy initiatives in the library context. The research design encompassed a pragmatic research paradigm and utilized a mixed-method approach, combining both quantitative and qualitative data collection methods. The data collection instruments included an online questionnaire and interviews, while the study sites encompassed five private universities within the Nairobi Metropolitan area. The research process prioritized ethical considerations and data confidentiality throughout.

258

Table 8- 1 : Summary of the study's research design

Study title	The feasibility of offering standardised data literacy services at selected private university libraries in Kenya
Main objective	To develop a standardised framework that would support feasible data literacy services at Kenyan university libraries.
Specific objectives of the study	1. Critically review the existing literature on data literacy, in order to establish a comprehensive framework encompassing essential services and components that would facilitate an essential discussion about the successful implementation of a data literacy initiative in selected private universities in Kenya 2. Asses the data literacy needs of faculty and postgraduate students in selected private universities in Kenya. 3. Asses the organizational infrastructure of selected private university libraries in Kenya to determine the feasibility of offering data literacy services. 4. Assess the technical infrastructure readiness in the selected private university libraries in Kenya to facilitate the implementation of data literacy training.
Research sub-questions	1. How does the existing literature on data literacy inform the development of a comprehensive framework with essential services and components for the successful implementation of a data literacy initiative in selected private universities in Kenya? 2. What are the data literacy needs of faculty members and postgraduate students as researchers in selected private universities in Kenya?

259

	3. What is the existing organisational infrastructure within the selected private university libraries in Kenya that supports data literacy initiatives for researchers? 4. To what extent are the selected private university libraries in Kenya technically equipped to implement data literacy programs?
Methodology phase	This phase offered a practical guide to put the introduced concepts into action, covering methodologies and tools for realizing the study's goals, effectively connecting theory with practical implementation.
Research paradigm	A pragmatic research paradigm
Research approach	A mixed method approach utilizing a convergence design was employed as the proposed by Creswell (2014). The researcher collected and analysed both quantitative and qualitative data on the same phenomenon separately. Subsequently, the obtained results were compared and contrasted to identify any patterns of alignment or discrepancy. Findings were interpreted by integrating and converging the findings from both data sources and comparing these to the guidance received from literature.
Data collection instruments	Two data collection instruments were used in the study. Questionnaire: An online questionnaire was used to collect quantitative data. The researcher developed a set of questions specifically for researchers, including PhD students and faculty members. The

260

	questionnaire consisted of both closed- and open-ended questions. The inclusion of open-ended questions aimed to enrich and augment the value of the collected data.
	Interview schedules:
	In addition to the questionnaire, interviews were conducted. For this study, the researcher opted for semi-structured interviews. Two different interview schedules were developed to gather data from a total of five University Librarians and five research/reference librarians.
Literature review	Although the researcher initially began working on the literature review in May 2019, literature was consistently tracked for new and emerging studies throughout the study. Resources were continuously integrated into the literature review.
Study sites and participants	The researcher gathered data from five private universities located within the Nairobi Metropolitan area. These universities included the Catholic University of Eastern Africa, United States International University Africa, African International University, Adventist University of Africa, and Daystar University.
	The participants were categorized into three groups: the researchers, consisting of PhD students and faculty members; the university (head) librarians; and research/reference librarians.
	A total of 614 researchers received the questionnaire, while 5 university librarians and 5 research/reference librarians were interviewed.
Data collection	Data collection was carried out from October 2022 to December 2022

261

	An online questionnaire was used to collect data from researchers. Participants were asked to indicate whether they provided informed consent for their data to be used in the research before they gained access to the questions.
	The researcher conducted interviews with the 6 university librarians and research/reference librarians. To ensure that the interviewees were well prepared and had a clear understanding of the interview process, they were sent the interview questions and a consent form at least one week prior to the scheduled interview. The interviewees had the choice of selecting the location for the interview, with most of the interviews being conducted at their place of work, except for one interviewee who chose to conduct the interview at a hotel.
	To maximize efficiency and minimize disruption, the interviewer carried out interviews with both staff members, based at a specific university, on the same day. The interviews lasted between 25-30 minutes, and prior to the start of the interview, the interviewer sought permission from the interviewee to record the session. The purpose of recording the interviews was to ensure that the data collected was accurate and could be used for later analysis.
Data ethics clearance	The study adhered to the ethical guidelines set by the Faculty of Engineering and Built Environment (EBIT) at the University of Pretoria. Prior to data collection, permission was requested and obtained from the university, ensuring compliance with ethical requirements, Reference number: EBIT/170/2022 (see Appendix IV).

	Furthermore, this research study fully complied with the ethical standards outlined by the National Commission for Science, Technology & Innovation (NACOSTI), which is responsible for issuing research permits in Kenya, Reference number: 660943 (see Appendix V). In addition, letters of permission to collect data were obtained from all the universities where the research took place. These letters have been included in the thesis's appendices (see Appendices VI, VII, VIII, IX, X) to provide evidence of ethical approval. To ensure informed consent, all participants were required to sign a form acknowledging their understanding and agreement to participate in the study.
Data confidentiality	The participants in this study were given assurances regarding the confidentiality of their participation and the anonymization of the collected data and its use for this specific study. Before the participants were provided with the self-administered questionnaire and before conducting the qualitative key informant interviews, a participant informed consent form was presented to address confidentiality concerns comprehensively. In order to maintain confidentiality, participants were not required to disclose any information that could potentially identify them. The researcher strictly adhered to the consent agreement and ensured that no confidential respondent information was included in the study, as doing so would breach the terms of consent.
Analysis	The data analysis involved descriptive statistics using SPSS, while thematic analysis was employed for the qualitative data.

Overall, the research design was carefully structured to address the study's objectives and provide a comprehensive understanding of the data literacy landscape in Kenyan universities. By combining quantitative and qualitative data collection methods and integrating literature findings, the study aimed to contribute to the development of effective data literacy initiatives in the library context, thus promoting good research data management in the region.

8.3   Research questions and sub-questions

Based on the research problem, the study was guided by the following central research question:

> *What would a generally accepted, evidence-based, broad framework of data literacy interventions, that includes policy, human resourcing and infrastructure development, (to be used by university libraries in Kenya), look like?*

To address the central research question, the study focused on the following sub-questions:

1. How does the existing literature on data literacy inform the development of a comprehensive framework with essential services and components for the successful implementation of a data literacy initiative in selected private universities in Kenya?
2. What are the data literacy needs of faculty members and postgraduate students as researchers in selected private universities in Kenya?
3. What is the existing organisational infrastructure within the selected private university libraries in Kenya that supports data literacy initiatives for researchers?
4. To what extent are the selected private university libraries in Kenya technically equipped to implement data literacy programs?

By addressing these questions, the study sought to contribute to evidence-based, effective data literacy programs and support the advancement of data literacy initiatives in the context of Kenyan universities.

8.4   Achieving study purpose and objectives

The purpose of this study was to assess the readiness of selected private universities in Kenya to provide data literacy services and to propose a generally accepted, evidence-based, broad data literacy framework for researchers in the country. By investigating these factors, the study aimed to make a meaningful contribution towards the establishment of effective data literacy initiatives within the library setting.

The broader objective was then to evaluate the overall feasibility and readiness of selected private university libraries in Kenya to provide standardized data literacy services. The overarching objective was subdivided into four specific study questions, which are outlined in Table 8-2 below.

Table 8- 2 : How the study answered each specific study questions

Specific study questions	Achieving the objective
How does the existing literature on data literacy inform the development of a comprehensive framework with essential services and components for the successful implementation of a data literacy initiative in selected private universities in Kenya?	The question was successfully answered, and the study findings have played a pivotal role in formulating a comprehensive data literacy framework that integrates essential requirements, derived from literature and study findings. This framework will ensure the viability and success of a data literacy initiative at universities in Kenya. The framework is accompanied by a process model that demonstrates the linear flow from the literacy need expressed by researchers and other stakeholders to the outputs of the process. In general, the study underlined the importance of developing a comprehensive data literacy framework and acknowledges the role of libraries in influencing data literacy among researchers. The findings emphasized the critical areas that should be prioritized in data literacy training programs and provided significant insights for establishing an efficient framework that will continuously equip researchers with the required abilities to deal with data efficiently and responsibly. All these have been included in the data literacy framework as captured in Figures 8.1 and 8.2.
What are the data literacy needs of faculty members and postgraduate students as researchers in selected private universities in Kenya?	The question was successfully resolved by presenting findings on the different data literacy needs of researchers, including faculty members and Ph.D. students (see Section 7.3). A gap exists between library services offered and perceived data literacy needs, there is a low adoption of Data Management Plans (DMPs), and a lack of

266

Specific study questions	Achieving the objective
	education and training on metadata, data storage options, and researchers' perceptions of competence in research data management, according to the study.
What is the existing organisational infrastructure within the selected private university libraries in Kenya that supports data literacy initiatives for researchers?	The question was resolved and the study presented findings on the organizational infrastructure of the selected private university libraries. The study revealed what was available at the selected university libraries, as well as gaps in organizational infrastructure that would aid in the viability of delivering data literacy services (see Section 7.4).

The findings emphasize the relevance of organizational infrastructure, policy, and library capabilities in fostering data literacy among researchers. By implementing particular policies, expanding library services, and providing training and support, it is possible to improve data management practices and research quality. |
| To what extent are the selected private university libraries in Kenya technically equipped to implement data literacy programs? | The goal of the objective was met, allowing the study to present findings on the state of technical infrastructure readiness in the selected private university libraries, which is critical for the implementation of data literacy training. The findings not only identified existing resources, but also showed deficiencies in the necessary technical infrastructure required to meet researchers' data literacy needs (see Section 7.5).

The study's findings show that there are gaps in the technical infrastructure's appropriateness for implementing data literacy. In conclusion, the study underlines the importance of better ICT infrastructure, software availability, consistent internet |

Specific study questions	Achieving the objective
	connection, and specialist data management solutions. These improvements are required to enhance the technology readiness for data literacy training.

268

In conclusion, this study successfully achieved its purpose of assessing the readiness of selected private universities in Kenya to provide data literacy services and proposing a comprehensive data literacy framework for researchers. By addressing the specific objectives outlined in Section 8.2, the study shed light on the data literacy needs of faculty and postgraduate students, examined the organizational and technical infrastructure within university libraries, and identified essential elements for a successful data literacy initiative. The findings emphasize the importance of developing a holistic data literacy framework that incorporates policy, organizational support, and technical infrastructure. This framework is crucial for fostering effective data management practices and enhancing researchers' data literacy skills, ultimately contributing to the advancement of open science and research quality within Kenyan universities.

8.5   Answering the research sub-questions

The findings below are organized and presented according to the research sub-questions that guided the study (see Section 1.6).

8.5.1   Data literacy framework

The first sub-question that had to be addressed was: How does the existing literature on data literacy inform the development of a comprehensive framework with essential services and components for the successful implementation of a data literacy initiative in selected private universities in Kenya?

The major finding of the study is that there is a need for a comprehensive data literacy framework to facilitate the development of a community of data literate researchers in Kenyan universities. The study identified key areas that should be prioritized in the development of a data literacy training program, including data organization and management, data analysis and statistics, data visualization and presentation, data sharing and collaboration, data ethics and governance, and data privacy and security (see Sections 5.11 and 7.2.3). The findings highlight the importance of libraries in spearheading data literacy initiatives, as they have the resources, infrastructure, and expertise to support researchers in managing and using research data effectively. The feedback from researchers, university librarians, and research/reference librarians indicate that the proposed framework was generally well-received and comprehensive, but there is a need for clarification of certain concepts and the explicit inclusion of key stakeholders. Overall, the study emphasized the significance of data literacy in

269

promoting effective data management, research outcomes, and evidence-based decision-making among researchers in Kenyan universities.

8.5.2 The data literacy and needs of faculty members and postgraduate students

The second sub-question that had to be addressed was: What are the data literacy needs of faculty members and postgraduate students as researchers in selected private universities in Kenya?

The study identified a number of data literacy needs of faculty members and postgraduate students which include, assistance in data creation, data processing support, data analysis services, data preservation services, access to research data, Developing Data Management Plans (DMPS), creating metadata, data storage options, competence in research data management, challenges with data analysis, developing data collection instruments, privacy and confidentiality issues, locating datasets, ethical data sharing and reuse, FAIR data principles, data storage solutions, comprehensive data management training, awareness of data management best practices, assistance with data analysis, awareness of open science and data sharing, and handling of research data challenges (see Sections 7.3.1.1-7.3.1.5 or the entire Section 7.3). The findings indicate that there may be a gap between the data literacy needs of researchers and the services provided by libraries. The majority of respondents felt that libraries did not assist in creating data, providing data processing services, or offering data analysis services. While data preservation, particularly data publication in the university's institutional repository, was recognized as an important service, a significant proportion of participants still believed that libraries did not offer adequate preservation services. Moreover, enhancing access to data was not considered a prominent service provided by libraries.

The study findings suggest that there is a need for libraries to increase their support for data creation, processing, and analysis services, as well as improve their efforts to promote and provide access to research data.

Furthermore, the study revealed a lack of awareness and adoption of Data Management Plans (DMPs) among researchers. Only a small percentage of respondents showed a clear grasp of DMPs, indicating issues such as a lack of awareness, institutional support, and expertise in data management procedures. The findings suggest the need for training and education initiatives to promote the use of DMPs and address the barriers to their implementation.

Additionally, the study highlighted the lack of knowledge and understanding of metadata creation among researchers. The low percentage of respondents who confirmed creating

270

metadata indicates a need for education and training on metadata to improve data management practices and enable data sharing and reuse.

Researchers also expressed limited competence in various areas of research data management, particularly in FAIR data, data reuse, and ethical data gathering. These findings emphasize the need for institutions to invest in growing researchers' skills and expertise in these areas, as well as providing comprehensive data management training programs that cover all stages of the research data life cycle.

Based on the researchers' data literacy needs, the study identified various data-related services sought by researchers, such as assistance with data analysis and data collection and evaluation. However, there was a perceived gap between the services already provided by libraries and the needs of researchers. To address this gap, libraries should improve communication and marketing activities to make researchers aware of the data-related services they offer.

Overall, the study's findings indicate the need for training and support in data management and sharing, including awareness and training in open data and FAIR data principles. Researchers expressed a desire for clarity and direction on best practices for data management and metadata generation tailored to their research needs. Institutions should invest in increasing researchers' skills and knowledge in research data management and address challenges faced in handling research data, such as data analysis using statistical software.

By addressing these findings and improving data literacy services, libraries and academic institutions in Kenya can better support researchers in managing research data, enhancing the quality and reliability of research outcomes, and promoting open science and data sharing practices.

8.5.3   Existing organisational infrastructures

The third sub-question: What is the existing organisational infrastructure within the selected private university libraries in Kenya that supports data literacy initiatives for researchers?

The study findings indicate that selected private university libraries in Kenya have varying levels of organizational infrastructure to support data literacy among researchers (see Section 7.4). The organizational infrastructure includes the roles and capabilities of university librarians and research/reference librarians, the presence of organizational structures supporting data management and data literacy, the availability of institutional policies supporting research data management (RDM) and data literacy, the role of libraries in

271

promoting data literacy, the research data services provided by libraries, and how libraries are addressing researchers' data literacy training needs and services.

University and research/reference librarians play crucial roles in providing library services and supporting data literacy. University librarians oversee library activities, policy development, staff management, and coordination of services, while research/reference librarians offer research-related services, teaching information literacy, and providing access to resources. From the findings, all the universities have university librarians in place who are aware of their administrative role. Furthermore, they have research librarians. However, the study reveals lack of relevant data management technical skills among research/reference librarian. The study notes the need to retrain research/reference librarian through professional development programs in order to make them competent enough to attend to data literacy needs of researchers.

Organizational structures supporting data management and data literacy, such as research offices or directorates, are present in some universities but may be lacking in others. Establishing such structures is important for efficient data management and usage.

Institutional policies supporting RDM and data literacy are recognized as essential. While research policies and research data management policies are commonly in place, there is a need for more targeted policies to ensure effective data management and preservation.

Libraries are recognized as having a critical role in promoting data literacy among researchers. They are seen as having the resources, technology, and personnel to support researchers' data needs, including data preservation, reuse, and improving discoverability.

Research data services provided by libraries are varied and need improvement. While libraries assist with institutional policies, intellectual property and privacy issues, and some data management training, there is a lack of services such as creating data management plans, managing institutional data repositories, metadata creation, and digital data archiving.

Libraries should prioritize data literacy instruction and support, enhance their capacity to provide resources and expertise, and contribute to transparent and reproducible research practices. Libraries should continue to improve infrastructure for long-term storage and access to research data, including establishing institutional data repositories, to ensure data preservation and accessibility for future researchers. Libraries should find a way of introducing

272

data literacy trainings among researchers for instance through seminars which can serve as platforms for addressing data literacy training needs.

Overall, the findings highlight the importance of organizational infrastructure, policies, and library support in fostering data literacy among researchers. Implementing targeted policies, enhancing library services, and providing training and support can improve data management practices, and research quality.

8.5.4 The extent to which libraries are technically equipped to implement data literacy programs

The fourth sub-question that had to be addressed was: To what extent are the selected private university libraries in Kenya technically equipped to implement data literacy programs?

The findings of the study indicate that the selected private university libraries in Kenya have gaps in their technical infrastructure readiness for the implementation of data literacy programs (see Section 7.5). In terms of data literacy-related support provided by the libraries to researchers throughout the research process, there were dissatisfaction and inadequacies in several areas. Participants expressed dissatisfaction with the availability of research data management guidelines, training modules, data management training events, data publishing guides, advice on copyright issues related to data publishing, and access to data management tools. The lack of support in these areas can lead to inefficient data management, limited access to research data, and hindered collaboration among researchers. The study emphasizes the critical role libraries play in supporting researchers' data management needs through investment in technical infrastructure and the provision of training interventions. It suggests that libraries should invest in technical resources, partnerships with research institutions, and the creation of a platform for disseminating training materials and resources.

Libraries can improve researchers' data literacy abilities and promote competent research practices by including data literacy training into the research process and taking a proactive approach through targeted marketing efforts. The availability of ICT infrastructure, such as data analysis software and data management tools, was discovered to differ between universities, leading to challenges for researchers in data collecting, manipulation, and sharing. The study emphasizes the significance of dedicated ICT infrastructure for effective data analysis, organization, storage, and sharing, hence assisting researchers in their scientific progress.

8.6   Recommendations

Based on the major findings presented, the study puts forward the following recommendations which are presented in the broad order of the research study questions which are:

1. How does the existing literature on data literacy inform the development of a comprehensive framework with essential services and components for the successful implementation of a data literacy initiative in selected private universities in Kenya?

2. What are the data literacy needs of faculty members and postgraduate students as researchers in selected private universities in Kenya?

3. What is the existing organisational infrastructure within the selected private university libraries in Kenya that supports data literacy initiatives for researchers?

4. To what extent are the selected private university libraries in Kenya technically equipped to implement data literacy programs?

The gap between the services rendered and the services needed is addressed first. Most of the recommendations could be linked to skills development. The infrastructure is the focus of the subsequent set of recommendations, while recommendations, linked to the updated framework are provided last since this stands out as the key and main aim of this study.

8.6.1   Data literacy needs of faculty members and postgraduate students

In view of the data literacy of faculty members and postgraduate students, the study recommends the following:

Bridging the gap: In view of the existing gap between the services offered by the libraries and data literacy needs of researchers (see Section 7.3) the study recommends that libraries should increase their efforts and support in these areas to close this gap. Some of the areas include providing assistance in data creation, data processing support, data analysis services, data preservation services, access to research data, development of data management plans (DMPs) and creation of metadata.

Libraries should facilitate research activities and improve the quality and impact of research output by efficiently supporting research data management. Libraries can play an important role in bridging the gap between the data literacy needs of researchers and the services they are offered. Libraries should improve data literacy support by offering services that meet

researchers' needs, such as data creation assistance, comprehensive data processing support, data analysis services, and data preservation services.

Promoting the adoption of Data Management Plans (DMPs): To make researchers more aware of the importance of data management planning, the study recommends that training and education programs should be developed. Libraries can take the initiative in creating such events to boost DMP adoption. To make DMP implementation easier, institutional support should be improved. This includes supplying resources, skills, as well as explicit criteria and procedures for developing DMPs. Libraries can work with other institutional stakeholders to build standardized approaches to DMP creation. Furthermore, standard DMP templates, such as the DMP Tool, should be encouraged to achieve uniformity and comparability in DMPs across research initiatives.

Addressing the lack of knowledge in metadata creation: In view of lack of knowledge in creation of metadata which has implications on discoverability, reuse, and sharing of data, the study recommends provision of education and training on metadata are in order to improve data management practices and enhance data sharing and reuse. The library should assume the role, in collaboration with other related stakeholders to offer the training.

Providing infrastructure and training for data storage: Regarding storage of data, the study recommends that institutions should provide infrastructure that would provide research data storage. Furthermore, libraries should provide training on best practices for data management storage especially on virtual storage choices.

Investing in researchers' skills and expertise in data management: Considering low or lack of competences among researchers in various areas of research data management such as FAIR data, data re-use, and ethical data collection the study recommends the need for institutions to invest more in growing their researchers' skills and expertise in these areas through training. Institutions must set rules and provide data management training to guarantee that researchers have the essential skills and knowledge to efficiently handle research data. Furthermore, institutions must invest in data management infrastructure, such as data repositories and data sharing platforms. The study recommends that raining in data management should address all areas of the research data life cycle, from planning to sharing and archiving. Furthermore, data management training should be integrated into the research process to ensure that researchers have the skills and knowledge needed to effectively manage research data.

Improving communication and marketing of data-related services: In view of a discrepancy between what the libraries offer as data related services and the most commonly sought research data services by researchers the study recommends that libraries to improve their communication and marketing activities in order to make researchers aware of the data-related services they provide. In order to better satisfy the demands of their users, libraries may need to examine their present data-related services and suggest areas for development. Libraries should also consider collaborating with other institutions or organizations to provide extra data-related services, such as data analysis training or access to specialized software, that are beyond the scope of what the library can give on its own. Overall, libraries must remain sensitive to the changing needs of researchers and constantly evaluate and adjust their services to best support their users.

Establishing comprehensive data-related support services: It was evident that researchers need assistance with data collection and analysis. As a result, it is recommended that comprehensive data-related support services be established, including training on effective data collection, expert guidance on data analysis, the provision of data search resources, fostering of collaboration between researchers and librarians, and ongoing assistance with data management tools and guidelines. These recommendations can improve the research process and ensure the integrity of study outputs by providing researchers with appropriate assistance and tools for good data management. The respondents' request for clarification and guidance on best practices for data management and metadata development highlights the importance of policies and guidelines customized to the individual needs of researchers and their institutions. Finally, respondents' strong preference for transparency and ethical data sharing practices emphasizes the importance of open science and data sharing.

Following some of the challenges indicated by researchers that they were experiencing while working with data during research, the study recommends the following:

Providing training on statistical software usage: Having indicated that they experience challenges with analysis of data while using various statistical software, the study recommends that the library in conjunction with relevant stakeholders such as the ICT and research department should provide or introduce related training which can be delivered in the form of workshops or seminars. These training initiatives should be ongoing in order to keep researchers up to date when new software and/or tools are released.

276

Training and guidance for data collection instrument development: Regarding challenges in development of data collection instruments, the study recommends that training and guidance during research process. The library could organise training on development of data collection instruments especially for researcher who are at the point of collecting data. This could be done alongside the research office and in conjunction with the research supervisors.

Training on privacy and confidentiality in research data management: With researchers having identified privacy and confidentiality issues associated with data as one of the challenges faced, this study recommends provision of training to researchers which should include instruction on ethical principles and research data management best practice, such as obtaining informed consent from participants, anonymizing data, and implementing secure data storage and transfer procedures. The library could partner with the National Commission for Science, Technology & Innovation (NACOSTI) in providing the training. Having these skills will allow researchers to navigate the challenges and adhere to privacy and confidentiality concerns in their research practices.

Identifying centralized repositories for research data: With researchers having a challenge in locating datasets, the study recommends that centralized repositories for research data must be identified, as well as metadata standards that allow the discovery and reuse of research data.

Adopting standardized data preservation strategies: In view of challenges related to data preservation that researchers face, the study recommends the need to adopt standardized data preservation strategies that ensure research data's long-term accessibility and usage.

Developing data storage systems for various research data types: Furthermore, to overcome the challenge of storing data, the study recommends the development and implementation of data storage systems that would be suitable for storage of various types of research data. This can be accomplished through collaboration among researchers, IT specialists, and data curators per institution. As an alternative it may be wise to collaborate across institutions and share the costs and resources required to maintain shared infrastructure.

8.6.2   Infrastructure that supports data literacy initiatives for researchers

Support for libraries: With the majority of the respondents having indicated that the library has a vital role to play in promoting data literacy, the study recommends that universities should offer the necessary support to libraries to enhance their capacity to provide high-quality resources, infrastructure, and expertise to researchers. The support should help libraries to

foster more transparent and reproducible research practices, improve research quality, and contribute to knowledge advancement in various fields.

Improve research data services: In view of inadequate research data services provided by the library, according to the findings, the study recommends the need for libraries to focus on improving their research data services in order to better assist researchers' data management needs. This involves assisting in the development of data management plans, establishing and managing institutional data repositories, creating metadata for datasets, and facilitating digital data archiving and preservation. For future scholars to be able to easily access and preserve research data, libraries should prioritize developing effective and long-term data storage infrastructure. Libraries should help in promoting transparent and reproducible research processes by strengthening research data services and infrastructure, while also assisting researchers in properly managing and protecting their data.

Considering that researchers identified lack of adequate data literacy-related support services from the library throughout the research process in research data management, data publishing and sharing, and access to data management tools, this study calls on the improvement of these services. Libraries in collaboration with the research office or department, should invest in research-related technical infrastructure and forge collaborations with research institutes to establish a platform for disseminating training materials and resources. Incorporating data literacy training into the research process will improve data management practices, increase access to research data, and foster researcher collaboration. Libraries can use digital resources to produce interactive training modules and take a proactive approach to marketing their services to researchers through targeted marketing efforts.

Enhance data services: In view of the challenges encountered by librarians while providing research data-related services to researchers the study makes the following recommendations; First, libraries should develop comprehensive data literacy training programs that cover basic ICT skills and the use of data analysis tools to bridge skill gaps among researchers. Secondly, there is need for libraries to enhance communication and collaboration with researchers to better understand their data-related needs and explore alternative solutions. Thirdly, institutions should invest in professional development to enhance librarians' data literacy skills. Finally, institutions should strengthen information infrastructure by developing contingency plans for internet downtime and advocating for reliable connectivity. These actions will improve the

delivery of data-related services, support research projects, and promote data literacy within the academic community.

Provide proper resources and support to data curation experts: Based on the findings, the study acknowledged the key role played by both university librarians and research/reference librarians in the universities. However, to enhance their role in serving data literacy needs of researchers, the study recommends that universities should provide proper resources and support to these experts. This way, they will be able to carry out their obligations more successfully, such as improving the university's strategic goals, advancing the research agenda, and increasing data literacy among students and faculty. Furthermore, universities should think about implementing research skills training into their curricula. Collaboration among librarians, faculty, and administrators should be encouraged in order to promote a holistic approach to data literacy and research support services within the university community. Universities can enhance the academic experience and research success of their students and researchers by valuing and empowering librarians.

Establishing research offices: According to the findings, some respondents mentioned the lack organisational data infrastructure in their institutions, emphasizing the necessity for universities to acknowledge and emphasize the importance of data management. Based on the findings it is recommended that universities take proactive steps to build and strengthen their organizational structures and setups to support data management and data literacy activities. This involves establishing research offices or directorates devoted to providing assistance with data-related tasks. Furthermore, universities should engage in improving data management knowledge and education among all stakeholders, including senior management. This can be accomplished through training programs, conferences, and other public awareness efforts emphasizing the value of data literacy in the academic setting.

Allocate adequate funds: The universities should also allocate adequate funds for training costs, infrastructure expenses, and facility and technology expenses related to data literacy programs. Collaboration across essential university departments, such as the research department, ICT department, and institutional research office, is critical for data literacy programs to succeed. Faculty members who supervise postgraduate students should be actively involved in the development and implementation of data literacy programs as stakeholders. By developing supporting organizational structures, raising awareness, and fostering cooperation, universities can efficiently manage and use research data.

Data literacy training and professional development programs for librarians: As an addition to developing and filling skills, competencies and knowledge gap in data literacy among librarians, the study recommends, that libraries and parent institutions prioritize data literacy training and professional development programs for librarians. Collaborating with external experts or universities can provide tailored training opportunities to enhance librarians' data literacy skills. Additionally, a focus on areas such as data curation, technical data management skills, and legal and copyright frameworks can further strengthen librarians' ability to support researchers effectively. Continuous support and mentorship programs should be implemented to ensure librarians stay updated with evolving data management practices and technologies. By investing in the development of librarians' data literacy competencies, libraries can position themselves as valuable resources in the era of data-intensive research.

Development and implementation of comprehensive data literacy-related policies: Even though the findings indicated that there were some policies available in support for data literacy, the study found out the absence of some key policies that would support the implementation of data literacy. The study therefore recommends that universities should prioritize the development and implementation of comprehensive policies that promote research data management and data literacy. These policies should address a variety of issues, such as research data management, ICT infrastructure, staff development, and digital preservation. Institutions should engage in developing policies that provide clear guidelines for researchers and other stakeholders regarding data handling, storage, access, sharing, and reuse. These policies should be consistent with global trends toward open science and open research data, with the goal of increasing the availability and accessibility of research data. Furthermore, universities should prioritize the development of digital preservation policies to ensure the long-term usability of research data. To enhance the effectiveness of these policies, they should be assessed, and revised on a regular basis to accommodate evolving needs and challenges in the research data landscape. Collaboration among various stakeholders, including researchers, librarians, IT departments, and senior management, is critical for the formulation, implementation, and enforcement of these policies. The establishment of robust and focused data-related policies will improve research integrity, visibility, and collaboration by supporting good research data management and data literacy practices.

8.6.3 Technical infrastructural readiness for the implementation of data literacy training

According to the findings of the study (see Section 7.5), there is glairing gap in terms of technical infrastructure intended to support data literacy. The study recommends that there is need to improve the accessibility and availability of ICT infrastructure and resources, such as data analysis software such as SPSS and NVIVO, as well as reliable internet connectivity. Furthermore, universities need to prioritize the development of specific data management tools and platforms to support efficient data collection, manipulation, storage, and sharing. These enhancements will help researchers analyse and interpret data more effectively, promote seamless collaboration, and boost data literacy skills among library users. To provide the required technical infrastructure, the library should work in collaboration with the ICT team as well as the university administration.

8.6.4 Towards developing a data literacy framework

The ultimate aim of this study was to develop a standardized data literacy framework to cultivate a community of data literate researchers at Kenyan universities. Findings from the study revealed existing practices in libraries aimed towards providing services to researchers. However, respondents indicated that there were needed attention as far as data literacy was concerned.

Developing a data literacy training program framework for researchers in Kenyan universities is critical to equipping them with the required skills to manage their data efficiently, engage with others, ensure data quality, and make decisions based on evidence. This framework is aimed at addressing critical concepts in research data management, identify stakeholders as well as key skills that researchers will need to attain.

The study, first of all identified the need to consider other literacies (information literacy, digital literacy and statistical literacy) in the wake of implementing data literacy. With majority of the respondents having indicated their lack of knowledge in the three literacies, this study recommends that there is a need for comprehensive training in information literacy, digital literacy, and statistical literacy for researchers even as they get trained in data literacy. Considering these literacies in the data literacy training of researchers can help researchers effectively navigate and analyse data, communicate their findings, and enhance research productivity. By emphasizing the development of these essential skills, researchers will be better equipped to address the challenges they encounter in their research endeavours.

Considering what the respondents indicated as areas that should be prioritised in the development of a data literacy training program, the study recommends the need for stakeholders to be alert to the needs of researchers. Stakeholders should consider developing a curriculum that covers each of these key areas in order to establish a comprehensive data literacy framework that captures these essential themes. The curriculum should also be developed to accommodate learners with varying degrees of data literacy proficiency, ranging from novice to advanced users. It should also include digital literacy skills and data processing techniques that will assist researchers work efficiently with data. Finally, the framework must stress ethical data gathering, sharing, and management procedures, as well as data privacy and security issues.

With the library having largely been identified by respondents as having the capability of spearheading the implementation of data literacy in universities, the study, therefore recommends that libraries be enabled and supported to pioneer data literacy programs within universities. Researchers, university librarians, and research/reference librarians identified libraries as having the resources, infrastructure, expertise, and active involvement required to foster data literacy among researchers. Universities should provide adequate resources, funding, and professional development opportunities for librarians to improve their knowledge and abilities in data administration, data analysis, and other data-related fields in order to promote the role of libraries in data literacy. Collaboration should be promoted between libraries and other major university stakeholders, such as research departments and faculty members, to create a holistic and coordinated approach to data literacy programs. Furthermore, libraries can use their knowledge of information literacy to create integrated data literacy programs that target the specific needs of researchers. This may entail working with faculty to incorporate data-related abilities into current courses and partnering with industry stakeholders to establish the data literacy competencies needed in the workforce. Recognizing and supporting libraries' critical role in developing data literacy, universities may improve their scholars' research capacities and outcomes, encouraging a culture of data-driven inquiry and innovation.

In view of respondents' feedback on various aspects of initially proposed data literacy framework as presented in Section 3.4, (see Figure 3.1), the study makes the following recommendations:

i. Since the respondents indicated that the proposed data literacy framework was necessary for implementing, the study took note of their suggestion on areas that need improvement. And therefore, the study recommends that the framework be improved in specific areas to improve its effectiveness to capture notable areas as suggested by respondents, this includes identifying and involving key stakeholders, such as faculty members and other relevant university departments, who play an important role in data literacy activities.

ii. In view of the positive input from respondents, it is recommended that the framework be further developed and refined for effective implementation. In its development, the framework should maintain its well-structured and broad nature, including all the fundamentals of data literacy. Furthermore, it is critical to underline the framework's relevance to the needs of researchers, particularly faculty members and students.

In conclusion, this section provides a comprehensive set of recommendations based on the major findings of the study regarding data literacy needs, existing organizational infrastructure, technical infrastructure readiness, and the development of a data literacy framework in Kenyan universities. The recommendations are presented in a systematic order, addressing various aspects related to data literacy.

Regarding data literacy needs of faculty members and postgraduate students, the study recommends that libraries should bridge the gap between the services offered and the data literacy needs of researchers. This can be achieved by providing support in research data management, increasing awareness of data management plans, offering training on metadata creation, improving data storage infrastructure, enhancing competences in research data management, improving communication about data-related services, and establishing comprehensive data-related support services.

The study emphasizes the role of libraries in promoting data literacy and recommends that universities provide necessary support to libraries to enhance their capacity in providing resources, infrastructure, and expertise to researchers. Libraries should focus on improving research data services, supporting data management needs, developing data literacy training programs, enhancing communication and collaboration with researchers, investing in professional development for librarians, and strengthening information infrastructure.

In terms of technical infrastructural readiness, the study suggests improving accessibility and availability of ICT infrastructure, data analysis software, and reliable internet connectivity.

Universities should prioritize the development of data management tools and platforms to support efficient data collection, manipulation, storage, and sharing.

The study emphasizes the need to develop a standardized data literacy framework to cultivate a community of data literate researchers. The framework should consider other literacies such as information literacy, digital literacy, and statistical literacy. Stakeholders should develop a comprehensive curriculum that covers key areas, accommodate learners with varying degrees of data literacy proficiency, and emphasize ethical data practices. Libraries should be enabled and supported to pioneer data literacy programs within universities, collaborating with other stakeholders and leveraging their expertise in information literacy.

Finally, the study recommends improving the initially proposed data literacy framework based on feedback from respondents. Key areas for improvement include involving key stakeholders and refining the framework to align with the needs of researchers.

Overall, implementing these recommendations can contribute to enhancing data literacy among researchers, improving research data management practices, promoting transparency and reproducibility in research, and advancing knowledge in various fields at Kenyan universities.

8.7   Framework discussion

After carefully integrating research, literature, and evaluating the study's discussion findings, the primary recommendation is the adoption of a data literacy framework for the implementation of data literacy in Kenyan universities. The initial framework (see Figure 3.1) presented in Section 3.4 received favourable reception among the respondents (see Section 7.2). Changes were incorporated to enhance the preliminary framework and to demonstrate key insights from the study's findings, serving as a comprehensive guide for promoting data literacy among researchers in universities. The framework's content is primarily derived from the outcomes of the study, as outlined in the accompanying Figure 8.1, which serves to demonstrate the core process. In addition, an output model was devised to demonstrate the core process but also all of the identified outputs associated with the framework (see Figure 8.2).
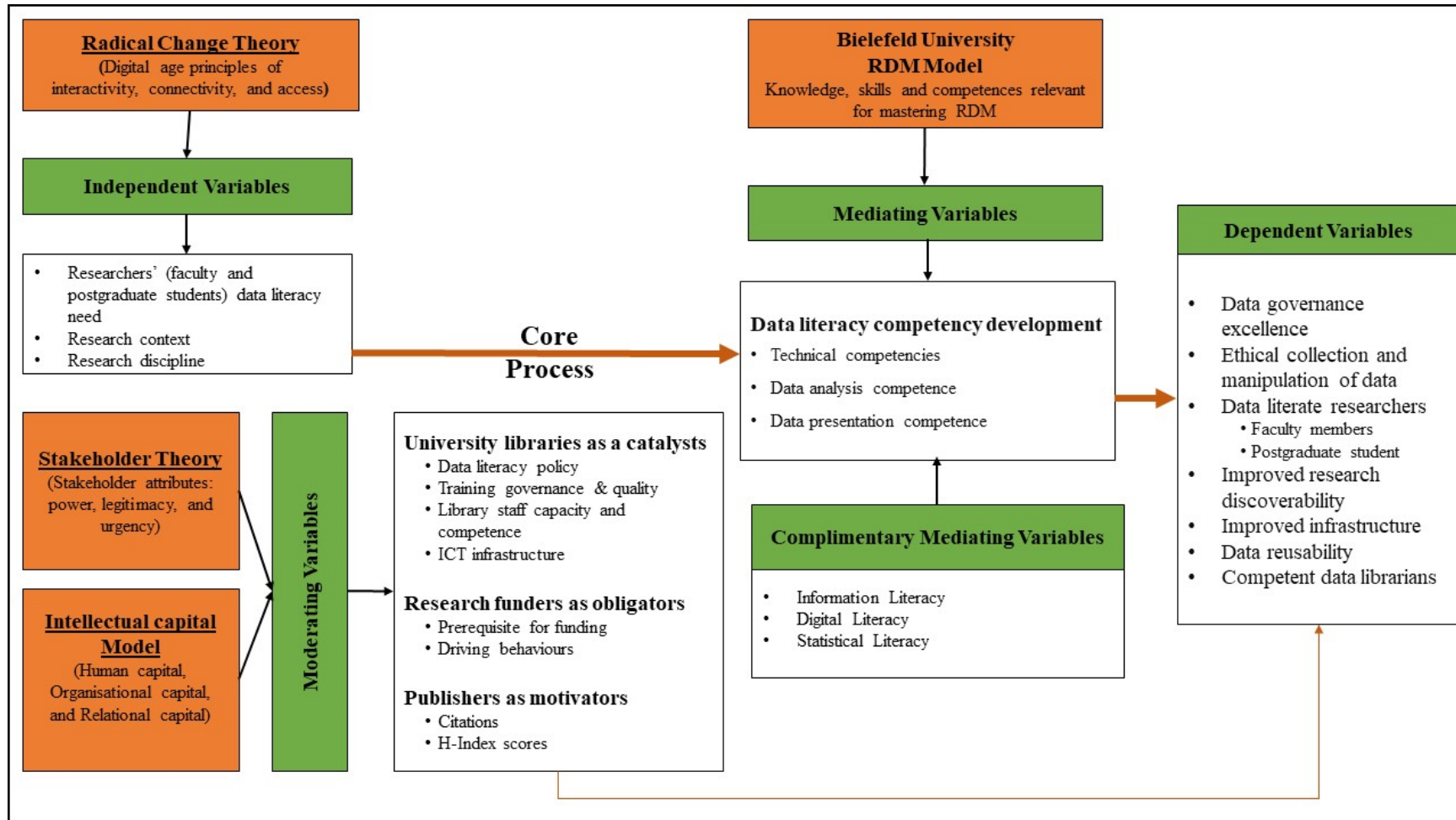
Figure 8- 1: Data literacy framework for Kenyan university libraries – process model

Source: Revised version of Figure 3.1

285
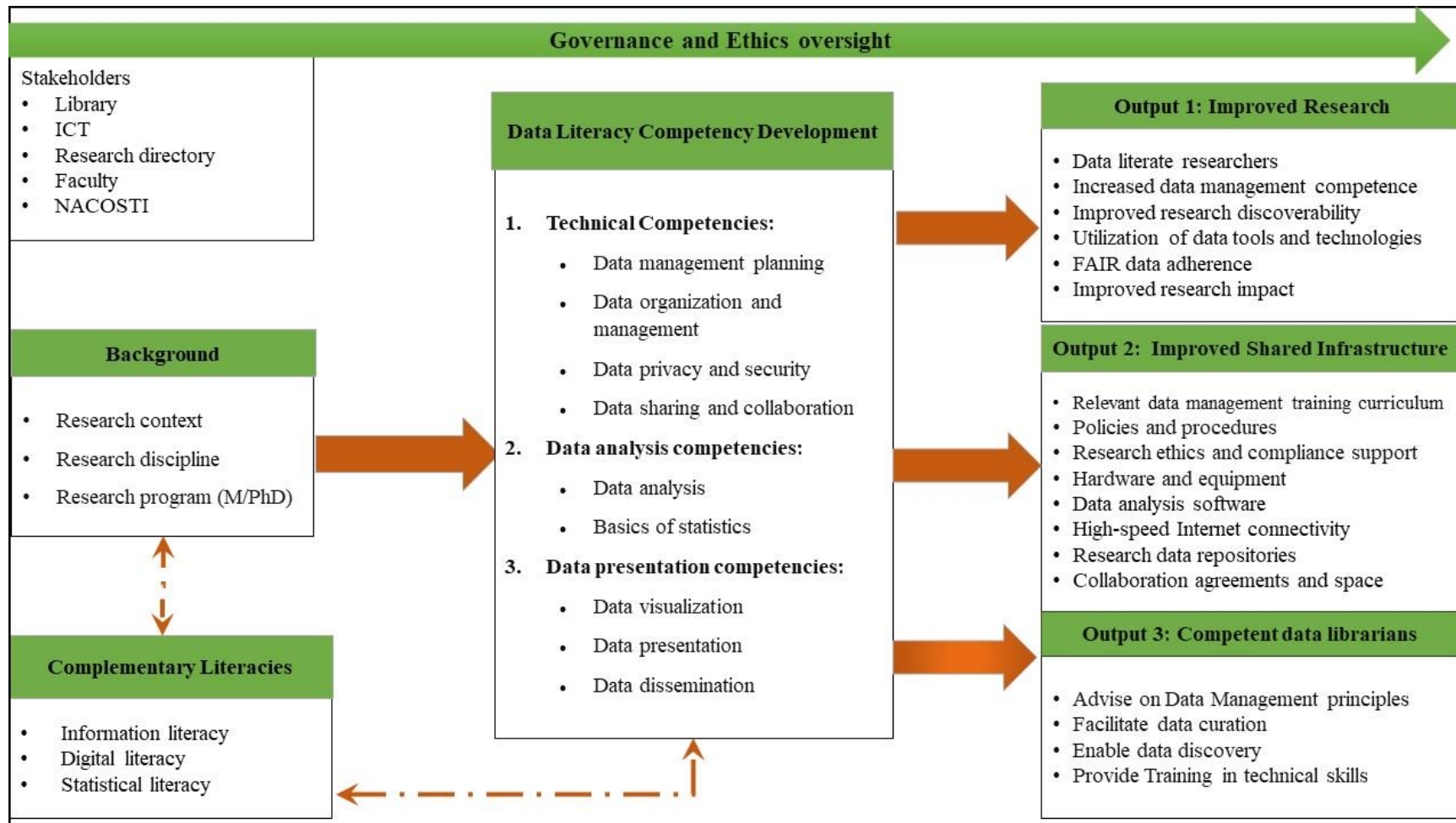
Figure 8- 2: Data literacy framework for Kenyan university libraries – output model

8.7.1 Stakeholders

The role of each stakeholder in the development and implementation of the data literacy program in universities is key.

1. Library: The library was identified by respondents as one that would play a central role in providing data literacy services. It is perceived to have the resources, technology, and staff to meet the data needs of researchers. In this case, would take the lead role of coordinating the whole process implementation and execution of data literacy. The library staff, particularly data librarians, would be involved in designing the curriculum, organizing training sessions, providing guidance and support to researchers and students, and managing data-related resources and infrastructure.

2. Information and Communication Technology (ICT) Department: The ICT department is crucial in providing the necessary technical support and infrastructure for data literacy services. The department plays a role in ensuring reliable internet connectivity, access to data analysis software, and the availability of hardware and software resources required for data literacy training and activities. Collaboration between the library and the ICT department is essential for the effective implementation.

3. Research Directorate: The research directorate or research office within the university is responsible for overseeing research activities and promoting research excellence. In the context of the data literacy program, the research directorate can collaborate with the library to incorporate data literacy training into the research process. They can provide support in terms of funding, research policy development, and facilitating collaborations between researchers and librarians.

4. Faculty: Faculty members are key stakeholders in the development and implementation of the data literacy program. They can contribute to the program by integrating data literacy concepts and skills into their respective courses and research projects. Faculty members can also play an active role in promoting the importance of data literacy among students and supporting their engagement in data-related activities.

5. National Commission for Science, Technology and Innovation (NACOSTI): NACOSTI is a national body in Kenya that is responsible for regulating and promoting science, technology, and innovation. In the context of the data literacy program, NACOSTI can provide guidance and support in terms of policy development, ethical considerations, and compliance with data management standards. They can collaborate

287

with universities and libraries to ensure that the program aligns with national guidelines and best practices.

These stakeholders play interconnected roles in the development and implementation of the data literacy program, with the library taking the lead in providing the necessary resources, infrastructure, and expertise, while collaborating with the ICT department, research directorate, faculty, and external bodies like NACOSTI to ensure standards are attained and maintained.

The yellow arrow at the top, running all the way from the stakeholders moving towards the end of the figure is an indication of the role of the stakeholders throughout. Their role does not only end once data literacy program has been developed and implemented, but they have to keep on monitoring it to ensure set standards are met.

8.7.2 Background

There are three items mentioned in forming the background to justify the development and implementation of data literacy.

1. Research context: The research context refers to the specific environment, challenges, and requirements within which researchers conduct their studies. Each research context may have unique data-related challenges, such as data collection, data management, data analysis, or data sharing. The background should justify the development of the data literacy framework by addressing the specific data-related needs and challenges present in the research context. The framework should be tailored to address the data literacy requirements within this specific context, taking into account the institutional culture, research priorities, and resources available.

2. Research discipline: Research disciplines encompass various fields of study, each with its own data characteristics, methodologies, and analytical techniques. The background should acknowledge the diversity of research disciplines and highlight the importance of discipline-specific data literacy skills. Different disciplines may require specialized knowledge and approaches to data management, analysis, and interpretation. The development and implementation of the data literacy framework should consider the specific needs and requirements of researchers in various disciplines, ensuring that the framework caters to the unique data literacy demands of different research domains.

   Research program (M/PhD): In addition to recognizing the importance of the research context and research discipline, it's essential to acknowledge the specific considerations

288

brought about by different research programs, particularly those pursued at the Master's (M) and Doctoral (PhD) levels. These advanced research programs delve deeply into specialized areas of study, demanding a heightened level of data literacy skills due to the rigorous nature of their inquiry.

Master's and PhD research programs are characterized by their emphasis on original research, critical analysis, and the production of new knowledge within a particular field. As researchers progress through these programs, they engage in complex data-driven projects, ranging from in-depth explorations to ground-breaking contributions that can significantly influence their respective domains.

Considering the unique challenges posed by these research programs, it becomes evident that a data literacy framework must cater to their specific demands. Data literacy is not only about understanding basic concepts but also encompasses the ability to navigate intricate datasets, apply advanced analytical methods, interpret nuanced findings, and effectively communicate complex results to both specialized and broader audiences.

Furthermore, researchers pursuing M and PhD degrees often work independently, requiring a heightened level of self-sufficiency in managing, analysing, and interpreting data. The data literacy framework should, therefore, empower them with the skills and knowledge necessary to make informed decisions throughout the research process, from experimental design to final publication.

By acknowledging the distinctive characteristics of research programs at the M and PhD levels, the development and implementation of the data literacy framework can foster a learning environment that equips researchers with the capabilities to excel in their endeavours. This entails providing resources, training, and support that are attuned to the intricate demands of their research journey.

By including research context, research discipline and research program (M/PhD) in the background to justify the development and implementation of data literacy, the study acknowledges the significance of addressing the specific needs of researchers within their research context and disciplinary frameworks. This approach ensures that the data literacy framework is relevant, practical, and tailored to the challenges and requirements faced by researchers in their specific fields and institutional contexts. Stakeholders should therefore be cognizant of these considerations.

8.7.3 Complementary literacies

In the development and implementation of a data literacy program, three literacies (information literacy, digital literacy, and statistical literacy) have been identified as being important.

1. Information Literacy: Information literacy is the ability to identify, locate, evaluate, and effectively use information from various sources. In the context of data literacy, information literacy is essential for researchers to access and retrieve relevant data sources and research literature. It enables researchers to critically evaluate the quality and reliability of data, understand data documentation and metadata, and effectively use information to inform their research questions and data analysis. Information literacy skills complement data literacy by providing researchers with the necessary foundation to navigate the vast amount of information and data available.

2. Digital Literacy: Digital literacy refers to the skills and knowledge required to effectively use digital technologies and tools. In the context of data literacy, digital literacy plays a crucial role in accessing, managing, and analysing digital data. Digital literacy skills enable researchers to work with data analysis software, database systems, data visualization tools, and other digital platforms for data management and analysis. Proficiency in digital literacy ensures researchers can effectively manipulate and interpret digital data, facilitating their understanding and utilization of research data.

3. Statistical Literacy: Statistical literacy involves understanding basic statistical concepts, reasoning, and interpreting statistical information. In the context of data literacy, statistical literacy is vital for researchers to analyse and interpret research data accurately. It includes knowledge of statistical techniques, data distributions, data sampling, hypothesis testing, and understanding statistical measures such as means, standard deviations, and correlations. Statistical literacy enables researchers to make informed decisions based on statistical evidence, effectively communicate their findings, and ensure the validity and reliability of their research conclusions.

The link between these three literacies and the successful implementation of data literacy is evident. Information literacy ensures researchers can locate and evaluate data sources, digital literacy enables researchers to work with digital tools and technologies for data management, and statistical literacy equips researchers with the skills to analyse and interpret research data accurately. Together, these literacies form a solid foundation for researchers to effectively

290

navigate the data landscape, critically evaluate data, analyse it using appropriate statistical techniques, and derive meaningful insights.

In the development and implementation of the data literacy framework, it is crucial to consider these three literacies and integrate them into the training and support provided to researchers. By addressing information literacy, digital literacy, and statistical literacy alongside data literacy, the framework can equip researchers with a comprehensive set of skills and knowledge needed to successfully manage, analyse, and interpret research data in their respective disciplines.

In the process of implementing data literacy initiatives, it is also essential for stakeholders to consider whether researchers possess complementary skills, such as Information literacy, Digital literacy, and Statistical literacy. The dotted arrows from the "Complementary Skills" box to the "Background" box and also from the "Complementary Skills" box to the "Data Literacy Competency Development" box indicate that researchers who already possess these complementary skills can seamlessly participate in the program.

On the other hand, researchers who currently lack these complementary skills could still have an opportunity to join the data literacy program. Stakeholders should consider introducing and incorporating these complementary skills into the program's implementation. This is crucial because these skills play a critical role in enhancing the researchers' ability to attain and benefit from data literacy skills.

In summary, the dotted lines in the developed outcomes framework represent the pathways for researchers with and without complementary skills to engage in the data literacy initiative. It highlights the importance of considering and incorporating complementary skills, such as Information literacy, Digital literacy, and Statistical literacy, to ensure a comprehensive and effective approach to developing data literate researchers.

8.7.4 Data literacy competency development

In view of the reviewed literature and findings of the study, data literacy competency development and implementation by universities should consist of the following areas:

1. Technical competencies

This competency involves the technical skills required to work with data and manage data-related tasks. Data organization and management, data privacy and security, and data sharing

291

and collaboration are all technical aspects of data literacy, as they involve handling data, ensuring its integrity, and facilitating collaboration in data-related activities.

Data organization and management: This is about the ability to effectively handle and manage research data throughout its lifecycle. Data organization includes practices such as data cleaning, formatting, and structuring to ensure data integrity and accessibility. Researchers with strong data organization skills can efficiently locate and retrieve data, saving time and effort during the research process. Proper data management enhances data quality, reproducibility, and the ability to draw meaningful insights from research data.

Data privacy and security: Data privacy and security are critical aspects of data literacy. Researchers must be aware of ethical considerations and legal requirements related to handling sensitive and confidential data. Competency in data privacy ensures that researchers protect the privacy rights of individuals whose data is collected and adhere to data protection regulations. Additionally, understanding data security measures prevents data breaches and unauthorized access, safeguarding the integrity and trustworthiness of research data.

Data sharing and collaboration: Data sharing and collaboration are essential components of open science and data-driven research. Data-literate researchers should know how to share research data responsibly and ethically. Competency in data sharing enables researchers to contribute to the broader scientific community, promote transparency, and encourage collaborations and interdisciplinary research. It fosters an environment of knowledge exchange and helps advance research in various fields.

2.   Data analysis competencies

This competency focuses on the skills needed to analyse and interpret research data accurately. Data analysis and basic statistics fall under this category as they involve using statistical methods to draw insights and conclusions from data.

Data analysis: Analytical competencies are fundamental for researchers to derive meaningful insights from research data. Data-literate researchers possess the skills to apply appropriate statistical techniques, data modelling, and data mining methods to analyse data effectively. Competency in data analysis empowers researchers to draw valid conclusions, identify patterns, and make evidence-based decisions. It ensures the reliability and accuracy of research findings and enhances the overall quality of research outputs.

Basics statistics: A solid grasp of basic statistics is essential within the realm of data analysis competencies. This understanding serves as the foundation for researchers to extract valuable insights from their data, enabling them to navigate datasets, identify trends, and formulate initial conclusions. Proficiency in these fundamental statistical concepts empowers researchers to perform meaningful data analysis, uncovering patterns and relationships within their datasets. This analytical prowess contributes to accurate insights and well-informed decisions, ultimately enhancing the credibility and impact of research outcomes. When integrated into the broader data literacy framework, proficiency in basic statistics complements other competencies, creating a comprehensive skill set that equips researchers to excel in data-driven endeavours. By cultivating a robust statistical foundation, researchers enhance their ability to navigate complex data analysis processes, ensuring the reliability and robustness of their research outputs.

3.   Data presentation competencies

This competency relates to the skills needed to effectively communicate research findings and insights derived from data. Data visualization, presentation and dissemination are essential in this context, as they enable researchers to present their data in a clear and visually appealing manner. Data ethics and governance are also relevant to communication competencies as they pertain to the responsible and ethical use of data in research, which is essential for communicating research outcomes with integrity.

Data visualization and presentation: Effective data visualization and presentation skills are crucial for researchers to communicate their research findings clearly and compellingly. Data-literate researchers can use visual elements such as charts, graphs, and infographics to represent complex data in a more understandable and impactful manner. Competency in data visualization enhances the accessibility of research findings, making them more engaging to diverse audiences, including fellow researchers, policymakers, and the general public.

By developing and implementing these competencies, universities can cultivate a data-literate research community. Data-literate researchers possess the skills and knowledge to effectively manage, analyse, interpret, and communicate research data, thereby contributing to research excellence and the advancement of knowledge. These competencies support data-driven research practices, enhance research integrity and reproducibility, and promote collaboration and data sharing within the academic community. Moreover, data literacy empowers researchers to tackle complex research questions, address societal challenges, and make

significant contributions to their respective fields. As the importance of data-driven research continues to grow, fostering data literacy among researchers becomes a critical imperative for universities.

8.7.5 Output 1: Research

The first expected outcome or output of the well developed and implemented data literacy program in a university will be on or about research. There are seven different areas related to the research output:

1. **Development of relevant training curriculum**: The development of a relevant training curriculum would be a key aspect considering what the stakeholders will put in place as well as in view of the background or context. The development of relevant training curriculum will be based on the identification of specific data literacy competencies, skills, and knowledge required by researchers. The resulting training curriculum will serve as a practical and tailored resource, addressing the unique data literacy needs and challenges in the research context. This output will contribute to the enhancement of research practices and the promotion of data literacy among researchers.

2. **Data literate researchers**: The development and implementation of the data literacy framework aims to cultivate a community of data-literate researchers. As an output, researchers who participate in the data literacy training programs and initiatives will acquire the necessary skills, knowledge, and competencies to effectively manage, analyse, and interpret research data. They will become proficient in data management best practices, data analysis techniques, and data visualization methods. The output of data-literate researchers reflects the success of the data literacy program in empowering researchers with essential data-related skills.

3. **Increased data management competence**: Another research output is the increased data management competence among researchers. Through the data literacy training, researchers will gain proficiency in data collection, organization, documentation, cleaning, and preservation. They will understand the importance of data quality, data security, and ethical considerations in research data management. Increased data management competence ensures that researchers can handle research data effectively, leading to improved research integrity, reproducibility, and the ability to derive meaningful insights from their data.

4. **Improved research discoverability**: As a research output, improved research discoverability stems from the implementation of data literacy practices. Researchers who possess data literacy skills will be more adept at documenting and describing their research data, adhering to metadata standards, and utilizing appropriate data repositories. This will enhance the discoverability of their research data, making it easier for other researchers to locate, access, and potentially reuse the data for further research. Improved research discoverability promotes collaboration, transparency, and the advancement of knowledge within the research community.

5. **Utilization of data tools and technologies**: The successful implementation of the data literacy framework will result in researchers effectively utilizing data tools and technologies. Researchers will acquire the skills to work with data analysis software, data visualization tools, and database systems relevant to their research disciplines. They will harness these tools and technologies to analyse, interpret, and present research data more effectively, thereby improving the quality and impact of their research outputs. The utilization of data tools and technologies enhances research capabilities and supports data-driven decision-making.

6. **FAIR data adherence**: FAIR data principles emphasize the Findability, Accessibility, Interoperability, and Reusability of research data. As a research output, the implementation of the data literacy framework would foster adherence to these principles among researchers. They will understand the importance of data documentation, metadata standards, and data sharing practices that align with FAIR principles. Adhering to FAIR data practices improves the accessibility and usability of research data, promotes transparency, and facilitates collaboration and knowledge exchange within the research community.

7. **Improved research output and impact**: Ultimately, the research output and impact of researchers will be enhanced as a result of data literacy initiatives. By acquiring data literacy skills, researchers will effectively manage, analyse, and interpret their research data, leading to higher-quality research outputs. Improved data management practices, enhanced research discoverability, utilization of data tools, adherence to FAIR data principles, and increased data literacy all contribute to the improvement of research output and impact. This, in turn, benefits the researchers, their institutions, and the broader academic community.

These research outputs demonstrate the positive outcomes of the study's efforts to develop and implement a data literacy framework. They indicate the tangible benefits and improvements in research practices, data management, and research impact resulting from the cultivation of data literacy among researchers.

8.7.6 Output 2: Infrastructure

The development and implementation of data literacy in universities will have an infrastructural impact in various ways.

1. **Policies and governance**: The development and implementation of a data literacy framework will contribute to the establishment of policies and governance structures related to research data management. These policies will provide guidelines for researchers and other stakeholders regarding data handling, storage, access, sharing, and reuse. The output of policies and governance ensures a structured and regulated approach to data management, protecting the rights of researchers and enhancing the integrity and security of research data.

2. **Research ethics and compliance support**: The implementation of the data literacy framework will have an impact on research ethics and compliance support infrastructure. Researchers will receive training on ethical principles and best practices in research data management. Researchers will have access to support and resources to ensure compliance with ethical standards, such as obtaining informed consent, anonymizing data, and implementing secure data storage and transfer procedures. The output of research ethics and compliance support infrastructure will strengthen the ethical foundations of research, protect human subjects' rights, and ensure data privacy and confidentiality. Depending on the discipline, institutions will have in place mechanisms and committees in place to attend to different ethical requirements and compliance to specific research.

3. **Hardware and equipment**: Implementation of data literacy program will require the need for appropriate hardware and equipment to support data literacy initiatives. There will be need for allocation of resources for acquiring hardware and equipment necessary for data literacy training and activities. This will include computers, laptops, data storage devices, and other hardware required to effectively manage and analyse research data. The output of hardware and equipment infrastructure ensures that

researchers have the necessary tools to engage in data-related activities and utilize data analysis software effectively.

4. **Data analysis software:** The development and implementation of the data literacy framework will impact the availability and utilization of data analysis software. For the purpose of data analysis and manipulation, universities will have to make for the provision of data analysis software that is essential for researchers to conduct data analysis and statistical computations. The output of data analysis software infrastructure ensures that researchers have access to licensed or open-source software tools relevant to their research disciplines. This enables them to perform data analysis tasks, draw meaningful insights, and produce reliable research outcomes.

5. **High-speed internet connectivity**: In view of the magnitude of data that researchers will have to deal with while also envisioning, data reuse and collaborations among researchers, there will be need for high-speed internet connectivity in supporting data literacy initiatives. This will facilitate access to online resources, data repositories, collaborative platforms, and data analysis tools. The output of high-speed internet connectivity infrastructure enables researchers to seamlessly access and utilize digital resources, engage in data-intensive activities, and stay connected to global research networks.

6. **Research data repositories**: The availability of research data repositories as part of the infrastructure supporting data literacy cannot be overemphasised. These repositories provide a secure and accessible platform for researchers to store, share, and preserve their research data. The output of research data repositories infrastructure will ensure the availability, accessibility, and long-term usability of research data, promoting data sharing, collaboration, and reproducibility.

7. **Collaboration agreements and space:** The implementation of the data literacy framework will impact collaboration agreements and the provision of suitable spaces. Fostering collaborations among stakeholders, such as libraries, research departments, faculty members, and external organizations will be necessary. Collaboration agreements will facilitate the sharing of expertise, resources, and training materials related to data literacy. Implementation of data literacy will advocate for dedicated spaces, such as data labs or research support centres, where researchers can access data-related services, receive training, and engage in collaborative activities. The output of

© University of Pretoria

collaboration agreements and space infrastructure encourages interdisciplinary collaboration, knowledge exchange, and effective utilization of data literacy resources.

These research outputs on infrastructure demonstrate the impact of the data literacy framework on various aspects of research infrastructure. They contribute to the development of a supportive and robust infrastructure that facilitates data literacy initiatives and enhances research capabilities within institutions.

8.7.7 Output 3: Librarian

The development and implementation of data literacy program will be setting in motion the development of skills and competences among librarian considered to be key players in the success of data literacy. This is because of their interaction with researchers as the findings of the study suggest.

1. **Data management**: The development and implementation of the data literacy framework will contribute to the development of skilled and competent data librarians in data management. Through training and professional development programs, librarians will acquire expertise in data management practices, including data collection, organization, documentation, cleaning, preservation, and sharing. They will understand the principles and best practices of data management throughout the research data lifecycle. The output of data management skills and competencies in data librarians enables them to effectively support researchers in managing their data, ensuring its quality, integrity, and accessibility.

2. **Data curation**: Data curation is the process of organizing, describing, and preserving research data for long-term usability and accessibility. In the context of the study, the development of a data literacy framework will lead to the development of skilled data librarians in data curation. Librarians will gain knowledge and skills in metadata creation, data documentation, data formatting, and data preservation techniques. They will understand the importance of data provenance, metadata standards, and quality control. The output of data curation skills and competencies in data librarians ensures they can effectively curate and maintain research data, enabling its discoverability, usability, and long-term preservation.

3. **Data discovery**: Data discovery refers to the process of locating and accessing relevant research data. Skilled data librarians will possess expertise in data discovery techniques and tools. They will be proficient in data search strategies, database querying, and

298

metadata-driven search methods. Librarians will understand the importance of data documentation, metadata standards, and indexing practices for facilitating data discovery. The output of data discovery skills and competencies in data librarians enables them to assist researchers in finding and accessing the most relevant research data for their projects, saving time and effort in data exploration.

4. **Technical skills**: Skilled and competent data librarians will possess technical skills required to work with data tools, software, and platforms. They will acquire proficiency in data analysis software, database systems, data visualization tools, and other technical resources relevant to data management and analysis. Librarians will understand data formats, data manipulation techniques, and technical considerations in data handling. The output of technical skills in data librarians ensures they can provide technical support to researchers in utilizing data tools, troubleshooting technical issues, and optimizing data management workflows.

By developing and implementing the data literacy framework, librarians will acquire these skills and competencies through training, professional development programs, and hands-on experience. The output of skilled and competent data librarians in data management, data curation, data discovery, and technical skills supports their role as valuable resources for researchers. They can effectively assist researchers in navigating the complexities of research data, provide guidance on best practices, and contribute to the successful management and utilization of data in research projects. Skilled data librarians play a crucial role in fostering data literacy among researchers, enhancing research outcomes, and promoting data-driven inquiry.

In conclusion, the development and implementation of a data literacy framework in universities will have a significant and far-reaching impact on various aspects of the research ecosystem. The comprehensive framework, guided by the study's findings, addresses key areas of data literacy competency development, stakeholder engagement, and infrastructure enhancement. The involvement of key stakeholders, such as the library, information and communication technology (ICT) department, research directorate, faculty, and the National Commission for Science, Technology, and Innovation (NACOSTI), ensures a collaborative and coordinated effort in promoting data literacy. The role of the library, in particular, is pivotal, as it leads the coordination of data literacy services, leverages its resources and technology, and supports researchers in their data-related endeavours. The incorporation of complementary literacies,

such as information literacy, digital literacy, and statistical literacy, further strengthens the data literacy program. These literacies provide researchers with the essential skills to navigate information resources, work effectively with digital tools, and conduct accurate data analysis, all of which are integral to successful data literacy practices.

The research outputs of the data literacy framework encompass the development of relevant training curriculum, data-literate researchers, increased data management competence, improved research discoverability, utilization of data tools and technologies, adherence to FAIR data principles, and improved research output and impact. These outputs highlight the positive outcomes of the program, demonstrating how researchers' skills, research practices, and collaboration are enhanced through data literacy initiatives. The impact on infrastructure emphasizes the importance of establishing policies and governance, supporting research ethics and compliance, providing essential hardware and equipment, ensuring access to data analysis software and high-speed internet connectivity, establishing research data repositories, and fostering collaboration agreements and dedicated spaces. These infrastructure enhancements create an enabling environment for researchers to engage in data literacy activities effectively and contribute to research excellence. Finally, the development of skilled and competent data librarians serves as a key output of the data literacy program. Data librarians play a critical role in supporting researchers' data management needs, promoting data curation and discovery, and offering technical assistance in data tools and technologies.

Overall, the data literacy framework will transform the research landscape in Kenyan universities, equipping researchers with essential skills, fostering data-driven decision-making, and ultimately contributing to the advancement of knowledge and research impact. Through collaboration among stakeholders and the cultivation of a data-literate community, the data literacy program will position universities at the forefront of data-enabled research and innovation.

8.8   Recommendations for further research

The study also makes some further recommendations for further study based on the findings:

1. **Assessing the impact of enhanced data literacy support**: Research is needed to assess the impact of implementing data literacy support in selected private university libraries in Kenya. The objectives of this study are to assess whether the interventions were effective in improving data literacy, improving data management practices, and improving the quality and impact of research outcomes.

2. **Investigating barriers to implementing data management plans (DMPs)**: Investigate barriers to DMP uptake and implementation among Kenyan university researchers. This research can look deeper into issues including institutional support, understanding, and skill in data management procedures, as well as offer solutions to overcome these barriers and promote the broad usage of DMPs.

3. **Evaluating the effectiveness of metadata training programs**: Conduct a study to determine the efficacy of metadata training programs offered by libraries in increasing researchers' knowledge and understanding of metadata development. The purpose of this study is to examine the impact of metadata training on data management practices, data sharing, and data reuse, as well as the best methods for teaching and training metadata.

4. **Qualitative research, to clarify some of the quantitative research results, will add value:** An investigation into the current practices would explain how researchers manage when the library does not provide enough support.

These areas of further study can contribute to the existing body of knowledge on data literacy, research data management, and library support in the Kenyan context, and provide insights for future improvements and interventions in supporting researchers and promoting effective data management practices.

8.9   Limitations of the study

Although this study was successful, it did encounter certain limitations throughout its process. Initially, the researcher planned to incorporate six private universities into the study population. However, one university's management declined to grant permission for data collection, even after receiving all the necessary supporting documents.

While efforts were made to analyze the current literature relevant to the study's focus on data literacy in the setting of selected private university libraries in Kenya, it is vital to recognize a challenge caused by a lack of the most recent relevant literature. Despite the application of strategies for searching and extensive study of academic resources, current research addressing specific nuances and developing trends in data literacy within the examined context were scarce. As a result, the reliance on older sources may have limited the depth of study of current advances and changing perspectives on the subject. The above limitation highlights the need for further studies to bridge the gap between known knowledge and current practices, thereby improving the robustness and currency of Kenyan data literacy training.

8.10 Summary

The major findings of the study shed light on crucial insights that can guide the development of effective data literacy initiatives. It revealed a gap between researchers' data literacy needs and the services provided by libraries. To bridge this gap, libraries should enhance their support for data creation, processing, analysis, and preservation services. Additionally, researchers expressed limited competence in various areas of research data management, highlighting the need for comprehensive training programs covering all stages of the research data life cycle.

The study identified varying levels of organizational infrastructure supporting data literacy initiatives within selected private university libraries. For successful data literacy implementation, universities should invest in resources and support for librarians to play a more active role in promoting data literacy and research support services. Furthermore, the development of policies addressing data management, ICT infrastructure, staff development, and digital preservation is essential.

In terms of technical infrastructure readiness, the study revealed that some areas were inadequate, emphasizing the need for improved accessibility and availability of ICT infrastructure, data analysis software, and reliable internet connectivity.

The study's success culminated in the development of a comprehensive data literacy framework encompassing essential elements such as data organization and management, data analysis and statistics, data visualization and presentation, data sharing and collaboration, data ethics and governance, and data privacy and security. This framework, once implemented, holds the potential to enhance researchers' data literacy skills and contribute to the advancement of open science and research quality in Kenyan universities.

Based on the findings, the study offers valuable recommendations to address the identified gaps and challenges. Libraries are urged to improve their data-related services, offer training on metadata creation and data management, and collaborate with stakeholders to provide technical resources. Universities should prioritize the development of policies supporting research data management and data literacy, while also investing in technical infrastructure and providing training opportunities for librarians. Strengthening data literacy among researchers will require collaboration among various stakeholders, and the adoption of policies and infrastructure in line with global trends towards open science and open research data.

In conclusion, this study has successfully explored the feasibility of offering standardized data literacy services at selected private university libraries in Kenya. Through a pragmatic research

paradigm and a mixed-method approach, the study addressed its objectives by thoroughly examining the data literacy landscape in the target universities.

This study presented significant insights into the data literacy needs and challenges faced by researchers in Kenyan universities. By implementing the recommended actions and adopting the developed data literacy framework, libraries and academic institutions in Kenya can better support researchers in managing research data, promoting open science practices, and enhancing the overall quality and impact of research outcomes. Ultimately, the successful implementation of standardized data literacy services in private university libraries can pave the way for increased research excellence, innovation, and positive societal impact in Kenya's academic community.

REFERENCES

Aaltonen, K. & Kujala, J. 2010. A project lifecycle perspective on stakeholder influence strategies in global projects. *Scandinavian Journal of Management*. 26(4): 381-397. DOI: 10.1016/j.scaman.2010.09.001.

Aborisade, O.P. 2013. Data collection and new technology. *International Journal of Emerging Technologies in Learning (IJET)*, 8(2): 48-52. DOI: https://doi.org/10.3991/ijet.v8i2.2157.

Acharjya, D.P. & Ahmed, K. 2016. A survey on big data analytics: challenges, open research issues and tools. *International Journal of Advanced Computer Science and Applications*, 7(2): 511-518.

ACRL Research Planning and Review Committee. 2015. Environmental Scan 2015. Available at: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/EnvironmentalScan15.pdf [20 October 2023].

Adamick, J., Reznik-Zellen, R. & Sheridan, M. 2012. Data management training for graduate students at a large research university. *Journal of eScience Librarianship*, 1(3): 180-188. DOI: 10.7191/jeslib.2012.1022.

Adika, F.O. & Kwanya, T. 2020. Research data management literacy amongst lecturers at Strathmore University, Kenya. *Library Management*. 41(6/7): 447-466. DOI: http://doi.org/10.1108/LM-03-2020-0043.

Akers, K.G. & Doty, J. 2013. Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation*. 8(2): 5-26. DOI: http://doi.org/10.2218/ijdc.v8i2.263.

Al Riyami, A.T. 2015. Main Approaches to Educational Research. *International Journal of Innovation and Research in Educational Sciences,* 2(5). Available at: https://www.researchgate.net/profile/Thuraya-Al-Riyami/publication/283071843_Main_Approaches_to_Educational_Research/links/5628a82d08ae518e347c5ee3/Main-Approaches-to-Educational-Research.pdf [4 November 2023].

Albers, M.J. 2017. Quantitative data analysis-In the graduate curriculum. *Journal of Technical Writing and Communication*, 47(2): 215-233. DOI: https://doi.org/10.1177/0047281617692.

Alexander, D.S. 2020. *Concurrent Triangulation Mixed Methods Research: Designing and Conducting a Childhood Obesity Study in a Rural Setting*. London SAGE Publications, Ltd. Available at: http://www.doi.org/10.4135/9781529709902 [22 October 2023].

Al-Jaradat, O.M. 2021. Research data management (RDM) in Jordanian public university libraries: Present status, challenges and future perspectives. *The Journal of Academic Librarianship*, 47(5): 102378. DOI: https://doi.org/10.1016/j.acalib.2021.102378.

Allen, M. 2017. *The sage encyclopaedia of communication research methods*, vol. 4. SAGE Publications, Inc, Thousand Oaks, CA. DOI: 10.4135/9781483381411.

American Library Association. 2015. Framework for information literacy for higher education. Available at: http://www.ala.org/acrl/standards/ilframework [22 October 2023].

Andrews, C.R. 2015. Embedded librarian ideas: best practices explored and redefined. *The International Journal of Educational Organization and Leadership*, 22(2): 1-14. Available at: https://academicworks.cuny.edu/bx_pubs/3/ [21 October 2023].

Andrews, T., 2012. What is social constructionism? *Grounded Theory Review: An International Journal*, 11(1). Available at: http://groundedtheoryreview.com/2012/06/01/what-is-social-constructionism/ [1 November 2023].

Ary, D., Jacobs, L.C., Irvine, C.K.S. & Walker, D. 2019. *Introduction to research in education*. United Kingdom: Cengage Learning.

Association of College and Research Libraries. 2000. *Information literacy competency standards for higher education*. Chicago: Association of College and Research Libraries. Available at: http://www.ala.org/ala/mgrps/divs/acrl/standards/informationliteracycompetency.cfm [25 October 2023]

Association of College and Research Libraries. 2018. Standards for Libraries in Higher Education. American Library Association. Available at: http://www.ala.org/acrl/standards/standardslibraries [7 November 2023].

Avuglah, B.K & Underwood, P.G. 2019. Research Data Management (RDM) Capabilities at the University of Ghana, Legon. *Library Philosophy and Practice*. 2258. Available at: https://digitalcommons.unl.edu/libphilprac/2258 [22 October 2023].

Azubogu, N. & Madu, C.C. 2019. Embedding a curriculum-based information literacy programme in Nigerian schools. *Information Technologist (The)*, 16(2): 199-204.

Babbie, E. 2004. *The practice of social research, 10th ed*. Belmont: Wadsworth.

Barbieri, B., Buonomo, I., Farnese, M.L. & Benevene, P. 2021. Organizational Capital: A resource for changing and performing in public administrations. *Sustainability*, 13(10): 5436. DOI: https://doi.org/10.3390/su13105436.

Barrett, D. & Twycross A. 2018. Data collection in qualitative research *Evidence-Based Nursing* 21:63-64. DOI: 10.1136/eb-2018-102939.

Barrie, S. C. 2004. A research-based approach to generic graduate attributes policy. *Higher Education Research & Development*, 23(3): 261-275. DOI: http://doi.org/10.1080/0729436042000235391.

Bazeley, P. 2018. Integration through Data Transformation 1: Qualitative Data to Statistical Variables. In: *Integrating Analyses in Mixed Methods Research*, 55 City Road, London: SAGE Publications Ltd pp.: 179-207. DOI: http://www.doi.org/10.4135/9781526417190.

Bell, G., Hey, T. & Szalay, A. 2009. Beyond the data deluge. *Science* 323(5919):1297-8. DOI: 10.1126/science.1170411.

Beretta, V., Desconnets, J.C., Mougenot, I., Arslan, M., Barde, J. & Chaffard, V. 2021. A user-centric metadata model to foster sharing and reuse of multidisciplinary datasets in environmental and life sciences. *Computers & Geosciences*, 154:1-10. DOI: https://doi.org/10.1016/j.cageo.2021.104807.

Bergin, T. 2018. *An introduction to data analysis: Quantitative, qualitative and mixed methods*. Sage: London.

Berman, E. A. 2017. An Exploratory Sequential Mixed Methods Approach to Understanding Researchers' Data Management Practices at UVM: Integrated Findings to Develop Research Data Services. *Journal of eScience Librarianship* 6(1): e1104. DOI: 10.7191/jeslib.2017.1104.

Bertam, C. & Christiansen, I. 2014. *Understanding research*. van Schaik Publisher: Pretoria.

Bhargava, R., Deahl, E., Letouzé, E., Noonan, A., Sangokoya, D. & Shoup, N. 2015. Beyond data literacy: reinventing community engagement and empowerment in the age of data. *Data-Pop Alliance White Paper Series. Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Lad and Overseas Development Institute) and Internews*. Available at: http://datapopalliance.org/wp-content/uploads/2015/11/Beyond-Data-Literacy-2015.pdf [23 October 2023].

Bhat, A. 2019. Data analysis in research: why data, types of data. data analysis in qualitative and quantitative research. Available at: https://www.questionpro.com/blog/data-analysis-in-research [20 October 2023].

Bierer, B. E., Crosas, M. & Pierce, H. H. 2017. Data authorship as an incentive to data sharing. *New England Journal of Medicine*, 376(17): 1684-1687. DOI: https://doi.org/10.1056/NEJMsb1616595.

Birkbeck, G., Nagle, T., & Sammon, D. 2022. Challenges in research data management practices: a literature analysis. *Journal of Decision Systems,* 31(Supplement 1): 153-167. DOI: https://doi.org/10.1080/12460125.2022.2074653.

Bjorkquist, C. 2011. *Stakeholder Regimes in Higher Education.* Waxmann: New York.

Blackburn, M., Alexander, J., Legan, J.D. & Klabjan, D. 2017. Big Data and the Future of R and D Management: The rise of big data and big data analytics will have significant implications for R and D and innovation management in the next decade. *Research-Technology Management*, 60(5): 43-51. Available at: https://dataanalytics.report/Resources/Whitepapers/cf88d0f4-1b4e-4e4a-97dd-a37841813b9e_Big%20Data%20Primer_0.pdf [28 October 2023].

Boateng, K.A. 2015, *Academic Library and Research Data Management Roles: The Case of Norwegian Libraries*, Master's thesis, Oslo and Akershus University College of Applied Sciences, Oslo, Norway, Available at: http://3.121.149.17/handle/10642/5093 [20 October 2023].

Bogdan, R. & Biklen, S.K. 2007. *Qualitative research for education; An introduction to theory and methods, 5th ed.* Allyn & Bacon: Boston, MA.

Bonnafous-Boucher, M. & Rendtorff, J. D. 2016. *Stakeholder Theory: A Model for Strategic Management*. Springer: Cham, Switzerland.

Borghi, J., & Van Gulick, A. (2022). Promoting Open Science Through Research Data Management. *Harvard Data Science Review*, 4(3). DOI: https://doi.org/10.1162/99608f92.9497f68e.

Borgman, C. L. (2015). *Big data, little data, no data: Scholarship in the networked* world, 1st ed. MIT Press: Cambridge, MA.

Borgman, C. L. 2012. The conundrum of sharing research data. *Journal of the Association for Information Science and Technology,* 63(6): 1059-1078. DOI: https://doi.org/10.1002/asi.22634.

Borycz, J., 2021. Implementing data management workflows in research groups through integrated library consultancy. *Data Science Journal*, 20:1-9. DOI: 10.5334/dsj-2021-009.

Bossaller, J. & Million, A.J. 2023. The research data life cycle, legacy data, and dilemmas in research data management. *Journal of the Association for Information Science and Technology*, 74(6): 701-706. DOI: 10.1002/asi.24645.

Brierley, J.A. 2017. The role of a pragmatist paradigm when adopting mixed methods in behavioural accounting research. *International Journal of Behavioural Accounting and Finance*, 6(2): 140-154. DOI: https://doi.org/10.1504/IJBAF.2017.10007499.

Briney, K. A., Coates, H. L. & Goben, A. 2020. Foundational practices of research data management. *Research Ideas and Outcomes*, 6: 1-17. DOI: 10.3897/rio.6.e56508.

Brodsky, M. 2016. The role of business librarians in teaching data literacy. *Ticker: The Academic Business Librarianship Review*, 1(3): 1-6. DOI: https://doi.org/10.3998/ticker.16481003.0001.301.

Brown, M.E. & Dueñas, A.N. 2020. A medical science educator's guide to selecting a research paradigm: building a basis for better research. *Medical Science Educator*, 30(1): 545-553. DOI: https://doi.org/10.1007/s40670-019-00898-9.

Brown, S., Alvey, E., Danilova, E., Morgan, H. & Thomas, A. 2018. Evolution of research support services at an academic library: specialist knowledge linked by core infrastructure, *New Review of Academic Librarianship*, 24(3-4): 337-348. DOI: 10.1080/13614533.2018.1473259.

Bryl, L. 2019. Intellectual capital measurement. In *Advanced Methodologies and Technologies in Library Science, Information Management, and Scholarly Inquiry*: 367-379. IGI Global. DOI: 10.4018/978-1-5225-7659-4.ch029.

Bryla, M. 2018. Data literacy: A critical skill for the 21st century. Available at: https://www.tableau.com/about/blog/2018/9/data-literacy-critical-skill-21st-century-94221 [21 October 2023].

Brymann, A. 2008. The end of the paradigm wars? In, Alasuutari, P., Bickman, L., and Brannen, J. (eds). *The SAGE handbook of social research methods*. Sage: Los Angeles.

Bueno, E., Salmador, M.P. & Longo-Somoza, M. 2014. Advances in the identification and measurement of Intellectual Capital and future developments in the Intellectual Capital research agenda: experience of the Intellectus Model and proposal of a synthetic index. *Knowledge Management Research and Practice*, 12(3): 339-349. DOI: https://doi.org/10.1057/kmrp.2014.11.

Bueno, E., Salmador, M.P., Rodríguez, Ó. & De Castro, G.M. 2006. Internal logic of intellectual capital: a biological approach. *Journal of Intellectual Capital.* 7(3): 394-405. DOI: 10.1108/14691930610681474.

Buhomoli, O.S. & Muneja, P.S. 2021. Research data handling by researchers in the selected universities in Tanzania. *University of Dar es Salaam Library Journal*, 16(2): 53-69. DOI: https://dx.doi.org/10.4314/udslj.v16i2.5.

Bunkar, A.R. & Bhatt, D.D. 2020. Perception of researchers; academicians of Parul University towards research data management system; role of library: a study. *DESIDOC Journal of Library & Information Technology*, 40(3):139-146. DOI: https://doi.org/10.14429/djlit.40.03.15302.

Burrell, N. A. & Gross, C. 2018. Quantitative Research, Purpose of. In Allen, M, *The sage encyclopaedia of communication research methods,* (Vols. 1-4). Thousand Oaks, CA: SAGE Publications, Inc. DOI: 10.4135/9781483381411.

Burress, T., Mann, E. & Neville, T. 2020. Exploring data literacy via a librarian-faculty learning community: A case study. *The Journal of Academic Librarianship*, 46(1): 1-7. DOI: 10.1016/j.acalib.2019.102076.

Cajal, C., Santolaria, J., Samper, D. & Garrido, A. 2015. Simulation of laser triangulation sensors scanning for design and evaluation purposes. *International Journal of Simulation Modelling*, 14(2): 250-264. Available at: http://w.ijsimm.com/Full_Papers/Fulltext2015/text14-2_250-264.pdf [24 October 2023].

Calzada, P. J. & Marzal, M.Á. 2013. Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2): 123-134. DOI: https://doi.org/10.1515/libri-2013-0010.

Campbell, H.A., Micheli-Campbell, M.A. & Udyawer, V. 2019. Early career researchers embrace data sharing. *Trends in ecology & evolution*, 34(2): 95-98. DOI: https://doi.org/10.1016/j.tree.2018.11.010.

Candy, P.C. 2002. Lifelong learning and information literacy. Available at: *https://www.researchgate.net/profile/Philip_Candy/publication/268299706* [24 October 2023].

Carballo-Garcia, A. and Boté-Vericad, J.J. 2022. Fair Data: History and Present Context. *Central European Journal of Educational Research*, 4(2): 45-53. DOI: https://doi.org/10.37441/cejer/2022/4/2/11379.

Carlson J & Bracke S M. 2015. Planting the seeds for data literacy: Lessons learned from a student-centered education. *International Journal of Digital Curation*. 10(1): 95-110. DOI: 10.2218/ijdc.v10i1.348.

Carlson, J, Nelson, M.S, Johnston, L.R. & Koshoffer, A. 2015. Developing data literacy programs: working with faculty, graduate students and undergraduates. *Bulletin of the American Society for Information Science and Technology* 41(6): 14-17. DOI: 10.1002/bult.2015.1720410608.

Carlson, J. & Johnston, L.R. (Eds). 2015. *Data information literacy: Librarians, data, and the education of a new generation of researchers. vol. 2*. Purdue University Press: Indiana.

Carlson, J. & Johnston, L.R. 2015. Introduction. In Carlson, J. and Johnston, L.R., *Data information literacy librarians, Data, and the Education of a New Generation of Researchers*:1-8. Purdue University Press: Indiana.

Carlson, J. & Kneale, R. 2011. Embedded librarianship in the research context: Navigating new waters. *College & Research Libraries News*, 72(3): 167-170. Available at: https://crln.acrl.org/index.php/crlnews/issue/view/565 [29 October 2023].

Carlson, J. 2014. The Use of Life Cycle Lifecycle Models in Developing and Supporting Data Services. In: Joyce M Ray (ed). 2014. *Research Data Management. Practical Strategies for Information Professionals*: 63-86. Purdue University Press, West Lafayette, IN.

Carlson, J., Fosmire, M., Miller, C.C. & Nelson, M.S. 2011. Determining data information literacy needs: a study of students and research faculty. *Portal: Libraries and the Academy*. 11(2): 629-657. DOI: 10.1353/pla.2011.0022.

Carlson, J., Jeffryes, J., Johnston, L.R., Nichols, M., Westra, B. & Wright, S.J. 2015. 'An exploration of the data information literacy competencies: Findings from the project interviews', in Carlson, J. and Johnston, L. R. (eds.) *Data information literacy: librarians, data, and the education of a new generation of researchers*: 51-69. Purdue University Press: Indiana.

Carlson, J., Nelson, M.S., Johnston, L.R. & Koshoffer, A. 2015. Developing data literacy programs: Working with faculty, graduate students and undergraduates. *Bulletin of the association for information science and technology*. 41(6): 14-17. DOI: https://doi.org/10.1002/bult.2015.1720410608.

Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. 2014. The use of triangulation in qualitative research. *Oncology nursing forum*, 41(5): 545–547. DOI: https://doi.org/10.1188/14.ONF.545-547.

Caruth, G.D. 2013. Demystifying mixed methods research design: A review of the literature. *Mevlana International Journal of Education*, 3(2): 112-122. DOI: http://dx.doi.org/10.13054/mije.13.35.3.2.

Chankseliani, M., Qoraboyev, I. & Gimranova, D. 2021. Higher education contributing to local, national, and global development: new empirical and conceptual insights. *Higher Education*, 81(1): 109-127. DOI: https://doi.org/10.1007/s10734-020-00565-8.

Chawinga, W.D. & Zinn, S. 2021. Research data management in universities: a comparative study from the perspectives of librarians and management. *International Information & Library Review*, 53(2): 97-111. DOI: https://doi.org/10.1080/10572317.2020.1793448

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

Childs, L. 2017. To uphold and resist: Protecting intellectual freedom through progressive librarianship. *The Serials Librarian*, 73(1): 58-67. DOI: https://doi.org/10.1080/0361526X.2016.1270248.

Chilisa, B. & Kawulich, B. 2012. Selecting a research approach: Paradigm, methodology and methods. In *Doing social research: A global context*, 5(1): 51-61. Available at: https://www.researchgate.net/profile/Barbara-Kawulich/publication/257944787_Selecting_a_research_approach_Paradigm_methodology_and_methods/links/56166fc308ae37cfe40910fc/Selecting-a-research-approach-Paradigm-methodology-and-methods.pdf [29 October 2023].

Chión, S.J., Charles, V. & Morales, J. 2020. The impact of organisational culture, organisational structure and technological infrastructure on process improvement through knowledge sharing. *Business Process Management Journal*, 26(6): 1443-1472. DOI: https://doi.org/10.1108/BPMJ-10-2018-0279.

Christensen-Dalsgaard, B., van den Berg, M., Grim, R., Horstmann, W., Jansen, D., Pollard, T. & Roos, A. 2012. Ten recommendations for libraries to get started with research data management, *LIBER*. Available at: Available at: http://www.libereurope.eu/news/ten-recommendations-forlibraries-to-get-started-with-research-data-management [7 November 2023].

Cibangu, K. S. 2012. Qualitative Research: The Toolkit of Theories in the Social Sciences. In Asunción López-Varela (Ed.). *Theoretical and Methodological Approaches to Social Sciences and Knowledge Management*. 95-126. InTech.

Clark-Kazak, C. 2017. Ethical considerations: Research with people in situations of forced migration. *Refuge: Canada's Journal on Refugees,* 33(2): 11-17. DOI: https://doi.org/10.7202/1043059ar.

Coghlan, D. & Brydon-Miller, M. 2014. *The SAGE encyclopaedia of action research* (Vols. 1-2). SAGE Publications Ltd: London. DOI: 10.4135/9781446294406.

Cohen, L.L., Manion & K. Morrison. 2007. *Research methods in education. 6th ed*. London: Routledge Falmer.

Collins, S., Genova, F.,Harrower, N., Hodson, S., Jones, S., Laaksonen, L., Mietchen, D., Petrauskaitė, R. & Wittenburg, P. 2018. Turning FAIR into reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data. Available at: https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf [21 October 2023].

Commission for University Education. 2014. *University Standards and Guidelines*, Nairobi: CUE.

Connelly, R., Playford, C. J., Gayle, V. & Dibben, C. 2016. The role of administrative data in the big data revolution in social science research. *Social science research*, 59: 1-12. DOI: https://doi.org/10.1016/j.ssresearch.2016.04.015.

Corrall, S. 2012. Roles and responsibilities: Libraries, librarians and data. In G. Pryor (Ed.), *Managing research data*:105-133. London: Facet.

Corrall, S. 2014. Designing libraries for research collaboration in the network world: An exploratory study. *Liber Quarterly*, 24(1): 17-48. DOI: https://doi.org/10.18352/lq.9525.

Corrall, S. 2019a. Data Literacy as a Human Right: Roles and Responsibilities of Libraries as Datamediaries. Available at: https://lida.ffos.hr/2018/datoteke/presentations_2018/LIDA_2018_Corrall.pdf [27 October 2023].

Corrall, S. 2019b. The Wicked Problem of Data Literacy: A Call for Action. In *LILAC: The Information Literacy Conference*, April 24-26, 2019, Nottingham, UK. Available at: http://d-scholarship.pitt.edu/id/eprint/36759 [20 October 2023].

Corrall, S. 2019c. The Social Mission of 21st Century Research Libraries: Building Data Literate Communities. In: *48th LIBER Annual Conference: Research Libraries for Society*, June 26-28, 2019, Dublin, Ireland. DOI: 10.5281/zenodo.3259728.

Corrall, S. 2019d. Repositioning Data Literacy as a Mission-Critical Competence. In: ACRL 2019: *Recasting the Narrative*. Available at: http://d-scholarship.pitt.edu/id/eprint/36975 [20 October 2023].

Corti, L. & Fielding, N. 2016. Opportunities from the digital revolution: Implications for researching, publishing, and consuming qualitative research. *Sage Open*, 6(4): 1-13. DOI: 10.1177/2158244016678912.

Cousijn, H., Kenall, A., Ganley, E., Harrison, M., Kernohan, D., Lemberger, T., ... & Clark, T. 2018. A data citation roadmap for scientific publishers. *Scientific data*, 5(1): 1-11. DOI: 10.1038/sdata.2018.259.

Couture, J.L., Blake, R.E., McDonald, G. & Ward, C.L. 2018. A funder-imposed data publication requirement seldom inspired data sharing. *PloS one*, 13(7): 1-13. DOI: 10.1371/journal.pone.0199789.

Cox, A. M., Kennan, M. A., Lyon, L., Pinfield, S., & Sbaffi, L. 2019. Maturing research data services and the transformation of academic libraries. *Journal of Documentation,* 75(6): 1432-1462. DOI: 10.1108/JD-12-2018-0211.

Cox, A.M. & Pinfield, S. 2014. Research data management and libraries: Current activities and future priorities. *Journal of librarianship and information science*, 46(4): 299-316. DOI: https://doi.org/10.1177/0961000613492542.

Cox, A.M. & Tam, W.W.T. 2018. A critical analysis of lifecycle models of the research process and research data management. *Aslib Journal of Information Management*. 70(2): 142-157. DOI: http://doi.org/10.1108/AJIM-11-2017-0251.

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

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. *Journal of the Association for Information Science and Technology*, 68(9): 2182-2200. DOI: https://doi.org/10.1002/asi.23781.

Cox, A.M., Kennan, M.A., Lyon, L., Pinfield, S. & Sbaffi, L. 2019. Maturing research data services and the transformation of academic libraries. *Journal of Documentation*, 75(6): 1432-1462. DOI: https://doi.org/10.1108/JD-12-2018-0211.

Creswell, J.W. & Clark. V.P. 2018. *Designing and conducting mixed methods research, 3rd ed*. SAGE: Los Angeles.

Creswell, J.W. & Creswell, J.D. 2018. Research design: *Qualitative, Quantitative and Mixed Methods Approaches*. Sage: London.

Creswell, J.W. 2013. *Research design: qualitative, quantitative, and mixed methods approaches 4th ed*. Sage Publications: Thousand Oaks, CA.

Creswell, J.W. 2014. *Research design: international student edition. 4th ed*. Sage: Los Angeles.

Critchlow & van Dam, (Eds). 2013. *Data-Intensive Science*. CRC Press: Boca Raton, FL.

D'Ignazio, C. & Bhargava, R. 2015. Approaches to building big data literacy. In *Proceedings of the Bloomberg Data for Good Exchange Conference*. Available at: http://www.kanarinka.com/wp-content/uploads/2021/01/DIgnazio-and-Bhargava-Approaches-to-Building-Big-Data-Literacy.pdf [20 October 2023].

Dai, Y. 2020. How many ways can we teach data literacy?. *IASSIST Quarterly*, 43(4): 1-11. DOI: https://doi.org/10.29173/iq963.

Dallmeier-Tiessen, S., Darby, R., Goerner, B., Hyppoelae, J., Igo-Kemenes, P., Kahn, D., ... and Stocker, M. 2014. Enabling sharing and reuse of scientific data. *New Review of Information Networking,* 19(1): 51-69. DOI: https://doi.org/10.1080/13614576.2014.883936.

Daniel, B. 2015. Big Data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*. 46(5): 904-920. DOI: 10.1111/bjet.12230.

Deja, M., Rak, D. & Bell, B. 2021. Digital transformation readiness: perspectives on academiathe andacademia and library outcomes in information literacy. *The Journal of Academic Librarianship*, 47(5), 102403. DOI: https://doi.org/10.1016/j.acalib.2021.102403.

Delaney, G. & Bates, J. 2018. How can the university library better meet the information needs of research students? Experiences from Ulster University, *New Review of Academic Librarianship*, 24(1): 63-89. DOI: http://doi.org/10.1080/13614533.2017.1384267.

Demir, S.B. & Pismek, N. 2018. A convergent parallel mixed-methods study of controversial issues in social studies classes: a clash of ideologies. *Educational Sciences: Theory and Practice*, 18(1): 119-149. DOI: 10.12738/estp.2018.1.029.

Deng, S. & Hu, X. 2014. Creating a knowledge map for the Research Lifecycle. *Knowledge Maps and Information Retrieval (KMIR) 2014 Conference Proceedings*. ACM/IEEE Joint Conference on Digital Libraries (JCDL) 2014, 8th - 12th September 2014,

London. Available at: http://ceur-ws.org/Vol-1311/paper5.pdf Slides available at: www.gesis.org/fileadmin/upload/kmir2014/paper5_slides.pdf [2 November 2023].

Denicolo, P., Long, T. and Bradley-Cole, K. 2016. *Constructivist approaches and research methods: A practical guide to exploring personal meanings*. Sage: London.

Denzin, N., & Lincoln, Y. 2011. *The SAGE handbook of qualitative research*. Thousand Oaks, CA: SAGE: Los Angeles.

Derakhshan, M. & Singh, D. 2011. Integration of information literacy into the curriculum: a meta-synthesis. *Library review*. 60(3): 218-229. DOI: http://doi.org/10.1108/00242531111117272.

Doonan, A., Akmon, D. & Cosby, E. 2019. An exploratory analysis of social science graduate education in data management and data sharing. Available at: https://deepblue.lib.umich.edu/handle/2027.42/150174 [20 October 2023].

Doyle, L., McCabe, C., Keogh, B., Brady, A., & McCann, M. 2020. An overview of the qualitative descriptive design within nursing research. *Journal of research in nursing: JRN*, 25(5): 443-455. DOI: https://doi.org/10.1177/1744987119880234.

Dresang, E T., & McClelland, K. 1999. Radical change: Digital age literature and learning. *Theory into Practice*. 38(3):160-7. DOI: 10.1080/00405849909543848.

Dresang, E. T. 1999. *Radical change: Books for youth in a digital age*. H. W: New York.

Dresang, E. T., & Kotrla, B. 2009. Radical change theory and synergistic reading for digital age youth. *The Journal of Aesthetic Education*, 43(2): 92-107. Available at: https://www.jstor.org/stable/40263787 [25 October 2023].

Dresang, E. T. 2006. Intellectual freedom and libraries: Complexity and change in the twenty-first-century digital environment. *Library Quarterly* 76(2): 169-92. DOI: 10.1086/506576.

Dresang, E.T. 2005. Access: The information-seeking behaviour of youth in the digital environment. *Library Trends*, 54(2): 178-196. DOI: 10.1353/lib.2006.0015.

Dudovskiy, J. 2022. *The ultimate guide to writing a dissertation in business studies: a step-by-step assistance, 6th edition*. Available at: https://research-methodology.net/about-us/ebook/ [22 October 2023].

Easterby-Smith, M., Thorpe, R., & Lowe, A. 2002. *Management research: An introduction*. Sage Publications: London.

Edmonds, W. & Kennedy, T. 2017. Embedded approach. In: *An Applied Guide to Research Designs: Quantitative, Qualitative, and Mixed Methods*, 2nd ed. Thousand Oaks, CA: SAGE Publications, Inc pp.: 189-195. Available at: http://www.doi.org/10.4135/9781071802779 [23 October 2023].

Eggers, W.D., Hamill, R. & Ali, A. 2013. Data as the new currency: government's role in facilitating the exchange. *Review of Deloitte*. (13): 18-31.

Elouataoui, W., El Alaoui, I. & Gahi, Y. 2022. Metadata quality in the era of big data and unstructured content. In *Proceedings 2nd International Conference Big Data, Modelling Machine Learning. (BML).* SciTePress, Setúbal, Portugal: 488-495. DOI: 10.5220/0010737400003101.

Etikan, I. & Bala, K. 2017. Sampling and sampling methods. *Biometrics & Biostatistics International Journal*, 5(6): 1-3. Available at: https://medcraveonline.com/BBIJ/BBIJ-05-00149.pdf [23 October 2023].

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

European Commission, Directorate-General for Research and Innovation. 2018. *Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data*. Publications Office. DOI: https://data.europa.eu/doi/10.2777/1524.

European Commission. 2014. Towards better access to scientific information: Boosting the benefits of public investments in research. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Brussels*. Available: https://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf [2 November 2023].

Family Health International-Kenya. 2005. Strengthening research and data management: new efforts at the Kenya division of reproductive health. Available at: https://pdf.usaid.gov/pdf_docs/Pnadk371.pdf [25 October 2023].

Farooq, O., Heppler, J.A. & Ehrig-Page, K.M. 2022. Creating capacity for research data services at regional universities: A Case Study. In Julia Bauder, (Eds), *Teaching Research Data Management*: 155-176. ALA Editions. Chicago.

Field, A. 2018. *Discovering statistics using IBM SPSS statistics*. Sage publications: London.

Fischer, G., Rohde, M. & Wulf, V. 2007. Community-based learning: The core competency of residential, research-based universities. *International Journal of Computer-Supported Collaborative Learning*, *2*(1): 9-40. DOI: https://doi.org/10.1007/s11412-007-9009-1.

Flak, L.S., Nordheim, S. & Munkvold, B.E. 2008. Analyzing stakeholder diversity in G2G efforts: Combining descriptive stakeholder theory and dialectic process theory. *E-Service Journal*, 6(2): 3-23. DOI: https://doi.org/10.2979/esj.2008.6.2.3.

Fleming, J. & Zegwaard, K.E. 2018. Methodologies, methods and ethical considerations for conducting research in work-integrated learning. *International Journal of Work-Integrated Learning*, 19(3): 205-213. Available at: https://files.eric.ed.gov/fulltext/EJ1196755.pdf [22 October 2023].

Flick, U. 2011. *Introducing research methodology: a beginner's guide to doing a research project*. Sage: Los Angeles.

Flores, J.R., Brodeur, J.J., Daniels, M.G., Nicholls, N., & Turnator, E. 2015. Libraries and the research data management landscape. *The Process of Discovery: The CLIR Postdoctoral Fellowship Program and the Future of the Academy*, 2010: 82-102. Available at: www.clir.org/wpcontent/uploads/sites/6/RDM.pdf [8 November 2023].

Fontichiaro, K. & Oehrli, J.A. 2016. Why data literacy matters. *Knowledge quest*, *44*(5): 21-27. Available at: https://files.eric.ed.gov/fulltext/EJ1099487.pdf. [8 November 2023].

Foo, S., Chaudhry, A.S., Majid, S. & Logan, E. 2002. Academic libraries in transition: Challenges ahead. *Proceedings of the World Library Summit, Singapore*, April 22-26. Available at: http://islab.sas.ntu.edu.sg:8000/user/schubert/publications/2002/02wls_fmt.pdf [1 November 2023].

Fraenkel, J.R., Wallen, N.E. & Hyun, H.H. 2012. *How to design and evaluate research in education*, *8th ed*. McGraw-hill: Boston.

Frank, E. P. & Pharo, N. 2016. Academic librarians in data information literacy instruction: a case study in meteorology. *College and Research Libraries,* 77(4): 536-532. DOI: https://doi.org/10.5860/crl.77.4.536.

Frederick, A. 2019. The role of academic libraries in research data management: A case in Ghanaian university libraries. *Open Access Library Journal*, 6(03): 1-16. DOI: https://doi.org/10.4236/oalib.1105286.

Freeman, R. E. 1984. *Strategic management: A stakeholder approach.* Pitman: Boston, MA.

Freeman, R. E. 2015. Stakeholder theory. *Wiley encyclopedia of management*: 1-6. DOI: https://doi.org/10.1002/9781118785317.weom020179.

Freeman, R. E., Phillips, R., & Sisodia, R. 2020. Tensions in stakeholder theory. *Business and Society*, 59(2): 213-231. DOI: 10.1177/0007650318773750.

Friedlander, A. & Adler, P. 2006. To stand the test of time: long-term stewardship of digital data sets in science and engineering. A Report to the National Science Foundation from the ARL Workshop on New Collaborative Relationships--The Role of Academic Libraries in the Digital Data Universe (Arlington, Virginia, September 26-27, 2006). *Association of Research Libraries*. Available at: https://files.eric.ed.gov/fulltext/ED528649.pdf [1 November 2023].

Fry, B.J. 2004. *Computational Information Design*, Massachusetts Institute of Technology, Boston.

Fuhr, J. 2022. Developing data services skills in academic libraries. *College & Research Libraries*, 83(3): 474. DOI: https://doi.org/10.5860/crl.83.3.474.

Gajbe, S.B., Tiwari, A. & Singh, R.K. 2021. Evaluation and analysis of data management plan tools: a parametric approach. *Information Processing & Management*, 58(3): 102480. DOI: https://doi.org/10.1016/j.ipm.2020.102480.

Gal, I. 2002. Adults' statistical literacy: Meanings, components, responsibilities. *International statistical review*, 70(1): 1-25. DOI: http://doi.org/10.1111/j.1751-5823.2002.tb00336.x.

Garnett, A., Leahey, A., Savard, D., Towell, B. & Wilson, L. 2017. Open metadata for research data discovery in Canada. *Journal of Library Metadata*, 17(3-4): 201-217. DOI: https://doi.org/10.1080/19386389.2018.1443698.

Gay, L.R. & Mills, G.E. 2019. *Educational research: Competencies for analysis and applications, 12th edition*. Pearson: New Jersey.

Ghasempour, Z., Bakar, M.N. & Jahanshahloo, G.R. 2014. Mix-method design in educational research: Strengths and challenges. *International Journal of Pedagogical Innovations,* 2(02): 83-90. Available at: https://www.researchgate.net/profile/Zahra-Ghasempour-4/publication/328565403_Mix-_Method_Design_in_Educational_ResearchStrengths_and_Challenges/links/6096cbc9299bf1ad8d8945e0/Mix-Method-Design-in-Educational-ResearchStrengths-and-Challenges.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail&_rtd=e30%3D [28 October 2023].

Gibson, J. P. & T. Mourad. 2018. The growing importance of data literacy in life science education. *American Journal of Botany* 105(12): 1953–1956. DOI:10.1002/ajb2.1195.

Giese, T.G., Wende, M., Bulut, S. & Anderl, R. 2020. Introduction of data literacy in the undergraduate engineering curriculum. *IEEE Global Engineering Education Conference (EDUCON), Porto, Portugal*: 1237-1245. DOI: 10.1109/EDUCON45650.2020.9125212.

Gilligan, D. 2016. Global data wars: building a thriving data economy for Australia. Available at: http://reinventure.com.au/wp-content/uploads/2016/10/GlobalDataWarsReportReinventure.pdf [21 October 2023].

Giudice, P.L., Musarella, L., Sofo, G. & Ursino, D. 2019. An approach to extracting complex knowledge patterns among concepts belonging to structured, semi-structured and unstructured sources in a data lake. *Information Sciences*, 478: 606-626. DOI: https://doi.org/10.1016/j.ins.2018.11.052.

Given, L. M. (Ed.). 2008. The SAGE Encyclopedia of Qualitative Research Methods. (Vols. 1-0). SAGE Publications, Inc.:DOI: https://doi.org/10.4135/9781412963909.

Goben, A. & Griffin, T. 2019. In aggregate: Trends, needs, and opportunities from research data management surveys. *College & Research Libraries*, 80(7): 903-924. DOI: https://doi.org/10.5860/crl.80.7.903.

Goldman, J., Chen, H.W. & Palau, J.P. 2023. Data management. In *Translational Surgery:* 451-459. Academic Press. DOI: https://doi.org/10.1016/B978-0-323-90300-4.00078-1.

Grant, C. & Osanloo, A. 2014. Understanding, selecting, and integrating a theoretical framework in dissertation research: creating the blueprint for your "house". Available at: http://jolle.coe.uga.edu/wp-content/uploads/2015/02/89596_manuscript-file_249104.pdf [20 October 2023].

Gray, J. & Darbishire, H. 2011. Beyond access: open government data and the right to (re) use public information. Available at: https://ssrn.com/abstract=2586400 [20 October 2023].

Groenewald, R. and Breytenbach, A. 2011. The use of metadata and preservation methods for continuous access to digital data. *The Electronic Library*, 29(2): 236-248. DOI: https://doi.org/10.1108/02640471111125195.

Groenland, E., and L. P. Dana. 2020. *Qualitative Methodologies and Data Collection Methods: Toward Increased Rigour in Management Research* (Vol. 1). World Scientific: Singapore.

Gross, M., Latham, D. & Julien, H. 2018. What the framework means to me: Attitudes of academic librarians toward the ACRL framework for information literacy for higher education. *Library & information science research*, 40(3-4): 262-268. DOI: https://doi.org/10.1016/j.lisr.2018.09.008.

Guba, E. 1990. *The Paradigm Dialogue*. SAGE Publications: Newbury Park, California.

Guss, S. 2016. A studio model for academic data services. In Kellam, L. and Thompson, K. (Eds), *Databrarianship: The Academic Data Librarian in Theory and Practice*, Association of College and Research Libraries, Chicago, IL: 9-24.

Hamad, F., Al-Fadel, M. & Al-Soub, A. 2021. Awareness of research data management services at academic libraries in Jordan: Roles, responsibilities and challenges. *New Review of Academic Librarianship*, 27(1): 76-96. DOI: https://doi.org/10.1080/13614533.2019.1691027.

Hamblin, Y. 2005. Library portals case studies. *Assignation*, 22(3): 26-29.

Hammersley, M. 2013. *What is Qualitative Research?* Bloomsbury: London.

Harris, S. 2012. *Moving towards an Open Access Future: The Role of Academic Libraries*. A report on a roundtable commissioned by SAGE, in association with the British Library, 2012. Web. 4 Mar. 2014. Available at: http://www.uk.sagepub.com/repository/binaries/pdf/Library-OAReport.pdf [20 October 2023].

Hart, E.M., Barmby, P., LeBauer, D., Michonneau, F., Mount, S., Mulrooney, P., Poisot, T., Woo, K.H., Zimmerman, N.B. & Hollister, J.W. 2016. Ten simple rules for digital data storage. *PLoS computational biology*, 12(10), e1005097. DOI: https://doi.org/10.1371/journal.pcbi.1005097.

Hart, K. 2019. Critical analysis of Big Data challenges and analytical methods. *Journal of Child Behaviour,* 1(1): 28-35. DOI: https://doi.org/10.1016/j.jbusres.2016.08.001.

Hassett, D.D. 2005. Reading hypertextually: Children's literature and comprehension instruction. *New Horizons in Learning*.

Hastings, S. 2010. Triangulation. In N. J. Salkind (Ed.), *Encyclopedia of research design:* pp. 1538-1540. SAGE Publications, Inc. DOI: https://dx.doi.org/10.4135/9781412961288.n469.

Heidorn, P. B. 2011. The Emerging Role of Libraries in Data Curation and Escience. *Journal of Library Administration,* 51, 7-8 (2011): 662-672. DOI: https://doi.org/10.1080/01930826.2011.601269.

Heidorn, P.B., 2011. The emerging role of libraries in data curation and e-science. *Journal of Library Administration*, 51(7-8): 662-672. DOI: https://doi.org/10.1080/01930826.2011.601269

Henty, M. 2014. Dreaming of data: the library's role in supporting e-research and data management. Available at: http://hdl.handle.net/1885/47617 [24 October 2023].

Hesse-Biber, S. & Johnson, R. B. (Eds.). 2015. *The Oxford handbook of multimethod and mixed methods research inquiry*. Oxford University Press: New York.

Hey, T., Gannon, D. & Pinkelman, J. 2012. The future of data-intensive science. *Computer.* 4 (5): 81-82. DOI: 10.1109/MC.2012.181.

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

Hilgartner, S. & Brandt-Rauf, S.I. 1994. Data access, ownership, and control: Toward empirical studies of access practices. *Knowledge*, 15(4): 355-372. DOI: https://doi.org/10.1177/107554709401500401.

Hong, Q.N., Pluye, P., Bujold, M. & Wassef, M. 2017. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative

and quantitative evidence. *Systematic reviews*, 6(1): 1-14. DOI: https://doi.org/10.1186/s13643-017-0454-2.

Houghton, T. 2011. Does positivism really "work" in the social sciences. *E-International Relations*: 1-4: Available at: https://www.e-ir.info/pdf/14310 [29 October 2023].

Hsieh, H.-F. & Shannon, S. E. 2005. Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, 15(9): 1277-1288. Doi: 10.1177/1049732305276687.

Huang, Y., Cox, A.M. & Sbaffi, L. 2021. Research data management policy and practice in Chinese university libraries. *Journal of the Association for Information Science and Technology*, 72(4): 493-506. DOI: https://doi.org/10.1002/asi.24413.

Huenneke, L.F., Stearns, D.M., Martinez, J.D. & Laurila, K. 2017. Key strategies for building research capacity of university faculty members. *Innovative higher education*, 42: 421-435. DOI: 10.1007/s10755-017-9394-y.

Islam, M., 2020. Data analysis: types, process, methods, techniques and tools. *International Journal on Data Science and Technology*, 6(10). DOI: 10.11648/j.ijdst.20200601.12.

Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., ... and Schultes, E. 2020. FAIR principles: interpretations and implementation considerations. Data Intelligence, 2(1-2): 10-29. DOI: http://doi.org/10.1162/dint_r_00024.

Janssen, M., Brous, P., Estevez, E., Barbosa, L.S. & Janowski, T. 2020. Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, 37(3): 101-493. DOI: https://doi.org/10.1016/j.giq.2020.101493.

Jao, I., Kombe, F., Mwalukore, S., Bull, S., Parker, M., Kamuya, D., Molyneux, S. & Marsh, V. 2015. Research stakeholders' view on benefits and challenges for public health research data sharing in Kenya: the importance of trust and social relations. PLoS ONE 10(9): e0135545. DOI: https://doi.org/10.1371/journal.pone.0135545.

Jharotia, A. K. 2018. Importance of digital preservation in digital era. Available at: https://www.researchgate.net/publication/326534604_IMPORTANCE_OF_DIGITAL _PRESER VATION_IN_DIGITAL_ERA/link/5b531283a6fdcc8dae35569d/download [20 October 2023].

Johnson, C.A. 2012. *The information diet: a case for conscious consumption*. O'Reilly Media: Sebastopol, CA.

Johnson, K.A. & Steeves, V. 2019. Research data management among life sciences faculty: Implications for library service. *Journal of EScience Librarianship*, 8(1). DOI: https://doi.org/10.7191/jeslib.2019.1159.

Johnson, R.B., Onwuegbuzie, A.J. & Turner, L.A. 2007. Toward a definition of mixed methods research. *Journal of Mixed Methods Research,* 1:112-133. DOI: 10.1177/1558689806298224

Jones, S., Pryor, G., & Whyte, A. 2013. *How to Develop Research Data Management Services - A Guide for HEIs*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available at: http://www.dcc.ac.uk/resources/how-guides [29 October 2023].

Jones, T.M., Wicks, A.C., Freeman, R.E. 2002. Stakeholder theory: the state of the art. In N.E Bowie (ed.), *The Blackwell Guide to Business Ethics*. Blackwell: Malden; 19–37.

Jung, H. 2019. The evolution of social constructivism in political science: past to present. *SAGE Open*, 9(1): 1-10. DOI: https://doi.org/10.1177/2158244019832703.

Kaehne, A. 2017. Integration as a scientific paradigm. *Journal of Integrated Care*, 25(4): 271-279. DOI: https://doi.org/10.1108/JICA-07-2017-0023.

Kahn, M., Higgs, R., Davidson, J. & Jones, S. 2014. Research Data Management in South Africa: How We Shape Up, *Australian Academic & Research Libraries*, 45(4): 296-308. DOI: http://doi.org/10.1080/00048623.2014.951910.

Kamal, S.S.L.B.A. 2019. Research paradigm and the philosophical foundations of a qualitative study. *International Journal of Social Sciences*, 4(3): 100-230. DOI: https://dx.doi.org/10.20319/pijss.2019.43.13861394.

Karchegani, M.R., Sofian, S. & Amin, S.M. 2013. The relationship between intellectual capital and innovation: A review. *International Journal of Business and Management Studies*. 2(1): 561-581. Available at: https://www.researchgate.net/profile/Mohammad-Rahmani-Karchegani/publication/236902495_The_Relationship_between_Intellectual_Capital_and_Innovation/links/0046351a3282ab1657000000/The-Relationship-between-Intellectual-Capital-and-Innovation.pdf [22 October 2023].

Kaushik, V. & Walsh, C.A. 2019. Pragmatism as a research paradigm and its implications for social work research. *Social Sciences*, 8(9): 255. DOI: https://doi.org/10.3390/socsci8090255.

Khan, H.R., Kim, J. & Chang, H.C. 2018. Toward an Understanding of Data Literacy. *iConference 2018 Proceedings*. Available at: https://pdfs.semanticscholar.org/0c23/38974a51cc7b562d31f5109547395c384186.pdf [24 October 2023].

Khan, N., Thelwall, M. & Kousha, K. 2023. Data sharing and reuse practices: disciplinary differences and improvements needed. *Online Information Review*, 47(6):1036-1064. https://doi.org/10.1108/OIR-08-2021-0423.

Kimbroughand, J. L. & Gasaway, L.N. 2015. Publication of government-funded research, open access, and the public interest. *Vand. J. Ent. & Tech. L.* 18: 266-302. Available

at:
https://heinonline.org/HOL/LandingPage?handle=hein.journals/vanep18&div=14&id
=&page= [21 October 2023].

Kirk, A. 2019. *Data visualisation: a handbook for data driven design, 2nd edition*. Sage:
London

Kitchin, R. 2014. *The data revolution: Big data, open data, data infrastructures and their
consequences*. Sage publishing.

Kivunja, C. & Kuyini, A. B. 2017. Understanding and Applying Research Paradigms in
Educational Contexts. *International Journal of higher education*, 6(5): 26-41. DOI:
https://doi.org/10.5430/ijhe.v6n5p26.

Kivunja, C. 2018. Distinguishing between theory, theoretical framework, and conceptual
framework: A systematic review of lessons from the field. *International journal of
higher education*, 7(6): 44-53. DOI: https://doi.org/10.5430/ijhe.v7n6p44.

Klain G.L. & Shoham, S. 2019. The role of academic libraries in research and teaching.
Journal of Librarianship and Information Science, 51(3): 721-736. DOI:
https://doi.org/10.1177/0961000617742462.

Knox, E.J. 2014. Supporting intellectual freedom: Symbolic capital and practical philosophy
in librarianship. *The Library Quarterly*, 84(1): 8-21. Available at:
https://www.journals.uchicago.edu/doi/full/10.1086/674033?casa_token=TnMrSPI35
ekAAAAA%3A21lpnUtk9pGZKnN5y2xiWtJ96ejkeQ_RhUOO_KH0BHWf7w1S7i1
O0lC8ECCUzhV36ucC1Rdt4LRu6A [21 October 2023].

Koh, K. 2015. Radical change theory: Framework for empowering digital youth. *Journal of
Research on Libraries and Young Adults*. 5:1-21. Available at:
http://www.yalsa.ala.org/jrlya/2015/01/radical-change-theory-framework-for-
empowering-digital-youth/#more-238 [1 November 2023].

Kohler, U., Kreuter, F. & Stuart, E.A. 2019. Nonprobability sampling and causal
analysis. *Annual review of statistics and its application*, 6:149-172. DOI:
https://doi.org/10.1146/annurev-statistics-030718-104951.

Köksal, H. 2019. The role of universities for workplace innovation: A Turkish
case. *European Journal of Workplace Innovation*, 5(1): 83-95. DOI:
https://doi.org/10.46364/ejwi.v5i1.561.

Koltay, T. 2015. Data literacy: in search of a name and identity. *Journal of Documentation*.
71(2): 401-415. DOI: 10.1108/JD-02-2014-0026.

Koltay, T. 2016, Are you ready? Tasks and roles for academic libraries in supporting
Research 2.0. *New Library World,* 117(1/2): 94-104. DOI: 10.1108/NLW-09-2015-
0062.

Koltay, T. 2016a. Data governance, data literacy and the management of data quality. *IFLA journal*. 42(4): 303-312. DOI: http://doi.org/10.1177/0340035216672238.

Koltay, T. 2016b. Are you ready? Tasks and roles for academic libraries in supporting research 2.0. *New Library World*. 117(1): 94-104. DOI: http://doi.org/10.1108/NLW-09-2015-0062.

Koltay, T. 2017a. Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science*. 49(1): 3-14. DOI: 10.1177/0961000615616450.

Koltay, T. 2017b. Research 2.0 and research data services in academic and research libraries: priority issues. *Library Management*. 38(6-7): 345-353. DOI: http://doi.org/10.1108/LM-11-2016-0082.

Koltay, T., 2019. Data curation in academic libraries as part of the digital revolution. *Zagadnienia Informacji Naukowej–Studia Informacyjne*, 57(1A (113A)): 28-36: DOI: 10.36702/zin.12.

Kothari, C.R. 2004. *Research methodology: methods &* techniques, 2nd edition. New Age International New: Delhi.

Kuhn, T. S. 1962. *The structure of scientific revolutions*. University of Chicago Press: Chicago, IL.

Kumar, R. 2014. *Research methodology: A step-by-step guide for beginners*. Sage Publications Limited: London.

Kvale, L. & Stangeland, E. 2017. Skills for research data management–developing RDM courses at the university of Oslo. *Proceedings of the Association for Information Science and Technology*, 54(1): 728-730. DOI: https://doi.org/10.1002/pra2.2017.14505401134.

Laboutkova, S. 2015. Open government data- a lesson to be learned. *Proceedings of the 12th International Conference Liberec Economic Forum 2015*. DOI: http://doi.org/10.13140/RG.2.1.4395.0808.

Lau, J. 2006. Guidelines on information literacy for lifelong learning. *IFLA, Veracruz*. Available at: https://repository.ifla.org/handle/123456789/193 [2 November 2023].

Leavy, P. (Ed.). 2014. *The Oxford handbook of qualitative research*. Oxford University Press: New York.

Leech, N., & Onwuegbuzie, A. 2009. A typology of mixed methods research designs. *Quality & Quantity: International Journal of Methodology*, 43(2): 265-275. DOI: https://doi.org/10.1007/s11135-007-9105-3.

Leedy, D. & Ormrod J. E. 2020. *Practical Research: Planning and Design, 12th Edition*. Pearson Education: Boston.

Leisyte, I. & Westerheijden, D.F. 2014. Stakeholders and Quality Assurance in Education. In Eggins, H., (ed.). *Drivers and barriers to achieving quality in higher education*,: 83-98, Boston: Sense publishers.

Lewis, M. 2010. Libraries and the management of research data. In S. Mcknight (Ed.), *Envisioning future academic library services: Initiatives, ideas and challenges*: 145–168. Facet: London.

Lin, T.C. 2018. Incorporating social activism. *Boston University Law Review,* 98:1535-1605.

Linde, P., Noorman, M., Wessels, B., & Sveinsdottir, T. 2014. How can libraries and other academic stakeholders engage in making data open? *Information Services & Use* 34(3/4): 211-219. DOI: http://doi.org/10.3233/ ISU-140741.

Liu, X. & Ding, N. 2016. Research data management in universities of central China: practices at Wuhan University Library. *The Electronic Library*. 34(5): 808-822. DOI: http://doi.org/10.1108/EL-04-2015-0063.

Llebot, C. & Castillo, D.J. 2023. Are institutional research data policies in the US supporting the FAIR principles? A content analysis. *Journal of eScience Librarianship*, 12(1): 1-16. DOI: https://doi.org/10.7191/jeslib.614.

Love, N. 2004. Taking data to new depths. *Journal of Staff Development*. 25(4): 22-26. Available at: https://all4ed.org/wp-content/uploads/2013/09/lovejsdarticle.pdf [28 October 2023].

Luce, R.E. 2008. A new value equation challenge: The emergence of eResearch and roles for research libraries. CLIR. Available at: https://www.researchgate.net/profile/Richard-Luce/publication/263261432_A_New_A_New_Value_Equation_Challenge_The_Emergence_of_eResearch_and_Roles_for_Research_Libraries/links/0a85e53ae07e9bcc77000000/A-New-A-New-Value-Equation-Challenge-The-Emergence-of-eResearch-and-Roles-for-Research-Libraries.pdf [20 October 2023].

Lukenchuk, A. 2017. Chapter Five: Methodology: choosing among paradigms and research designs. *Counterpoints,* 428: 57-85. Available at: http://www.jstor.org/stable/45177753 [9 November 2023].

Lund, T. 2012. Combining qualitative and quantitative approaches: Some arguments for mixed methods research. *Scandinavian Journal of Educational Research,* 56(2): 155-165. DOI: 10.1080/00313831.2011.568674.

Majid S., Foo S. & Zhang X. 2018. Research data management by academics and researchers: Perceptions, knowledge and practices. *In: 20th international conference on asia-pacific digital libraries, ICADL 2018* (eds Dobreva M, Hinze A and Zumer M), Hamilton, New Zealand, 19–22 November 2018: 166–178. Cham: Springer. https://doi.org/10.1007/978-3-030-04257-8_16 [26 October 2023].

Makani, J. 2015. Knowledge management, research data management, and university scholarship: Towards an integrated institutional research data management support-

system framework. *VINE*, 45(3): 344-359. DOI: http://doi.org/10.1108/VINE-07-2014-0047.

Mandel, M. 2017. The economic impact of data: why data is not like oil. Available at: https://www.progressivepolicy.org/wp-content/uploads/2017/07/PowerofData-Report_2017.pdf [28 October 2023].

Mandinach, E.B. & Gummer, E.S. 2012. Navigating the landscape of data literacy: it is complex. *WestEd*. Available at: https://files.eric.ed.gov/fulltext/ED582807.pdf [28 October 2023].

Mandinach, E.B. & Gummer, E.S. 2013. A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, 42(1): 30-37. DOI: 10.3102/0013189X12459803.

Mandinach, E.B. 2012. A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, 47(2): 71-85. DOI: https://doi.org/10.1080/00461520.2012.667064.

Mannheimer, S. 2018. Toward a better data management plan: The impact of DMPs on grant funded research practices. *Journal of eScience Librarianship*, 7(3): 1-30. DOI: https://doi.org/10.7191/jeslib.2018.1155

Marques, P., Bernardo, M., Presas, P. & Simon, A. 2019. Corporate social responsibility in a local subsidiary: internal and external stakeholders' power. *EuroMed Journal of Business*, 15(3): 377-393. DOI: 10.1108/EMJB-01-2019-0013.

Martin, E.R. 2016. The role of librarians in data science: a call to action. *Journal of eScience Librarianship*, 4(2). DOI: https://doi.org/10.7191/jeslib.2015.1092.

Masinde, J., Chen, J., Wambiri, D. & Mumo, A. 2021. Research librarians' experiences of research data management activities at an academic library in a developing country. *Data and Information Management*, 5(4): 412-424. DOI: https://doi.org/10.2478/dim-2021-0002.

Maxcy, S. J. 2003. Pragmatic threads in mixed methods research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research*:51-89. Sage: Thousand Oaks, CA.

Maybee, C., & Zilinski, L. 2015. Data informed learning: A next phase data literacy framework for higher education. Proceedings of the Association for Information Science and Technology 52(1): 1–4. DOI: https://doi.org/10.1002/pra2.2015.1450520100108.

Mayer-Schoenberger, V. & K. Cukier. 2013. *Big Data. A Revolution that will transform how we live, work, and think*. John Murray Publishers: London.

Merrill, A. 2011. Library+. *Public Services Quarterly*. 7(3-4):144-148. DOI: 10.1080/15228959.2011.623598.

Mertens, D.M. 2017. Transformative research: Personal and societal. *International Journal for Transformative Research*, 4(1): 18-24. DOI: https://doi.org/10.1515/ijtr-2017-0001.

Mertler, C. A. 2019. *Introduction to educational research, 2nd edition*. SAGE: Los Angles.

Mihas, P. 2019. Qualitative analysis. In Burkholder, G.J., Cox, K.A., Crawford, L.M. and Hitchcock, J.H. (Eds.), *Research design and methods: An applied guide for the scholar-practitioner*. Sage Publications: Los Angeles.

Miles, S. 2017. Stakeholder theory classification: A theoretical and empirical evaluation of definitions. *Journal of Business Ethics*, 142(3): 437-459. DOI: https://doi.org/10.1007/s10551-015-2741-y.

MIT Libraries, n.d. Data management. Available: https://libraries.mit.edu/data-management/services/workshops/ [8 October 2023].

Mitchell, R. K., Agle, B. R., & Wood, D. J. 1997. Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review*, 22(4): 853-886. DOI: ttps://doi.org/10.5465/amr.1997.9711022105.

Mohajan, H.K. 2018. Qualitative research methodology in social sciences and related subjects. *Journal of Economic Development, Environment and People*, 7(1): 23-48. Available at: https://www.ceeol.com/search/article-detail?id=640546 [8 October 2023].

Mojtahedi, M., & Oo, B. L. 2017. The impact of stakeholder attributes on performance of disaster recovery projects: The case of transport infrastructure. *International Journal of Project Management*, 35(5): 841-852. DOI: https://doi.org/10.1016/j.ijproman.2017.02.006.

Monino, J.L. 2021. Data value, big data analytics, and decision-making. *Journal of the Knowledge Economy*, 12(1): 256-267. DOI: https://doi.org/10.1007/s13132-016-0396-2.

Mooney, H. and Newton, M.P. 2012. The anatomy of a data citation: discovery, reuse, and credit. *Journal of Librarianship and Scholarly Communication*, 1(1): 1-14. DOI: https://doi.org/10.7710/2162-3309.1035.

Moore, J., Smith, D., Schultz-Jones, B. & Marino, J. 2019. Data literacy: School librarians as data coaches. DOI: https://doi.org/10.29173/iasl7390.

Moore, R. 2014. Democratizing data: Why data literacy will be the most important new skill of the 21st century. *Venture Beat,* 7 November 2014. Available at: https://venturebeat.com/2014/11/07/democratizing-data-why-data-literacy-will-be-the-most-important-new-skill-of-the-21st-century/ [22 October 2023].

Morales, L.G., Hsu, Y., Poole, J., Rae, B. & Rutherford, I. 2014. A world that counts: mobilising the data revolution for sustainable development. *Report prepared at the request of the United Nations secretary-general, by the independent expert advisory group on a data revolution for sustainable development.* Available at: http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf [20 October 2023].

Moran, G. 2019. We're in a data literacy crisis. Could librarians be the superheroes we need? *Tech: Human Intelligence*. Available at: https://fortune.com/2019/08/31/data-literacy-crisis-librarians-library/ [20 October 2023].

Mortelmans, D. 2019. Analyzing qualitative data using NVivo. In Van den H. Bulck, M. Puppis, K. Donders, & Van L. Audenhove (Eds.), *The Palgrave handbook of methods for media policy research*: 435–450. Cham: Palgrave Macmillan. DOI: https://doi.org/10.1007/978-3-030-16065-4_25.

Mugenda, A.G. & Mugenda, O.M. 2012. *Research methods dictionary*. Arts Press: Nairobi.

Muhammad, S. 2016. *Basic Guidelines for Research: An Introductory Approach for All Disciplines*. Book Zone Publication: Bangladesh.

Munene, I. 2016a. University branch campuses in Kenya to close. *University World News*, Issue No. 176. Available at: http://www.universityworldnews.com/article.php?story=2016070109174334 [25 October 2023].

Munene, I. 2016b. Part-time lecturers are the norm in Kenya. There's a plan to reverse the practice. *The Conversation Africa*, Inc. Available at: http://theconversation.com/part-time-lecturers-are-the-norm-in-kenya-theres-a-plan-to-reverse-the-practice-65977 [25 October 2023].

Muriithi, P., Horner, D. & Pemberton, L. 2016. Factors contributing to adoption and use of information and communication technologies within research collaborations in Kenya. *Information Technology for Development*, 22(1): 84-100. DOI: https://doi.org/10.1080/02681102.2015.1121856.

Mushi, G.E., Pienaar, H. & van Deventer, M. 2020. Identifying and Implementing Relevant Research Data Management Services for the Library at the University of Dodoma, Tanzania. *Data Science Journal*, 19(1): 1-9. DOI: http://doi.org/*10.5334/dsj-2020-001*.

Nagpal, D., Kornerup, I. & Gibson, M.P. 2020. Mixed-method research: A basic understanding. *CODS Journal of Dentistry,* 12(1): 11–16. DOI: 10.5005/jp-journals-10063-0065.

National Commission for Science, Technology and Innovation. 2020. *Strategic plan 2018-2022*. Available at:

https://www.nacosti.go.ke/images/docs/2019/NACOSTI%20STRATEGIC%20PLAN .pdf [22 October 2023].

National Research Fund. 2019. *Background*. https://researchfund.go.ke/background/ [22 October 2023].

Nayek, J.K. & Sen, D. 2015. Data literacy and library: an overview. *West Bengal College Librarians' Association (WBCLA)*, 30(I- II): 38-51.

Ndanu, M. C. & Syombua, M. J. 2015. Mixed methods research: The hidden cracks of the triangulation design. *General Education Journal*, 4 (2): 46-67.

Ng'eno, E. & Mutula, S. 2018. Research Data Management (RDM) in agricultural research institutes: a literature review. *Inkanyiso: Journal of Humanity and Social Sciences*, 10 (1): 28-50. Available at: https://www.ajol.info/index.php/ijhss/article/view/183277 [22 October 2023].

Ngulube, P. 2019. *Handbook of research on connecting research methods for information science research*. PA IGI Global: Hershey.

OECD Organisation for Economic Cooperation and Development. 2004. *Science and innovation policy. Key challenges and opportunities.* OECD: Paris. Available at: https://www.oecd.org/science/inno/23706075.pdf [8 November 2023].

Okamoto, K. 2017. Introducing open government data. *The Reference Librarian*, 58(2): 111-123. DOI: http://doi.org/10.1080/02763877.2016.1199005.

Olakunle, S. A. & Olanrewaju, P. S. 2019. Relationship between information literacy skills and research productivity of researchers in Nigeria, and the mediating role of socio-economic factors. *Libers,* 29(1): 51-76. Available at: https://bpb-us-e1.wpmucdn.com/blogs.ntu.edu.sg/dist/8/644/files/2020/01/LIBRESv29i1p51-76.SimisayePopoola.2019.pdf [29 October 2023].

Olum, G. H. 2013. *Enhancing migration data management in Kenya: assessment and recommendations*. Nairobi: ACP Observatory on Migration.

Ong, M.H.A. & Puteh, F. 2017. Quantitative data analysis: Choosing between SPSS, PLS, and AMOS in social science research. *International Interdisciplinary Journal of Scientific Research*, 3(1): 14-25.Available at: https://www.researchgate.net/profile/Fadilah-Puteh-2/publication/322885790_Quantitative_Data_Analysis_Choosing_Between_SPSS_PLS_and_AMOS_in_Social_Science_Research/links/5a73d1320f7e9b20d490034b/Quantitative-Data-Analysis-Choosing-Between-SPSS-PLS-and-AMOS-in-Social-Science-Research.pdf [22 October 2023].

Onwuegbuzie, A. J. & Combs, J. P. 2011. Data analysis in mixed research: A primer. *International Journal of Education*, 3(1):1-25. Available: https://hdl.handle.net/20.500.11875/2950 [22 October 2023].

Orman, T.F. 2016. Paradigm as a central concept in Thomas Kuhn's thought. *International Journal of Humanities and Social Science*, 6(10): 47-52. Available at: https://www.researchgate.net/profile/Turkan-Firinci-Orman/publication/316550322_Paradigm_as_a_Central_Concept_in_Thomas_Kuhn's_Thought/links/59034205a6fdccd580cfc732/Paradigm-as-a-Central-Concept-in-Thomas-Kuhns-Thought.pdf [22 October 2023].

Palinkas, L.A., Horwitz, S.M., Green, C.A., Wisdom, J.P., Duan, N. & Hoagwood, K. 2015. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health,* 42(5): 533-544. DOI: https://doi.org/10.1007/s10488-013-0528-y.

Palsdottir, A. 2021. Data literacy and management of research data–a prerequisite for the sharing of research data. *Aslib Journal of Information Management*, *73*(2): 322-341. DOI: https://doi.org/10.1108/AJIM-04-2020-0110.

Pantaleo, S. 2017. *Exploring student response to contemporary picturebooks*. University of Toronto Press: Toronto.

Pantzos, P., Gumaelius, L., Buckley, J. & Pears, A. 2020. Considerations in the development of a follow-up exploratory quantitative design for student's motivation regarding to work industry-related activities in higher engineering education. *IEEE Frontiers in Education Conference (FIE)*, Uppsala, Sweden: 1-5. DOI: 10.1109/FIE44824.2020.9274011.

Paoloni, P., Cesaroni, F.M. & Demartini, P. 2018. Relational capital and knowledge transfer in universities. *Business Process Management Journal*. 25(1): 185-201. DOI: 10.1108/BPMJ-06-2017-0155.

Paradis, E., O'Brien, B., Nimmon, L., Bandiera, G. & Martimianakis, M.A. 2016. Design: Selection of data collection methods. *Journal of graduate medical education*, 8(2): 263-264. DOI: https://doi.org/10.4300/JGME-D-16-00098.1.

Parent, M.M. & Deephouse, D.L. 2007. A case study of stakeholder identification and prioritization by managers. *Journal of business ethics*, 75(1):1-23. DOI: 10.1007/s10551-007-9533-y.

Park, J.H., Younas, M., Arabnia, H.R. & Chilamkurti, N. 2021. Emerging ICT applications and services—Big data, IoT, and cloud computing. *International Journal of Communication Systems*, 34(2):1-4. DOI: 10.1002/dac.4668.

Park, S., Bekemeier, B., Flaxman, A. & Schultz, M. 2022. Impact of data visualisation on decision-making and its implications for public health practice: a systematic literature review. *Informatics for Health and Social Care*, 47(2): 175-193. DOI: https://doi.org/10.1080/17538157.2021.1982949.

Park, Y.S., Konge, L. & Artino, A.R. 2020. The positivism paradigm of research. *Academic Medicine*, 95(5): 690-694. DOI: https://doi.org/10.1097/ACM.0000000000003093.

Pasquetto, I. V., Borgman, C. L., & Wofford, M. F. 2019. Uses and Reuses of Scientific Data: The Data Creators' Advantage. *Harvard Data Science Review*, 1(2): 1-34. DOI: 10.1162/99608f92.fc14bf2d.

Patel, D. 2016. Research data management: A conceptual framework. *Library Review*. 65(4-5): 226-241. DOI: 10.1108/LR-01-2016-0001.

Patil, D. J., and H. Mason. 2015. *Data Driven: Creating a Data Culture*. O'Reilly Media Inc.: Sebastopol, CA.

Peck, B. & Mummery, J. 2018. Hermeneutic constructivism: An ontology for qualitative research. *Qualitative health research*, 28(3): 389-407. DOI: https://doi.org/10.1177/1049732317706931

Penev, L., Mietchen, D., Chavan, V., Hagedorn, G., Remsen, D., Smith, V. & Shotton, D. 2011. Pensoft data publishing policies and guidelines for biodiversity data. *Pensoft Publ*. Available at: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5602faa0928e5b19568ab53b648dd6a82e1113d7 [21 October 2023].

Pentland, A.S. 2013. The data-driven society. *Scientific American,* 309(4): 78-83. Available at: https://www.jstor.org/stable/26018109 [21 October 2023].

Perrier, L. & Barnes L. 2018. Developing research data management services and support for researchers: A mixed methods study. *Partnership: The Canadian Journal of Library and Information Practice and Research* 13(1): 4115. DOI: https://doi.org/10.21083/partnership.v13i1.4115.

Perrier, L. and Barnes, L. 2018. Developing research data management services and support for researchers: a mixed methods study. *Partnership: The Canadian Journal of Library and Information Practice and Research*, 13(1): 1-23. DOI: 10.21083/partnership.v13i1.4115.

Pham, H.T. & Tanner, K. 2015. Collaboration between academics and library staff: A structurationist perspective. *Australian Academic & Research Libraries*, 46(1): 2-18. DOI: http://doi.org/10.1080/00048623.2014.989661.

Pham, L. 2018. A review of advantages and disadvantages of three paradigms: positivism, interpretivism and critical inquiry. *Qualitative approach to research*,:1-7. DOI: 10.13140/RG.2.2.13995.54569.

Piracha, H.A. & Ameen, K. 2018. Research data management practices of faculty members. Pakistan Journal of Information Management and Libraries, 20: 60–75. Available at: http://111.68.103.26/journals/index.php/pjiml/article/view/1321/667 [21 October 2023].

Plale, B. & Kouper, I. 2017. The centrality of data: data lifecycle and data pipelines. In M. Chowdhury, A. Apon, and K. Dey, eds. *Data analytics for intelligent transportation*

systems. Amsterdam, Netherlands: Elsevier, 91-111. DOI: http://doi.org/10.1016/B978-0-12-809715-1.00004-3.

Pothier, W.G. & Condon, P.B. 2020. Towards data literacy competencies: Business students, workforce needs, and the role of the librarian. *Journal of Business & Finance Librarianship*, 25(3-4): 123-146. DOI: https://doi.org/10.1080/08963568.2019.1680189.

Prado, J.C. & Marzal, M. A. 2013. Incorporating data literacy into information literacy programs: core competencies and contents. *Libri*, 63(2):123-134. Available at: https://e-archivo.uc3m.es/bitstream/handle/10016/27173/incorporating_calzada_LIBRI_2013.pdf?sequence=1 [28 October 2023].

Press, G. 2018. How Apple, Amazon, Facebook, Google and Microsoft Made 2018 the Year that it mattered a lot. *Forbes Dec 30, 2018.* Available at: *https://www.forbes.com/sites/gilpress/2018/12/30/how-apple-amazon-facebook-google-and-microsoft-made-2018-the-year-that-it-mattered-a-lot/#1a06e85e1cee* [28 October 2023].

Prewitt, K. 1999. Of Significance... the statistical literacy challenge: an introduction. *Of Significance: A Topical Journal of the Association of Public Data Users*, 1(1): 1-2.

Proctor, L., Wartho, R. & Anderson, M. 2005. Embedding information literacy in the sociology program at the University of Otago. *Australian Academic & Research Libraries*. 36(4):153-168. DOI: http://doi.org/10.1080/00048623.2005.10755307.

Project Management Institute (PMI). 1996. *A guide to the project management body of knowledge* (PMBOK Guide). Project Management Institute: Newtown Square, PA.

Pryor, G. 2014. Who's doing data? A spectrum of roles, responsibilities, and competencies. In G. Pryor, S. Jones and A. Whyte (Eds.), *Delivering research data management services: Fundamentals of good practice*: 41-58. Facet: London.

Qi, Z. 2018. Research on scientific data literacy education system. *Open journal of social sciences*, 6(6): 187-199. DOI: https://doi.org/10.4236/jss.2018.66017.

Qin, J. & D'Ignazio, J. 2010a. Lessons learned from a two-year experience in science data literacy education. In *Proceedings of the 31st annual IATUL conference*: 20-24 June 2010, 2. Available at: http://docs.lib.purdue.edu/iatul2010/conf/day2/5 [28 October 2023].

Qin, J. & D'ignazio, J. 2010b. The central role of metadata in a science data literacy course. *Journal of Library Metadata*, 10(2-3): 188-204. DOI: https://doi.org/10.1080/19386389.2010.506379.

QuestionPro. 2019. Data Analysis in Research: Types & Methods. Available at: https://www.questionpro.com/blog/data-analysis-in-

research/#:~:text=The%20motive%20behind%20data%20analysis,mining%2C%20or
%20developing%20graphical%20representation [24 October 2023].

Raffaghelli, J.E. & Manca, S. 2023. Exploring the social activity of open research data on
ResearchGate: Implications for the data literacy of researchers. *Online Information
Review*, 47(1): 197-217. DOI: https://doi.org/10.1108/OIR-05-2021-0255.

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

Rahimi, M., & Weisi, H. 2018a. The impact of research practice on professional teaching
practice: Exploring EFL teachers' perception. Cogent Education, 5(1):1-15. DOI:
http://doi.org/10.1080/2331186X.2018.1480340 [24 October 2023].

Rahimi, M., & Weisi, H. 2018b. Reflective practice, self-efficacy and research practice of
EFL teachers: Examining possible relationships. *Issues in Educational Research*,
28(3): 756-780. Available at:
https://search.informit.org/doi/10.3316/ielapa.864357661266441 [2 November 2023].

Rahlf, T. 2019. *Data Visualisation with R: 111 Examples*. Springer Nature: Switzerland.

Ranaweera, P. 2008. Importance of information literacy skills for an information literate
society. Available at: https://core.ac.uk/download/pdf/11884153.pdf [5 November
2023].

Rasul, A. & Singh, D. 2010. The role of academic libraries in facilitating postgraduate
students' research. *Malaysian Journal of Library & Information Science*. 15(3): 75-
84. Available at: https://www.researchgate.net/publication/228839574 [24 October
2023].

Ravitch, S. M. & Riggan, M. 2017. *How conceptual frameworks guide research, 2nd edition*.
Sage: Los Angeles, CA.

Rehman, A. A. & Alharthi, K. 2016. An Introduction to Research Paradigms. *International
Journal of Educational Investigations*. 3(8): 51-59. Available at:
http://www.ijeionline.com/attachments/article/57/IJEI.Vol.3.No.8.05.pdf [5
November 2023].

Reitz, J. M. 2004. *Dictionary for library and information science*. Libraries Unlimited:
London.

Research Data Alliance. 2020. FAIR Data Maturity Model: specification and guidelines. *Res.
Data Alliance* DOI: 10.15497/rda00045.

Rice, R. & Southall, J. 2016. *The data librarian's handbook*. Facet publishing: London.

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D.,
Matwin, S. & Wuetherick, B. 2015. Strategies and best practices for data literacy

education: Knowledge synthesis report. Available at:
https://dalspace.library.dal.ca/bitstream/handle/10222/64578/Strategies%20and%20Best%20Practices%20for%20Data%20Literacy%20Education.pdf?sequence=1 [4 November 2023].

Robelo, O.G. and Bucheli, M.G.V. 2018. Comparative analysis of research skills and ICT: A case study in higher education. *International Journal of Educational Excellence*, 4(1): 15-27. DOI: 10.18562/IJEE.031.

Robertson, A., McMurray, I., Ingram, J. & Roberts, P.I. 2012. Embedding a curriculum-based information literacy programme at the University of Bedfordshire. *Journal of pedagogic development*, 2(1): 4-10. Available at: http://hdl.handle.net/10547/336260 [9 November 2023].

Rowley, J. 2002. Using case studies in research. *Management research news*, 25(1): 16-27. DOI: https://doi.org/10.1108/01409170210782990.

Rubel, A. 2014. Libraries, electronic resources, and privacy: The case for positive intellectual freedom. *The Library Quarterly*, 84(2): 183-208. DOI: 10.1086/675331.

Rudo, P. 2013. 6 important stages in the data processing cycle. Available at: https://medium.com/@peerxp/the-6-stages-of-data-processing-cycle-3c2927c466ff [4 November 2023].

Sabzwari, M. N., Bhatti, R. & Ahmed, B. 2012. ICT skills and computer self-efficacy of research students: The case of institute of pure & applied biology and biotechnology, Bahauddin Zakariya University, Multan, Pakistan. *Library Philosophy and Practice*, 1: 1-15. Available at: https://digitalcommons.unl.edu/libphilprac/844/ [9 November 2023].

Safdar, M., Rehman, S.U., Arif, M. and Ashiq, M., 2023. Research data services in libraries: a systematic literature review. *Information Discovery and Delivery*, 51(2): 151-165. DOI: https://doi.org/10.1108/IDD-04-2021-0044.

Sample, A. 2020. Historical development of definitions of information literacy: A literature review of selected resources. *The journal of academic librarianship*, *46*(2), 102116. DOI: https://doi.org/10.1016/j.acalib.2020.102116.

Sayogo, D. S., & Pardo, T. A. 2013. Exploring the determinants of scientific data sharing: Understanding the motivation to publish research data. *Government information quarterly*, *30*: S19-S31. DOI: 10.1016/j.giq.2012.06.011.

Scherbaum, C. & Shockley, K. 2015. Introduction. In: *Analysing Quantitative Data for Business and Management Students*, Mastering Business Research Methods. 1st ed. 55 City Road, London: SAGE Publications, Inc.:1-8. DOI: https://doi.org/10.4135/9781529716719.

Schield, M. 2005. Information literacy, statistical literacy, data literacy. *IASSIST quarterly*, 28(2-3): 6-11. Available at: https://iassistquarterly.com/public/pdfs/iqvol282_3shields.pdf [9 November 2023].

Schmidt, B., Gemeinholzer, B., & Treloar, A. 2016. Open data in global environmental research: The Belmont Forum's open data survey. *PLoS ONE*, 11(1), e0146695: 1-29. DOI: http://doi.org/10.1371/journal.pone.0146695.

Schultz-Jones, B., Moore, J.E. & Marino, J. 2019. *Data Literacy Leadership Preparation for School Librarians*. Paper presented at: IFLA WLIC 2019 - Athens, Greece - Libraries: dialogue for change in Session 190 - School Libraries. Available at: https://library.ifla.org/id/eprint/2545/1/190-schultz-jones-en.pdf [6 November 2023].

Searle, S., Wolski, M., Simons, N. & Richardson, J. 2015. Librarians as partners in research data service development at Griffith University. *Program: electronic library and information systems*. 49(4): 440-460. DOI: http://doi.org/10.1108/PROG-02-2015-0013.

Semeler, A.R., Pinto, A.L. & Rozados, H.B.F. 2019. Data science in data librarianship: Core competencies of a data librarian. *Journal of Librarianship and Information Science*, 51(3): 771-780. DOI: https://doi.org/10.1177/0961000617742.

Serap K.S. 2003. Self-efficacy: a concept closely linked to information literacy and lifelong learning. *Journal of Documentation*. 59(6): 635-646. DOI: https://doi.org/10.1108/00220410310506295.

Sewell, C., & Kingsley, D. 2017. Developing the 21st century academic librarian: The Research Support Ambassador Programme. *New Review of Academic Librarianship*, 23(2-3): 148–158. DOI:10.1080/13614533.2017.1323766.

Sewerin, C. 2015. Research data management faculty practices: A Canadian perspective. *Proceedings of the IATUL Conferences. Paper 2*. Available at: https://docs.lib.purdue.edu/iatul/2015/mrd/2 [3 November 2023].

Shah, S.R. & Al-Bargi, A. 2013. Research paradigms: researchers' worldviews, theoretical frameworks and study designs. *Arab World English Journal*, 4(4): 252 -264. Available at: https://bpb-us-w2.wpmucdn.com/sites.gsu.edu/dist/7/3504/files/2016/12/Shah-2013-Research-paradigms-Researchers-worldviews-theoretical-frameworks-and-study-designs-1duh1ty.pdf [23 October 2023].

Sharma, G. 2017. Pros and cons of different sampling techniques. *International journal of applied research*, 3(7): 749-752.

Sharma, S. & Qin, J. 2014. Data management: Graduate student's awareness of practices and policies. *Proceedings of the Association for Information Science & Technology* 51(1):1-3. DOI: 10.1002/meet.2014.14505101130.

Sharma, S. 2017. Definitions and models of statistical literacy: a literature review. *Open Review of Educational Research*, 4(1): 118-133. DOI: http://doi.org/10.1080/23265507.2017.1354313.

Shelly, M. & Jackson, M. 2018. Research data management compliance: is there a bigger role for university libraries? *Journal of the Australian Library and Information Association*, 67(4): 394-410. DOI: https://doi.org/10.1080/24750158.2018.1536690.

Shin, A.Y.M. D. 2020. Paradigm Shift. *Techniques in Hand & Upper Extremity Surgery.* 24 (2): 53-54. DOI: 10.1097/BTH.0000000000000285.

Shorten, A. & Smith, J. 2017. Mixed methods research: expanding the evidence base. *Evidence Based Nursing*, 20(3): 74-75. DOI: 10.1136/eb-2017-102699.

Si, L., Xing, W.M., Zhuang, X.Z., Hua, X.Q., & Zhou, L.M. 2015. Investigation and analysis of research data services in university libraries. Electronic Library, (33): 417–449. DOI: http://doi.org/10.1108/EL-07-2013-0130.

Simons, N. & Searle, S. 2014. Redefining 'the librarian' in the context of emerging eResearch services. *VALA2014 Proceedings*: 3-6. Available at: http://hdl.handle.net/10072/63599 [23 October 2023].

Sinaeepourfard, A., Masip-Bruin, X., Garcia, J. & Marín-Tordera, E. 2015. A survey on data lifecycle models: Discussions toward the 6vs challenges. *Technical Report* (UPC-DAC-RR-2015–18). Available at: https://www.ac.upc.edu/app/research-reports/html/RR/2015/18.pdf [23 October 2023].

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. 2017. Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research,* 70: 263-286. DOI: https://doi.org/10.1016/j.jbusres.2016.08.001.

Smith, K. 2008. Embedding enterprise education into the curriculum at a research-led university. *Education + Training*, 50(8/9): 713–724. DOI: http://dx.doi.org/10.1108/00400910810917082.

Solmaz, D.Y. 2017. Relationship between lifelong learning levels and information literacy skills in teacher candidates. *Universal Journal of Educational Research*. 5(6): 939-946. DOI: http://doi.org/10.13189/ujer.2017.050605.

Someh, I., Davern, M., Breidbach, C.F. & Shanks, G. 2019. Ethical issues in big data analytics: A stakeholder perspective. *Communications of the Association for Information Systems*, 44(1): 718-747 DOI: 10.17705/1CAIS.04434.

Stephenson, E. & Caravello, S. P. 2007. Incorporating data literacy into undergraduate information literacy programs in the social sciences: A pilot project. *Reference Services Review*. 35(4): 525-540. DOI: 10.1108/00907320710838354.

Stucki, T. & Wochner, D. 2019. Technological and organizational capital: Where complementarities exist. *Journal of Economics and Management Strategy*, 28(3): 458-487. DOI: 10.1111/jems.12269.

Suchman, M.C. 1995. Managing legitimacy: Strategic and institutional approaches. *Academy of management review*, 20(3): 571-610. DOI: 10.2307/258788.

SurveyMonkey. 2021. Survey sample size calculator. Available at: https://www.surveymonkey.com/mp/sample-size-calculator/ [23 October 2023].

Swan, A. & Brown, S. 2008. *The skills, role and career structure of data scientists and curators: an assessment of current practice and future needs.* Truro: Key Perspectives. Available at: http://www.jisc.ac.uk/publications/documents/dataskillscareersfinalreport.aspx [30 October 2023].

Swart, J. 2006. Intellectual capital: disentangling an enigmatic concept. *Journal of Intellectual capital*, 7(2): 136-159. DOI 10.1108/14691930610661827.

Taherdoost, H. 2016. Sampling methods in research methodology; how to choose a sampling technique for research. How to Choose a Sampling Technique for Research: *International Journal of Academic Research in Management (IJARM)* 5(2): 18-27. DOI: http://dx.doi.org/10.2139/ssrn.3205035.

Tait, E., Martzoukou, K. & Reid, P. 2016. Libraries for the future: the role of IT utilities in the transformation of academic libraries. *Palgrave Communications*, 2(1):1-9. DOI: https://doi.org/10.1057/palcomms.2016.70.

Talley, P.C. 2017. Preparing Students for the Globalized Workforce: A Mission Statement for Taiwanese Universities. *International Journal of Humanities and Social Science,* 7(3): 57-60.

Tammaro, A. M., Matusiak, K.K., Sposito, F. A., Pervan, A., & Casarosa, V. 2016. Understanding roles and responsibilities of data curators: An international perspective. *Libellarium* 9(2): 39-47. DOI: https://doi.org/10.15291/libellarium.v9i2.286.

Tang, R. & Hu, Z. 2019. Providing research data management (RDM) services in libraries: Preparedness, roles, challenges, and training for RDM practice. *Data and information management*, 3(2): 84-101. DOI: https://doi.org/10.2478/dim-2019-0009.

Tansley, S. & Tolle K. 2009. Jim Gray on eScience: A Transformed Scientific Method. In Hey, T, Tansley, S. and Tolle, K, (Eds), *The fourth paradigm: data-intensive scientific discovery*. USA: Microsoft Corporation: xvii-xxxi.

Taylor, P.C. & Medina, M. 2011. Educational research paradigms: From positivism to pluralism. *College Research Journal*, 1(1): 1-16. Available at: https://researchportal.murdoch.edu.au/esploro/outputs/journalArticle/Educational-research-paradigms-From-positivism-to/991005540537707891 [27 October 2023].

Taylor, S.P. 2018. Critical realism vs social constructionism & social constructivism: application to a social housing research study. *International Journal of Sciences: Basic and Applied Research*, 37(2): 216-222. Available at:

https://insight.cumbria.ac.uk/id/eprint/3596/1/8701-25730-1-PB.pdf [27 October 2023].

Tchamyou, V.S. 2019. The role of information sharing in modulating the effect of financial access on inequality. *Journal of African Business*, 20(3): 317-338. DOI: https://doi.org/10.1080/15228916.2019.1584262.

Tenopir, C., Dalton, E.D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D. & Dorsett, K. 2015. Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PloS one*, 10(8), p.e0134826. DOI: https://doi.org/10.1371/journal.pone.0134826.

Tenopir, C., Hughes, D., Allard, S., Frame, M., Birch, B., Baird, L., Sandusky, R., Langseth, M. & Lundeen, A. 2015. Research Data Services in Academic Libraries: Data Intensive Roles for the Future?. *Journal of eScience Librarianship*, 4(2):1-21. DOI: https://doi.org/10.7191/jeslib.2015.1085.

Tenopir, C., Sandusky, R.J., Allard, S. & Birch, B. 2014. Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*. 36(2): 84-90. DOI: http://doi.org/10.1016/j.lisr.2013.11.003.

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

Thomas, G. 2014. *A critical analysis of the use of community intelligence in local neighbourhood policing in south Wales*. University of South Wales.

Tibor, K. 2019. Identifying new roles for academic libraries in supporting data-intensive research. *Библиосфера*, (4): 97-102. DOI: 10.20913/1815-3186-2019-4-97-102.

Torres. Z., I., 2019. University research centres: Organizational structures and performance. *Journal of technology management & innovation*, 14(3): 23-43. DOI: http://dx.doi.org/10.4067/S0718-27242019000300023.

Tripathi, M, Shukla A. & Sonker S.K. 2017. Research data management practices in university libraries: A study. *DESIDOC Journal of Library & Information Technology* 37(6): 417-424. DOI: https://doi.org/10.14429/djlit.37.11336.

Turner, S. F., Cardinal, L. B. & Burton, R. M. 2017. Research Design for Mixed Methods: A Triangulation-based Framework and Roadmap. *Organizational Research Methods*, 20(2): 243–267. DOI: 10.1177/1094428115610808.

Unal, Y., Chowdhury, G., Kurbanoğlu, S., Boustany, J. & Walton, G. 2019. Research data management and data sharing behaviour of university researchers'. *Information Research: an international electronic journal*, 24(1): 1-29. Available at: https://e-space.mmu.ac.uk/id/eprint/623415 [22 October 2023].

United Nations Economic Commission for Europe. 2012. *Making Data Meaningful, Part 4: a guide to improving statistical literacy*. Geneva: UNECE. Available at: http://www.unece.org/fileadmin/DAM/stats/documents/writing/Making_Data_Meaningful_Part_4_for_Web.pdf [27 October 2023].

United States Office of Science and Technology Policy. 2012. Obama Administration Unveils 'Big Data' Initiative: Announces $200 Million in New R&D Investments. Available at: https://obamawhitehouse.archives.gov/the-press-office/2015/11/19/release-obama-administration-unveils-big-data-initiative-announces-200 [22 October 2023].

Urbano, F., Cagnacci, F., & Initiative, E. C. 2021. Data management and sharing for collaborative science: Lessons learnt from the Euromammals Initiative. *Frontiers in Ecology and Evolution,* 9, 727023. DOI: https://doi.org/10.3389/fevo.2021.727023.

Uwe, F. 2022. *The SAGE Handbook of Qualitative Research Design*. Sage Publications: London.

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

van Dijck, J. 2014. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance and Society*, 12(2):197-208. DOI: https://doi.org/10.24908/ss.v12i2.4776.

van Reisen, M., Stokmans, M., Mawere, M., Basajja, M., Ong'ayo, A. O., Nakazibwe, P., ... & Chindoza, K. 2020. FAIR practices in Africa. *Data Intelligence*, 2(1-2): 246-256. DOI: https://doi.org/10.1162/dint_a_00047.

van Wyk, J. 2017. Research Data Management at UP and the role of NeDICC. Available at: http://hdl.handle.net/2263/63458 [11 November 2023].

Varpio, L., Paradis, E., Uijtdehaage, S. and Young, M., 2020. The distinctions between theory, theoretical framework, and conceptual framework. *Academic Medicine*, 95(7): 989-994. DOI: https://doi.org/10.1097/ACM.0000000000003075.

Venkatesh, V., Brown, S.A. & Bala, H. 2013. Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS quarterly*, 1(37): 21-54. Available at: https://www.jstor.org/stable/43825936 [4 November 2023].

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

Vickers, A.J. 2011. Making raw data more widely available. *BMJ*, *342*. d2323. DOI:10.1136/bmj.d2323.

Wald, N. & Harland, T. 2017. A framework for authenticity in designing a research-based curriculum. Teaching in Higher Education, 22(7): 751–765. DOI: http://doi.org/10.1080/13562517.2017.1289509.

Wallis, J.C., Rolando, E. & Borgman, C.L. 2013. If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PloS one*. 8(7):1-17. DOI: http://doi.org/10.1371/journal.pone.0067332.

Wanner, A. 2015. Data literacy instruction in academic libraries: best practices for librarians. *See Also*. 1(1). DOI: https://doi.org/10.14288/sa.v1i1.186335.

Weng'ua, F.N., Rotich, D.C. & Kogos, E.J. 2017. The role of Kenyan universities in promoting research and scholarly publishing. *South African Journal of Libraries and Information Science*, 83(2): 23-29. Available at: https://hdl.handle.net/10520/EJC-e45b7c838 [7 November 2023].

White, B. & Cossham, A. 2017. Partnerships or parallel lines? The contributions of practitioners and academics to library and information research. *Information Research*, 22(4). Available at: https://informationr.net/ir/22-4/rails/rails1605.html [27 October 2023].

Wiley, C. & Mischo, W.H. 2016. Data management practices and perspectives of atmospheric scientists and engineering faculty. *Issues in Science and Technology Librarianship,* 85. DOI: https://doi.org/10.29173/istl1688.

Wiley, C.A. & Kerby, E.E. 2018. Managing research data: graduate student and postdoctoral researcher perspectives. *Issues in Science and Technology Librarianship,* 89. DOI:10.5062/F4FN14FJ.

Wiljes, C. & Cimiano, P. 2019. Teaching research data management for students. *Data Science Journal*, 18(38):1-9. DOI: http://doi.org/10.5334/dsj-2019-038.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E. & Bouwman, J. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1): 1-9. DOI: https://doi.org/10.1038/sdata.2016.18.

Williams, G. & Pigeot, I. 2017. Consent and confidentiality in the light of recent demands for data sharing. *Biometrical journal*, 59(2): 240-250. DOI: https://doi.org/10.1002/bimj.201500044.

Williams, M. & Vogt, W. P. 2011. *The SAGE handbook of innovation in social research methods*. SAGE Publications: London. DOI: https://doi.org/10.4135/9781446268261.

Williams, M., Bagwell, J. & Zozus, M.N. 2017. Data management plans: the missing perspective. *Journal of biomedical informatics*, 71: 130-142. DOI: https://doi.org/10.1016/j.jbi.2017.05.004.

Willis, J. W. 2007. *Foundations of qualitative research: Interpretive and critical approaches*. Thousand Oaks, CA: Sage Publications.

Wilson, J.A. & Jeffreys, P. 2013. Towards a unified university infrastructure: The data management roll-out at the University of Oxford. *International Journal of Digital Curation*. 8(2): 235-246. DOI: https://doi.org/10.2218/ijdc.v8i2.287.

Wisdom, J. & Creswell, J. W. 2013. Mixed methods: integrating quantitative and qualitative data collection and analysis while studying patient-centred medical home models. *Rockville: Agency for Healthcare Research and Quality. AHRQ Publication*. 13: 0028-EF. Available at: https://www.ahrq.gov/sites/default/files/wysiwyg/ncepcr/tools/PCMH/mixed-methods.pdf [29 October 2023].

Wissik, T. & Durco, M. 2016. Research data workflows: from research data lifecycle models to institutional solutions. In *Selected Papers from the CLARIN Annual Conference 2015, October 14–16, 2015, Wroclaw, Poland*, 123: 94-107. Linköping University Electronic Press. Available at: https://ep.liu.se/ecp/123/008/ecp15123008.pdf [21 October 2023].

Wolff, A., Gooch, D., Cavero, M. J.J, Rashid, U. & Kortuem, G. 2016. Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics.* 12(3): 9-26. DOI: https://doi.org/10.15353/joci.v12i3.3275.

Wong, G. K. W. 2010. Facilitating students' intellectual growth in information literacy teaching. *Reference & User Services Quarterly*, 50(2): 114-118. https://www.jstor.org/stable/20865377 [21 October 2023].

Wood, D.J., Mitchell, R.K., Agle, B.R. & Bryan, L.M. 2018. Stakeholder identification and salience after twenty years: progress, problems, and prospects. *Business and Society*. DOI: 10.1177/0007650318816522.

Xu, Q., 2023. Investigating and understanding library data services to support college student data literacy competencies: A conceptual framework. *Journal of Librarianship and Information Science*. 09610006231196604. DOI: https://doi.org/10.1177/09610006231196604.

Xu, Z., Zhou, X., Kogut, A. & Clough, M. 2022. Effect of online research data management instruction on graduate students' RDM skills. *Library & Information Science Research,*44(4): 101-190. DOI: https://doi.org/10.1016/j.lisr.2022.101190.

Yang, N. & Li, T. 2020. How stakeholders' data literacy contributes to student success in higher education: a goal-oriented analysis. *International Journal of Educational Technology in Higher Education*, 17(1): 41. DOI: https://doi.org/10.1186/s41239-020-00220-3.

Yang, R. J., Wang, Y. & Jin, X.H. 2014. Stakeholders' attributes, behaviours, and decision-making strategies in construction projects: importance and correlations in practice. *Project Management Journal*, 45(3): 74-90. DOI: 10.1002/pmj.21412.

Yeasmin, S. & Rahman, K.F. 2012. Triangulation research method as the tool of social science research. *BUP journal*, 1(1): 154-163.

Younas, M. 2019. Research challenges of big data. *Service Oriented Computing and Applications*, 13:105–107. DOI: 10.1007/s11761-019-00265-x.

Yu, X. & Khazanchi, D. 2017. Using embedded mixed methods in studying is phenomena: Risks and practical remedies with an illustration. *Communications of the Association for Information Systems*, 41: 18-42. DOI: https://doi.org/10.17705/1CAIS.04102.

Yusof, I.J. 2021. Assessing statistical literacy level of postgraduate education research students in Malaysian research universities. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(5): 1318-1324. DOI: https://doi.org/10.17762/turcomat.v12i5.1800.

Zamawe, F.C. 2015. The implication of using NVivo software in qualitative data analysis: Evidence-based reflections. *Malawi Medical Journal*, 27(1): 13-15. https://www.ajol.info/index.php/mmj/article/view/116229 [25 October 2023].

Zamorski, B. 2002. Research-led teaching and learning in higher education: A case. *Teaching in Higher Education*, 7(4): 411-427. DOI: http://doi.org/10.1080/135625102760553919.

Zegwaard, K. E., Campbell, M., & Pretti, T. J. 2017. Professional identities and ethics: The role of work-integrated learning in developing agentic professionals. In T. Bowen and M. T. B. Drysdale (Eds.), *Work-integrated learning in the 21st century: Global perspectives on the future*: 145-160. Emerald Publishing Limited: Bingley, UK. DOI: https://doi.org/10.1108/S1479-367920170000032009.

Zimmerman, A.S. 2008. New knowledge from old data: The role of standards in the sharing and reuse of ecological data. *Science, Technology, & Human Values*. 33(5): 631-652. DOI: http://doi.org/10.1177/0162243907306704.

Zohrabi, M. 2013. Mixed method research: Instruments, validity, reliability and reporting findings. *Theory & practice in language studies*, 3(2): 254-262. DOI: 10.4304/tpls.3.2.254-262.

APPENDICES

Appendix I: Online Questionnaire for Researchers (PhD Students and Faculty)

You are kindly invited to participate in this research which aims to inform the development of data literacy support services in academic libraries in Kenya. This invitation is based on your capacity as a researcher (Faculty/Postgraduate student). Your participation will require that you complete this questionnaire. You will also be required to express your opinion regarding a proposed data literacy framework.

As a participant, you are assured that there are no known risks if you accept to participate, nor are there any cost implications that you will incur for participating in the study. The information you provide will help to inform the development of data literacy support services framework to be used in academic libraries in Kenya. The title of the study is: *The feasibility of offering standardised data literacy services at selected private university libraries in Kenya*.

If you have any questions about this study, you may contact me (the PI) or my supervisors through the contact detail provided below:

Principal Investigator	**Supervisor**	**Co-Supervisor**
Agava Stanislaus	Dr MJ van Deventer	Prof JTD Bothma
agavastanislaus1759@gmail.com	Vandeventer.martha@up.ac.za	Theo.Bothma@up.ac.za
Phone: +254722935932	+2782 924 6650	

Department: Information Science, University of Pretoria

Your anonymity as a respondent is guaranteed. You are not required to share any information that could be traced back to you. From the data collected, no one will be able to identify you. All data collected will be reported in aggregated format. Furthermore, no information provided when completing the questionnaire will in any way affect or influence your present or future status as a researcher.

Your participation in this study is voluntary. You have the right to withdraw your approval to participate at any moment after consenting and without giving a reason. Withdrawing from this study will have no impact on your relationship with the researcher, neither will it attract any penalty.

Filling the questionnaire will take a time range of between 30-40 minutes, depending on the depth of your responses.

Consent

I have read and I do understand the provided information about this study. I understand that my participation in this study is voluntary and that I am free to withdraw at any time, without giving a reason and without any penalty.

If you voluntarily agree to participate, click here:	YES	Continue to questionnaire
If you do not wish to participate, please click here:	NO	Link to: Thank you very much for your time and support.

Section One: Biographical Information

1. a) Please indicate the capacity in which you are completing this questionnaire:

PhD student/Candidate	[ ]
Faculty member	[ ]

Section Two: Research Data Management (RDM)

Definition: RDM entails the (a) creation, (b) processing, (c) analysis, (organisation), the (d) preservation (storage), and publication of research data in order to improve its sharing, (e) accessibility, and (f) re-use.

2. The following indicates procedures in the research data management cycle. Please select only the areas within which the library has assisted you, or has provided you with advisory services (select those that apply using "X")

   a. The library has assisted me in data creation in the following ways:

Preparing a data management plan	[ ]
Selecting data formats	[ ]
Planning consent for data sharing	[ ]
Locating existing data	[ ]
Collecting data	[ ]
Data capturing	[ ]

© University of Pretoria

Other	[ ]
(Please mention below)	
None of the above	[ ]

b. The library has assisted me in data processing in these aspects:

Data entry	[ ]
Data translation	[ ]
Data transcription	[ ]
Data validation	[ ]
Data cleaning	[ ]
Data anonymisation	[ ]
Data description	[ ]
Other (Please mention below)	[ ]
None of the above	[ ]

c. The library has assisted me in data analysis in these aspects:

Data interpretation	[ ]
Production of research output	[ ]
Author publications	[ ]
Data visualisation	[ ]
Preparation of data documentation	[ ]
Other (Please mention below)	[ ]
None of the above	[ ]

d. The library has assisted me in the preservation of data in:

Migrating data to appropriate formats	[ ]
Data back-up and storage	[ ]
Metadata creation	[ ]
Data archiving in an open access repository	[ ]
Data publication in the University's institutional repository	[ ]
Data publication in disciplinary repository	[ ]
Other (Please mention below)	[ ]
None of the above	[ ]

e. The library has assisted me in enhancing access to my data through:

Data sharing	[ ]
Data control	[ ]
Copyright establishment	[ ]
Data promotion	[ ]
Creating a reference for citing my data	[ ]
Other (Please mention below)	[ ]
None of the above	[ ]

3. Have you ever created a data management plan (DMP) (a document that provides details on how your data will be stored, accessed, secured, shared, re-used) for your research?

Yes	[ ]
No	[ ]
Not sure	[ ]

Please comment on your choice

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

4. Have you ever created any metadata (descriptors such as author, subject etc.) in your research process?

Yes	[ ]
No	[ ]
Not sure	[ ]

Please comment on your choice

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

5. Where do you normally store your research data? (Select all of those that apply)

Library server	[ ]
PC hard drive	[ ]
USB stick	[ ]
Portable hard drive	[ ]
My laptop	[ ]
Cloud storage	[ ]
External servers	[ ]
Other (specify)	[ ]

Section Three: Data management competence

6. Rate your competence (knowledge/skills/abilities) in the following areas?

	No Competence	Little Competence	Somewhat Competent	Very Competent	Not Applicable
Data planning	[ ]	[ ]	[ ]	[ ]	[ ]
Data collection	[ ]	[ ]	[ ]	[ ]	[ ]
Data processing	[ ]	[ ]	[ ]	[ ]	[ ]
Data analysis	[ ]	[ ]	[ ]	[ ]	[ ]
Data preservation	[ ]	[ ]	[ ]	[ ]	[ ]
Data sharing/publishing	[ ]	[ ]	[ ]	[ ]	[ ]
Data re-use	[ ]	[ ]	[ ]	[ ]	[ ]
FAIR data	[ ]	[ ]	[ ]	[ ]	[ ]
Ethical collection of data	[ ]	[ ]	[ ]	[ ]	[ ]

Section Four: Research process

7. The following statements indicate ways through which the library can offer data literacy-related support to researchers throughout the research process. Please select the option that is the most appropriate answer in relation to the statement. The options are 'Agree, 'Disagree' and 'Not sure'

	Statements	Agree	Disagree	Not sure
Managing data				
a.	My library has guides on research data management	[ ]	[ ]	[ ]
b.	My library has made available training modules on research data management	[ ]	[ ]	[ ]
c.	My library hosts data management training events.	[ ]	[ ]	[ ]
Data publishing and sharing				

		[ ]	[ ]	[ ]
a.	I am aware of open data as a publishing option	[ ]	[ ]	[ ]
b.	The library provides guides on data publishing	[ ]	[ ]	[ ]
c.	The library provides advice on copyright issues related to data publishing	[ ]	[ ]	[ ]

Tools

a.	The library provides data analysis tools	[ ]	[ ]	[ ]
b.	The library provides tools for online collaboration among researchers	[ ]	[ ]	[ ]
c.	I have access to 'ready to use' Data Management Plans	[ ]	[ ]	[ ]

Section Five: Importance of data literacy

(*Data literacy is the ability to read, understand, create, and communicate data. It focuses on the competencies involved in working with data*.)

8. What, in your opinion, is the importance of data literacy?
....
....
……………………………………………………………………………………………………

9. Indicate your level of agreement with the following statements.

Statement	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I am familiar with open data requirements	[ ]	[ ]	[ ]	[ ]	[ ]
I am familiar with FAIR data	[ ]	[ ]	[ ]	[ ]	[ ]

I am comfortable to share my research data with others	[ ]	[ ]	[ ]	[ ]	[ ]
I am willing to share my research data with others	[ ]	[ ]	[ ]	[ ]	[ ]
I would like to store my research datasets beyond the lifetime of the project	[ ]	[ ]	[ ]	[ ]	[ ]
Academic libraries should have templates for Data Management Plans	[ ]	[ ]	[ ]	[ ]	[ ]
I understand that participants should be made aware that data collected from them will be shared with others	[ ]	[ ]	[ ]	[ ]	[ ]
Academic libraries should provide a template with an embedded metadata set for uploading data into a repository	[ ]	[ ]	[ ]	[ ]	[ ]

Section Six: Research data management challenges

10. I face the following challenges while working with data during research (select all that apply)

Challenges	
Developing an appropriate data management plan	[ ]
Developing data collection instruments	[ ]
Processing collected data	[ ]

349

Analysing data using various statistical software	[ ]
Storing data	[ ]
Preserving data	[ ]
Creating metadata	[ ]
Locating datasets	[ ]
Privacy and confidentiality issues associated with data	[ ]
Others (please specify all that are applicable in the textbox below)	[ ]

Section Seven: Data literacy framework

In view of the framework for successful data literacy training (Refer back to the document that was send to you by email):

11.   Are you familiar with any of the listed literacies (select all that apply)?

Information literacy	[ ]
Statistical literacy	[ ]
Digital literacy	[ ]
None of these literacies	[ ]

12. What are some of the key areas in research data management that you think should be prioritised in the development of a data literacy training program?

…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
……………………

13. Do you think the library has a role to play in promoting data literacy among researchers?

Yes	[ ]
No	[ ]

350

Not sure	[ ]

Please comment on your choice

……

……

……

14. What is it that does not make sense to you when you look at the proposed framework for successful data literacy training?

```



```

15. What do you like about the framework?

```



```

16. What gaps could you identify in the framework?

```



```

17. Given a chance, what would you exclude from the framework?

```



```

18. Kindly give any other comment you have about the attached data literacy framework.

```



```

Appendix II: Interview Guide for University Librarians

Thank you for agreeing to participate in this research which aims to inform the development of data literacy support services in academic libraries in Kenya.

As a reminder: This invitation is based on your capacity as the university librarian. Your participation will require that you respond to the interview questions in this schedule. As a participant, you are assured that there are no known risks if you accept to participate, nor are there any cost implications that you will incur for participating in the study. The information you provide will help to inform the development of a data literacy support services framework to be used in academic libraries in Kenya. The title of the study is: *The feasibility of offering standardised data literacy services at selected private university libraries in Kenya*.

If you have any questions about this study, you may contact me (the PI) or my supervisors. Our contact details are on the consent form. My supervisors are:

Supervisor	Co-Supervisor
Dr MJ van Deventer	Prof JTD Bothma
Vandeventer.martha@up.ac.za	Theo.Bothma@up.ac.za
+2782 924 6650	

Department: Information Science, University of Pretoria

Your anonymity as a respondent is guaranteed. You are not required to share any information that could be traced back to you. From the data collected, no one will be able to identify you. All data collected will be reported in aggregated format. Furthermore, no information provided during the interview will in any way affect or influence your present or future status as a researcher.

Your participation in this study is voluntary. You have the right to withdraw your approval to participate at any moment after consenting and without giving a reason. Withdrawing from this study will have no impact on your relationship with the researcher, neither will it attract any penalty. The interview will take a time range of between 40 - 60 minutes, depending on the depth of your responses.

I would like to record the interview. The recording will be used to create accurate notes for the interview. The recording will be destroyed once the thesis has been examined.

When you sign the consent form, you are indicating that you have read and understand the detail provided above and that you agree to take part.

Thank you very much for your time and support.

Do you have any questions before we start the interview?

Note:

All interviewees will receive the interview schedule one week before the day of the interview.

This is to allow the responded time to familiarise him-/herself with the interview questions.

Interview questions

1. Tell me about your role and how long you have been in your position as a University Librarian at this university.

   a. *What are the main responsibilities in your role?*

   b. *What does your job entail?*

   c. *Explain a normal day in your life as a university librarian.*

2. Which of the following research data services are provided by your library?

   The interviewer will tick the services that the interviewee confirms are provided by the library

   a. Helping researchers with creating data management plans [ ]

   b. Creation and management of institutional data repositories [ ]

   c. Providing tools for data mining and visualization [ ]

   d. Training researchers on data management activities [ ]

   e. Guidance on institutional policies [ ]

   f. Helping with the creation of metadata for data sets [ ]

   g. Assistance with intellectual property and privacy issues surrounding research data [ ]

   h. Digital data archiving and preservation [ ]

   i. Others (please specify)

   ....

   j. None of the items listed [ ]

3. As an institution offering postgraduate research, which of the following policies that support research data services are already in place? *The interviewer will tick policies confirmed by the interviewee.*

   a. Research policy [ ]

   b. Research Data Management policy [ ]

   c. ICT policy [ ]

   d. Staff development policy [ ]

   e. Open access policy [ ]

   f. Digital preservation policy [ ]

   g. Others (please specify) ....

   h. None of the items listed [ ]

4. How does the organisational structure and arrangement in your university support data management and data literacy for researchers?

    a. What are some of the organisational structures that are in place in your institution that are aimed at spearheading data literacy?

    b. How is data management a key component in the organisational structure of your university?

    c. What do you foresee as some of the cost implications of implementing a data literacy programme?

    d. Other than the library, who else do you think should be involved in data literacy programme activities?

5. What level of attention does your university give to research and related services?

    a. How is research captured in the strategic plan of the university?

    b. How are researchers recognised at your university?

    c. How are research activities funded in your university?

6. Please explain the ICT infrastructure that has been provided, to researchers, by the university to enable data collection, data manipulation and data sharing?

    a. Does the university have a data repository? If not, does it provide access to any national data repository? Please explain.

    b. What do researchers currently do when they have to share their data with external research partners?

7. What are your views about the library spearheading the data literacy program in your university?

   Could you explain why you feel that way?

8. How are you addressing data literacy training needs and services within your library?

9. What are your thoughts regarding the 'research data' skills levels among librarians at your university?

    a. How well equipped, do you think, are the current set of librarians when you consider the skills and knowledge necessary to offer data literacy training? Please explain your point of view.

    b. If they are not ready to train researchers: what can be done to enhance their skills and knowledge in data literacy to take up the role of data literacy experts?

    c. If they are ready to assist: please explain the process you went through to get to the appropriate level of proficiency.

10. Please refer to the "Explanation of the study" document. When considering the proposed data literacy framework please respond to the following questions:

   a. *What does not make sense to you?*

   b. *What do you like about the framework?*

   c. *What gaps could you identify in the framework?*

   d. *What would you like to remove?*

11. In view of the proposed framework (refer to the "Explanation of the study" document), what would be your response if you were requested to collaborate in the development of a single data literacy curriculum that would provide standardised data literacy training to all Kenyan researchers - similar to what was done with information literacy? *If positive:*

   a. *What are some of the key components that should be included in the data literacy curriculum?*

   b. *Which roles do you think the library should play in the development and implementation of data literacy?*

   c. *Other than the library who are some of the other stakeholders that would play a key role in the development and implementation of a data literacy program and what would this role be?*

   d. *What factors within your organisation can be the enablers for the emergence of this role in data literacy?*

12. Do you have any other information, related to data literacy, that you think will be of use to this study?

Appendix III: Interview guide for research/reference librarians

Thank you for agreeing to participate in this research which aims to inform the development of data literacy support services in academic libraries in Kenya.

As a reminder: This invitation is based on your capacity as a research/reference librarian. Your participation will require that you respond to the interview questions in this schedule. As a participant, you are assured that there are no known risks if you accept to participate, nor are there any cost implications that you will incur for participating in the study. The information you provide will help to inform the development of a data literacy support services framework to be used in academic libraries in Kenya. The title of the study is: *The feasibility of offering standardised data literacy services at selected private university libraries in Kenya*.

If you have any questions about this study, you may contact me (the PI) or my supervisors. Our contact details are on the consent form. My supervisors are:

Supervisor	Co-Supervisor
Dr MJ van Deventer	Prof JTD Bothma
Vandeventer.martha@up.ac.za	Theo.Bothma@up.ac.za
+2782 924 6650	

Department: Information Science, University of Pretoria

Your anonymity as a respondent is guaranteed. You are not required to share any information that could be traced back to you. From the data collected, no one will be able to identify you. All data collected will be reported in aggregated format. Furthermore, no information provided during the interview will in any way affect or influence your present or future status as a researcher.

Your participation in this study is voluntary. You have the right to withdraw your approval to participate at any moment after consenting and without giving a reason. Withdrawing from this study will have no impact on your relationship with the researcher, neither will it attract any penalty. The interview will take a time range of between 40 - 60 minutes, depending on the depth of your responses.

I would like to record the interview. The recording will be used to create accurate notes for the interview. The recording will be destroyed once the thesis has been examined.

When you sign the consent form, you are indicating that you have read and understand the detail provided above and that you agree to take part.

Thank you very much for your time and support.

Do you have any questions before we start the interview?

Note:

All interviewees will receive the interview schedule one week before the day of the interview. This is to allow the responded time to familiarise him-/herself with the interview questions.

358

Interview questions

1. Tell me about your current role and how long you have been in your position as a research/reference librarian in this university.

   For example:

   a. *What are the main responsibilities in your role?*

   b. *What does your job entail?*

   c. *Briefly describe your academic qualifications in relation to your roles as a research/reference librarian*

   d. *Explain a normal day in your life as a research librarian.*

2. Which of the following research data services are provided by your library?

   The interviewer will tick the services that the interviewee confirms that they are provided by the library

   a. Helping researchers with creating data management plans [ ]

   b. Creation and management of institutional data repositories [ ]

   c. Providing tools for data mining and visualization [ ]

   d. Training researchers on data management activities [ ]

   e. Guidance on institutional policies [ ]

   f. Helping with the creation of metadata for data sets [ ]

   g. Assistance with intellectual property and privacy issues surrounding research data [ ]

   h. Digital data archiving and preservation [ ]

   i. Others (please specify)

      ....

   j. None of the items listed [ ]

3. What are some of the most common research data services sought by researchers in your university?

4. As a research/reference librarian would you consider yourself as having skills and competency in the following areas as related to research data management?

   The interviewer will tick skills and competencies confirmed by the interviewee

   a. Development of data management plans [ ]

   b. Data curation [ ]

   c. Technical skills in data management [ ]

   d. Research methodologies [ ]

e. The Research lifecycle [ ]

f. Data description and documentation [ ]

g. Institutional repository management [ ]

h. Open access initiatives [ ]

i. Legal and copyright frameworks [ ]

j. Others (please specify)..

k. None of the items listed [ ]

5. As a research/reference librarian what would you consider as gaps in your knowledge base and skills (technical and non-technical) in relation to your current roles and duties of serving researchers?

6. As an institution offering research data management services, which of the following policies are in place that support research data services? *The interviewer will tick policies confirmed by the interviewee.*

a. Research policy [ ]

b. RDM policy [ ]

c. ICT policy [ ]

d. Staff development policy [ ]

e. Open access policy [ ]

f. Preservation policies [ ]

g. Others (please specify) ....

h. None of the items listed [ ]

7. Which of the following training interventions are offered to researchers by the library? *The interviewer will tick trainings confirmed by the interviewee.*

a. Research data management [ ]

b. Statistical data analysis tools [ ]

c. Research data publication [ ]

d. Ethics considerations in research [ ]

e. Research metrics and impact [ ]

f. Data visualization [ ]

g. Data archiving [ ]

h. Others (please specify) ....

i. None of the items listed [ ]

8. What are your views about the library spearheading the data literacy program in your university?

Could you explain why you feel that way?

9. To what extent do librarians already work with researchers at your university?

10. What challenges do you face when providing research data-related services to researchers?

11. You were provided with a copy of the "Explanation of the study" document. Please have a look at the proposed data literacy framework and consider the following:
    a. *What does not make sense to you?*
    b. *What do you like about the framework?*
    c. *What gaps could you identify in the framework?*
    d. *What would you like to remove?*

12. In view of the proposed framework (provided in the "Explanation of the study" document), what would be your response if you were requested to collaborate in the development of a single data literacy curriculum that would provide standardised data literacy training to all researchers in Kenya -similar to what was done with information literacy?

    If positive:

    a. *What are some of the key components that should be included in the data literacy curriculum?*
    b. *Which roles do you think the library should play in the development and implementation of data literacy?*
    c. *Other than the library who are some of the other stakeholders that would play a key role in the development and implementation of a data literacy program and what would this role be?*
    d. *What factors within your organisation can be the enablers for the emergence of this role in data literacy?*

13. Do you have any other information related to data literacy that you think will be of use to this study?

Appendix IV: University of Pretoria ethical approval

Faculty of Engineering, Built Environment and Information Technology

Fakulteit Ingenieurswese, Bou-omgewing en Inligtingtegnologie / Lefapha la Boetšenere, Tikologo ya Kago le Theknolotši ya Tshedimošo

8 November 2022

Reference number: EBIT/170/2022

Mr SL Agava
Department: Information Science
University of Pretoria
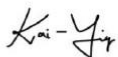Pretoria
0083

Dear Mr SL Agava,

FACULTY COMMITTEE FOR RESEARCH ETHICS AND INTEGRITY

Your recent application to the EBIT Research Ethics Committee refers.

Approval is granted for the application with reference number that appears above.

1. This means that the research project entitled "The feasibility of offering standardised data literacy services at selected private university libraries in Kenya" has been approved as submitted. It is important to note what approval implies. This is expanded on in the points that follow.

2. This approval does not imply that the researcher, student or lecturer is relieved of any accountability in terms of the Code of Ethics for Scholarly Activities of the University of Pretoria, or the Policy and Procedures for Responsible Research of the University of Pretoria. These documents are available on the website of the EBIT Research Ethics Committee.

3. If action is taken beyond the approved application, approval is withdrawn automatically.

4. According to the regulations, any relevant problem arising from the study or research methodology as well as any amendments or changes, must be brought to the attention of the EBIT Research Ethics Office.

5. The Committee must be notified on completion of the project.

The Committee wishes you every success with the research project.

Prof K.-Y. Chan
Chair: Faculty Committee for Research Ethics and Integrity
FACULTY OF ENGINEERING, BUILT ENVIRONMENT AND INFORMATION TECHNOLOGY

Appendix V: NACOSTI research permit

REPUBLIC OF KENYA

NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY & INNOVATION

Ref No: **660943**

Date of Issue: **03/August/2022**

RESEARCH LICENSE

This is to Certify that Mr.. Stanislaus Litsalia Agava of University of Pretoria, has been licensed to conduct research in Nairobi on the topic: The feasibility of offering standardized data literacy services at selected private university libraries for the period ending : 03/August/2023.

License No: **NACOSTI/P/22/19566**

660943

Applicant Identification Number

Director General
NATIONAL COMMISSION FOR
SCIENCE,TECHNOLOGY &
INNOVATION

Verification QR Code

NOTE: This is a computer generated License. To verify the authenticity of this document,
Scan the QR Code using QR scanner application.

363

Appendix VI: Catholic University of Eastern Africa

THE CATHOLIC UNIVERSITY OF EASTERN AFRICA

Office of the Deputy Vice Chancellor
ACADEMIC AFFAIRS & RESEARCH

Our Ref: DVC/AA&R/SN/esm/075/2022

23rd August 2022

Stanislaus Litsalia Agava
Department of Information Science
Faculty of Engineering, Built Environment & Information Technology
University of Pretoria
Ref: 16383053
Email: agavastanislaus1759@gmail.com

Dear Stanislaus,

RE: Permission to Conduct Research at The Catholic University of Eastern Africa (CUEA)

Greetings in the Mighty Name of our Lord and Savior Jesus Christ!

I am glad to inform you that your request to conduct research on the topic: *"The Feasibility of offering standardized data literacy services at selected private University libraries in Kenya"*, has been granted. You are therefore authorized to collect data from the targeted participants at The Catholic University of Eastern Africa. You are expected to strictly observe the normal ethical cautions and discretions while conducting the research.

I wish you well with your study and I look forward to you sharing your findings with the Directorate of Research and Innovation of the The Catholic University of Eastern Africa.

Sincere regards

Prof. Samuel Nyanchoga
Ag. Deputy Vice Chancellor/Academic Affairs & Research

CC   *Vice Chancellor*
      Directorate of Research and Innovation
      University Librarian

Appendix VII: United States International University

Mr. Stanislaus Litsalia Agava
17th August, 2022

Dear Mr. **Agava**,

REF: PERMISSION TO CONDUCT RESEARCH AT USIU-AFRICA

Following your request to conduct research at USIU-Africa on the topic **"The feasibility of offering standardized data literacy services at selected private university libraries"**, the university's Research Office has authorized you to collect data from the respondents in the Institution.

However, the university imposes the following conditions:

1. No personal information will be asked of the respondents.

2. You will share the preliminary report findings with us prior to completion.

3. You will provide a copy of the completed research to us.

4. Under no circumstances will the information obtained from USIU-Africa be re-used or disclosed for other purposes.

Your research period expires on 18th August, 2023. Kindly contact the undersigned to confirm your acceptance to the condition stated above.

Sincerely,

Prof. Amos Njuguna,
Dean- School of Graduate Studies, Research and Extension.
Tel: 0730116442
Email: amnjuguna@usiu.ac.ke

p.o.box 14634-00800 Nairobi, Kenya | tel:254-730-116-000 | info@usiu.ac.ke
www.usiu.ac.ke

365

Appendix IX: African International University

AFRICA INTERNATIONAL UNIVERSITY

Committed to His mission
Connected to His world

19th August 2022

Mr. Agava Stanislaus Litsalia
University of Pretoria
Pretoria.

Dear Mr. Agava,

RE: PERMISSION TO UNDERTAKE RESEARCH

Receive greetings from Africa International University!

I acknowledge receipt of your letter dated 7th August 2022, requesting for permission to conduct research for a doctoral degree on the topic: "**The feasibility of offering standardized data literacy services at selected private university libraries in Kenya**".

Your request is hereby granted trusting that the findings of the research will be beneficial to the AIU. Upon completion of research, kindly submit a copy of the thesis to the university.

Meanwhile, we wish you success in your studies.

Sincerely,

Prof. James Nkansah-Obrempong, PhD
Ag. Deputy Vice - Chancellor Academic Affairs & Research

CONSTITUENT SCHOOLS: SCHOOL OF BUSINESS AND ECONOMICS (SBE) INSTITUTE FOR THE STUDY OF AFRICAN REALITIES (ISAR) SCHOOL OF EDUCATION, ARTS, AND SOCIAL SCIENCES (SEASS) NAIROBI EVANGELICAL GRADUATE SCHOOL OF THEOLOGY (NEGST)

P.O. Box 24686 00502 Karen Nairobi-Kenya TEL: +254-(0)20-260 3663, 0715-247540 Admission Hotlines: 0725-841885
Email: pr@africainternational.edu Web: www.aiu.ac.ke

366

Appendix X: Adventist University of Africa

AUA
Adventist University of Africa

Private Bag Mbagathi
00503 Nairobi, Kenya
Tel (254) 733 333 451/2
Email: Info@aua.ac.ke
Web: www.aua.ac.ke
Location:
Advent Hill, Magadi Road, Ongata Rongai

August 18, 2022

To Whom It May Concern:

This letter confirms that **Stanislaus Litsalia Agava,** a student at the University of Pretoria, is allowed to collect data from the Adventist University of Africa full-time faculty members and Ph.D. students for his research project entitled "*The feasibility of offering standardized data literacy services at selected private university libraries in Kenya.*"

Sincerely,

Josephine Ganu, PhD
Director, Research & Grants Development
Associate Professor of Management
Email: ganuj@aua.ac.ke

367

Appendix XI: Daystar University

1st September 2022

Agava Stanislaus Litsalia
University of Pretoria, SA

Dear Mr. Agava,

RE: PERMISSION TO CONDUCT RESEARCH AT DAYSTAR UNIVERSITY

Your request for permission to conduct your research entitled: "*The feasibility of offering standardised data literacy services at selected private university libraries in Kenya*", *has* been received.

We have examined your research proposal and have satisfactorily established the purpose of your research. We have also looked at your research tools critically and note that you have the relevant approval from your School.

Having done all the above, I wish to confirm that you have been granted permission to carry out the study. The permission granted is for a period of one year from the date of this letter.

We wish you all the best in your data collection and eventual completion of your dissertation.

Yours sincerely,

Prof. Faith Nguru
Deputy Vice-Chancellor – Academic, Research & Student Affairs

"*...until the day dawn and the daystar arise in your hearts*"
2 Peter 1.19 KJV