sciendo

Research Article **Open Access**

Johnson Masinde*, Jing Chen, Daniel Wambiri, Angela Mumo

Research Librarians' Experiences of Research Data Management Activities at an Academic Library in a Developing Country

https://doi.org/10.2478/dim-2021-0002
received November 3, 2020; accepted January 28, 2021.

Abstract: University libraries have archaeologically augmented scientific research by collecting, organizing, maintaining, and availing research materials for access. Researchers reckon that with the expertise acquired from conventional cataloging, classification, and indexing coupled with that attained in the development, along with the maintenance of institutional repositories, it is only rational that libraries take a dominant and central role in research data management and further their capacity as curators. Accordingly, University libraries are expected to assemble capabilities, to manage and provide research data for sharing and reusing efficiently. This study examined research librarians' experiences of RDM activities at the UON Library to recommend measures to enhance managing, sharing and reusing research data. The study was informed by the DCC Curation lifecycle model and the Community Capability Model Framework (CCMF) that enabled the Investigator to purposively capture qualitative data from a sample of 5 research librarians at the UON Library. The data was analysed thematically to generate themes that enabled the Investigator to address the research problem. Though the UON Library had policies on research data, quality assurance and intellectual property, study findings evidenced no explicit policies to guide each stage of data curation and capabilities. There were also inadequacies in skills and training capability, technological infrastructure and collaborative partnerships. Overall, RDM faced challenges in all the examined capabilities. These challenges limited the managing, sharing, and reusing of research data. The study recommends developing an RDM unit within the UON Library to oversee the implementation of RDM activities by assembling all the needed capabilities (policy guidelines, skills and training, technological infrastructure and collaborative partnerships) to support data curation activities and enable efficient managing, sharing and reusing research data.

Keywords: research data, research data management activities, research librarians' perceptions, academic libraries, developing countries

1 Introduction

Research data is ordinarily complex, irreplaceable, costly, and laborious to generate. They are invaluable knowledge assets in an electronic or non-digitized form generated by the scholarly community in the course of their research process (Ohaji, 2016). Accordingly, it is indispensable for research organizations to effectually capture, describe, manage, and make available the data for discovery, sharing, and reusing. The propensity to manage and preserve data has been prompted by the increasingly data-intensive science, an acknowledgment of the critical need for a coordinated and fixated approach to data management, and the need to develop immense potentials for competitiveness, productivity, and liveability (Ingram, 2019).

Today, RDM has become a global subject for academic libraries due to their previous experiences with open access services and their proactive role in shaping scholarly communication. Similarly, funders, publishers, and government agencies are mandating researchers to develop data management plans and make the data occasioning from their research openly

***Corresponding author: Johnson Masinde,** School of Information Management, Central China Normal University, Wuhan, China. Email: masindejohnson@mails.ccnu.edu.cn
Jing Chen, School of Information Management, Central China Normal University, Wuhan, China
Daniel Wambiri, School of Education, Kenyatta University, Nairobi, Kenya
Angela Mumo, Library and Information Services, University of Nairobi, Nairobi, Kenya

available for sharing and reuse. Therefore, researchers are turning to academic libraries for the management of their data. In this context, the contemporary study attempts to examine research librarians' experiences with RDM activities at an academic library in an emerging economy to recommend measures to enhance managing, sharing, and reusing research data. The framing of the research problem is motivated by the need to ascertain the current state of RDM in academic libraries of emerging economies from a librarians' perspective. At the core of this was identifying the presence of the needed capabilities (Policy guidelines, knowledge and skills, technological infrastructure and collaborative partnerships) to support data curation activities to enable efficient managing, sharing and reusing research data. The research approach undertaken enabled the research problem to be addressed in such a way as to permit comparisons over time.

1.1 Research Problem

RDM activities in most academic libraries of emerging economies have been described as lacking, rudimental and unstructured (Chigwada, Chiparausha, & Kasiroori, 2017). Similarly, very little research was conducted to examine how the libraries capture, describe, manage, and make available research data for discovery, sharing, and reusing (Josiline, Hwalima, & Kwangwa, 2019). Literature review also shows that not much attention was paid to key stakeholders such as researchers, librarians, the IT office, the research office, and the legal office. In view of that, the current study seeks to examine the research librarians' experiences of RDM activities at an academic library in an emerging economy to recommend measures to enhance managing, sharing, and reusing research data.

2 Literature Review

Librarians and researchers reckon that preservation and management of research data ought to be a key role of libraries. Banking on the reputable grounds of institutional repositories, academic libraries are expected to amass multiple data from scientific research activities, select, classify, index, preserve, and avail it for discovery, sharing and reuse. Accordingly, Cox and Pinfield (2014) in a report titled 'The Last Mile' mention that academic libraries have a duty to employ novel stratagems and assemble capabilities in the form of policies, expertise, technologies, and partnerships to enable efficient managing, sharing and reuse of data.

When research data is appropriately managed, it can ensure more productive and efficient science as new knowledge is created by building on discoveries, innovations, and open scientific inquiry. Furthermore, research data reuse can reduce the costs of researching by making data readily available while limiting its duplication and acting as a way of meeting funder mandates (Woeber, 2017). Universities, particularly in developed countries, have been addressing this strategic objective. Key RDM stakeholders, for example, academic libraries, the legal office, the IT office, the research office, and the parent institution have been engaged in developing capabilities such as supportive policies, skills and training, technological infrastructure, and started collaborating with other relevant institutions and organizations to ensure efficient managing, sharing and reusing research data (Cox, Kennan, Lyon, & Pinfield, 2017).

In contrast, Ng'eno (2018) observes that research institutions in emerging economies have mainly restricted their RDM efforts to advocacy programs, with most institutions lacking or at the initial stages of RDM implementation. In this regard, the researchers have repeatedly had to cope with multifaceted realities of making decisions in the absence of harmonized data management systems; which has proved challenging. Similarly, Nhendodzashe and Pasipamire (2017) who investigated academic libraries' readiness to offer research data services in Zimbabwe, found that the libraries lack the needed capabilities such as supportive laws and policies, human resource capabilities, robust technological infrastructure, and collaborative partnerships to support RDM activities.

In Kenya, a study conducted by Anduvare (2019) that investigated data support services among Kenyan private universities established that researchers managed their own data that was in different formats and scattered all over the shelves and computers. There was no framework to guide RDM activities. Additionally, Iinformation professionals were deficient in RDM competencies and were not engaged in any form of RDM activities despite knowing its significance (Anduvare, 2019). Consequently, the research institutions played little to no role in managing, sharing, and reusing research data, which resulted in the loss of valuable knowledge assets, duplication of findings, lack of innovations, increased research costs, inability to meet funder mandates and poor decision making.

To check RDM challenges, Flores, Brodeur, Daniels, Nichools, and Turnator (2015) propose an organizational

change within the academic libraries, to establish a specific unit to oversee the managing of research data activities. According to the study, the unit should guide the whole research data management process, including the roles, responsibilities, services and coordination. Such functions and responsibilities should include the formulation of research data management strategies and mechanisms, implementation and effecting of policy guidelines and operationalization of awareness and advocacy campaigns. Additionally, the unit should develop capture instruments, clean and validate research data, design database systems and facilitate researcher collaboration at national and international levels. Further still, Flores et al. (2015) advance the need for the unit to manage the portal services, metadata standards and all the data curation activities, including developing an ethical framework to guide the management of research data.

2.1 Theories Underpinning the Study

The Open Archival Information System (OAIS), the Maturity Model, the DCC Curation Lifecycle Model, and the CCMF have attempted to explain RDM. The OAIS does not include a phase before the ingest and pre-ingest function making it short of the theoretical basis to effectively underpin the study (Ng'eno 2018). Similarly, the maturity model may be taken to infer a specific development course, resulting in a fixed and established definitive setting. However, this is rarely the case. A term like "underdeveloped," or "immature" when linked to maturity frameworks, may be regarded as disapproving (Cox et al., 2017). In this study, the theories that align well with the study problem are the Community Capability Model Framework (CCMF) and the DCC Curation Lifecycle Model

The CCMF provides details of the roles and responsibilities of each capability for effective management of research data. However, the model does not focus in detail on data curation which is the nucleus of RDM (Lyon, Ball, Duke, & Day, 2012). Similarly, the DCC Curation lifecycle model offers a high-level view of data curation allowing the activities to be planned at different granularity levels including defining tasks and responsibilities, building frameworks of standards and technologies, and ensuring processes and policies are adequately documented (Higgins, 2012). However, like the CCMF, the model does not support RDM capabilities. Higgins (2012) asserts that the model can be utilized in conjunction with other relevant models to plan and explain RDM. T*he investigation adopts* the DCC Curation lifecycle model and the CCMF to elucidate research librarians' experiences of RDM activities.

3 Methodology

The investigation adopted thematic analysis of qualitative research that was underpinned by the DCC Curation Lifecycle Model and the CCMF. The DCC Curation Lifecycle Model underpinned data curation activities while the CCMF informed the capabilities (Policy guidelines, skills and training, technological infrastructure and partnerships).

Gibbs (2007) states that thematic analysis is a form of qualitative research method that focuses on identifying passages of text that are interconnected by a common idea or theme, allowing the Investigator to index the text into categories or groupings, consequently generating a pattern of thematic ideas about it. Thematic analysis identifies themes in the data that are important or interesting and uses them to address the research problem.

3.1 Location of the Study

The UON Library stands as a hub for providing research services to postgraduate students and academic staff. The library offers a suite of research data services, including partnerships in the form of linkages and workgroups, interdisciplinary and emerging research methods, and expert guidance at all stages of the research process. The UON Library is also manned by expert personnel with competencies in research-related services. The Investigator selected the library because:

i. In 2020, the US news Education ranked University of Nairobi position 792 globally, 18th in Africa, and 1st in Kenya depicting its research standing and suitability as a case (U.S.News, & World Report, 2020);
ii. University of Nairobi has an annual research kitty of 45 million USD, the highest in Kenya. This funding signifies its involvement in rich, valuable, and high intensive research (Nacosti, 2020);
iii. It has the largest student population in Kenya at 84,000, with at least 14,500 postgraduates (University of Nairobi, 2020);
iv. It was the first to be ISO certified in 2008 (University of Nairobi, 2015). With ISO certification comes amplified expectations in quality, customer care, and satisfaction. Therefore, the researchers expect efficient and high-quality RDS.

3.2 Target Population

The investigation targeted the research librarians who were familiar with research data services at the *UON Library*.

3.3 Sampling Techniques

The study considered it appropriate to select participants based on their knowledge of the research problem. This kind of sampling is referred to as purposive sampling (Etikan & Bala, 2017).

3.3.1 Sample Size

The Investigator purposively selected 7 participants. The Interviewees were senior research librarians at the UON Library. They had previously been involved in RDM projects and were either directly or indirectly involved in the delivery of RDS.

3.4 Data Collection

The investigation employed a semi-structured interview to capture qualitative data from respondents. The primary goal of selecting unstructured interviews was to have the participants guided by the study questions and at the same time, express their views freely and in their own words (Pickard, 2013).

Block and Erskine (2012) aver that online interviews are cost-effective and time-efficient compared to other methods such as face to face interviews. Additionally, Block and Erskine (2012) mention that they have a personal touch with broader geographical access. Considering the researcher's primary focus, which was to utilize the best possible data collection methods as per the prevailing conditions (COVID19 pandemic) and exploit contemporary trends, the study selected the online interviews.

Janghorban, Roudsari, and Taghipour (2014) identify two online interview methods; asynchronous and synchronous. In asynchronous interviews, the interviewer and interviewee do interviews at different time frames. The interviews include emails, bulletin boards, and discussion groups. On the contrary, synchronous interviewees are executed in real-time and include instant messenger protocols, text-based chat rooms, and videoconferencing such as zoom. Accordingly, zoom interviews presented the best option for the Investigator who had to try new things in place of the conventional methods.

The questions were designed to explore issues reflecting the study's purpose and were developed using information gained from the literature review, contacts in the field, and for comparative purposes, with prior surveys. Consequently, the data collected covered the following major themes.

i. Legal and policy guidelines influencing research data management activities;
ii. Digital curation activities;
iii. Human resource capabilities (knowledge, skills, and training requirements for research data management activities);
iv. Technology and infrastructure readiness for research data management activities;
v. Collaborative partnerships influencing research data management activities.

3.5 Validity and Reliability

Before data collection, the investigator subjected the interview instrument to both face and content validity. Connell et al. (2018) state that face validity is the extent with which a test is subjectively observed as blanketing the concept or idea it measures while content validity is the degree with which a measure represents all the items of a construct. To assess the instrument's face and content validly, the Investigator invited professionals to critic the interpretation of the items in the constructs. The Investigator revised three items as suggested by the professionals.

3.6 Pilot Studies

Upon authorization to conduct the study, the investigation piloted the instrument among the research librarians at the KU Post Modern Library. The Investigator chose a different library to check study contamination.

The study participants were 3 senior research librarians. According to Crouch and McKenzie (2006), pilot studies are not intended for data collection but to ground truth. After the pilot study, the Investigator invited RDM experts for further assessment.

3.7 Data Collection Procedures

Subsequently, the Investigator contacted the respondents between February and May 2020 for briefing and subsequent invitation to participate in the study. Of the 7 Interviewees who had confirmed participation, only 5 honored their appointments. Before the commencement of the interviews, each participant was given a chance to confirm their consent. Participation was on the interviewee's sole discretion, explicit accord and de-identified reporting.

Respondents were urged to share their experiences in-depth about the themes with interviews lasting 40 minutes on average. The Investigator made notes and audio recordings of all the interview conversations. This was necessitated by the need to make reference and have a permanent record for verification and future use (Tessier, 2012). The Investigator also did the transcription after each interview to increase the accuracy and circumvent a mix-up.

The Investigator aimed at supplementing the study with document reviews due to the small number of respondents. However, upon searching and inquiring, the study established no meaningful information about the research problem. The available documents were deemed not sharable because they carried other sensitive information. Therefore, the Investigator was left with no choice but to make maximum use of interviews. Nonetheless, Dworkin (2012) explains that qualitative research interviews involving 5-50 participants are considered sufficient to make meaningful conclusions.

3.8 Data Analysis

The study utilized Braun (2006)'s six-step framework of analysis. Maguire and Delahunt (2017) describe the process as the most effective method in that it offers a rich and practical framework for thematic analysis. Braun (2006) suggests familiarizing oneself with the data, which was attained by re-listening to the recorded interviews, re-reading the excerpts, transcribing or re-writing the recorded data, and the excerpts; generating codes to pinpoint features of interest and placing the codes together to form possible themes that are used to address the study problem. Maguire and Delahunt (2017) explain that the themes or patterns generated are much more than merely summarizing the data; an excellent thematic analysis construes and makes meaning of the data. The investigator connected several themes with the coded excerpts. The themes were indexed to give richer meanings that addressed the research questions.

3.9 Ethical Considerations

Ethics in research warrant that no individual suffers the harmful consequence of research activities (Masinde, Wambiri, & Chen, 2020). Consequently, the Investigator took time to inform respondents of the nature and aim of the research before requesting their consent.

4 Data Analysis and Discussion of Findings

The findings are presented in the context of the key themes as itemized below:

i. Policy guidelines that influence RDM activities;
ii. The data curation activities;
iii. The human resource capabilities (skills and training requirements for RDM activities);
iv. Technological infrastructure readiness for RDM activities;
v. Collaborative partnerships influencing RDM activities.

4.1 The Availability or Absence of Policy Guidelines Influencing RDM Activities

Policy guidelines respond to RDM activities such as preservation, storage, quality, security, sharing, compliance, and jurisdiction to improve managing sharing and reusing research data (Nhendodzashe & Pasipamire, 2017). They are supported by the DCC Curation lifecycle model which defines roles and policies at each stage and the CCMF, ensuring processes and policies are adequately documented and adhered to at all stages after being defined and premeditated (Higgins, 2012; Lyon et al., 2012). Consequently, the study investigated the policy guidelines influencing RDM activities at the UON Library

4.1.1 Policy Guidelines Governed RDM Activities

Respondents were asked to reveal if the library had policies that guided RDM activities such as data curation, technological infrastructure, skills and training capabilities, and collaborative partnerships.

Study findings revealed that the library had a policy on research data that guided research data activities such as data capture, appraisal, description, preservation, access, use, reuse, and sharing. Furthermore, the interviewees added that they were continually refining the policy through engagements with stakeholders to meet user needs fully. One respondent mentioned that:

> *"We have a policy on research data that guides these activities. We are also refining it to ensure it meets all our needs." (RL3)*

Asked whether the library had explicit policies on data curation activities such as data capture, data appraisal, data description, data preservation, data access, data sharing and data reuse, etc. The interviewees disclosed that the policy on research data guided data curation activities collectively. However, the library had two more policies, the quality assurance policy that ensured quality was maintained and enhanced and the intellectual property rights policy that protected creativity and innovations from being appropriated by others. One respondent stated that:

> *"The policy on research data covers all the data curation activities collectively. We also have the quality assurance policy and the Intellectual property rights policy that also support research services." (RL1)*

4.1.2 Research Data Ownership

The study further sought to ascertain who owned the rights to research data produced by the university's research community.

Interviews with most research librarians revealed that the university shared ownership rights with the researchers, clearly signaling no clear-cut guidelines on research data ownership. One respondent commented that:

> *"We share the rights. Most of the output is usually part of the requirements for completing their academic programs or part of requirements as an academic staff. "(RL6)*

4.2 Investigation of the Data Curation Activities

Data curation is the active and continuous management of research data through its lifecycle of interest and usefulness for discovery, sharing, and reuse (Koltay, 2019). It is underpinned by the DCC Curation lifecycle model; which supports digital materials' management, permitting their effective curation and preservation from initial conceptualization to disposal or selection for reuse, and long-term preservation. The activities involved in data curation include data capture, appraisal, description, preservation, access, reuse, and transformation of research data (Higgins, 2011).

Similarly, the CCMF imparts data curation activities by supporting capabilities (policy guidelines, skills and training, technological infrastructure and partnerships) in data curation (Lyon et al., 2012). Accordingly, the study investigated data curation activities at the UON Library.

4.2.1 Data Curation Activities

Interviewees were requested to describe the data curation activities at the library. The findings suggest that the library strictly adhered to the data curation lifecycle. Data was curated for value addition, discovery, sharing and reusing. One respondent, while describing the activities mentioned that:

> *"Data capture, which is the first stage involves the collection and storage of data using descriptive metadata to facilitate its later discovery and access. The second stage is the appraisal stage and involves the assessment and selection of research data for long-term curation and preservation. The third is the disposal stage, and in this stage, data that is not selected for preservation and curation is discarded. Sensitive information is usually detracted through shredding. The fourth is the ingestion, and this is the stage where research data is transferred to the repository. The fifth is the preservation stage, and here, several actions are executed to ensure long term retention and preservation of data. These actions include audits and privacy checks to guarantee the integrity and controlled access to research data. The data is frequently assessed to maintain quality. There is also the re-appraisal stage where we re-evaluate data that failed validation for further appraisal and reselection. The seventh is the storage stage, and here, research data is stored in a safe and secure place free from unauthorized access and also availed for access, sharing and reusing. Lastly, we have the transformation stage. This stage concerns value addition to original data. We also maintain hard copies for reference purposes" (RL3).*

4.2.2 Research Data Appraisal

Interviewees were also asked to mention the items in the appraisal checklist. Accordingly, the respondents cited the date, authority, accuracy, coverage, objectivity, documentation, volume, added value, reuse value, substantivizes and accessibility.

Asked about the tools they used in appraisal of research data, the interviewees stated that the library had a policy on research data that guided research data appraisal. However, further inquiry suggested that they seldom applied the policy. One respondent commented that;

> *"Our library has a policy on research data that guides appraisal of the data. It dictates the data to keep for long-term and that to discard." (RL2).*

4.2.3 Metadata Creation

Metadata makes datasets discoverable, accessible, citable and reusable. The metadata elements considered when describing and documenting research data include author, title, subject, date of creation, description, funding, rights, publishing organization, and a unique identifier (Dalhousie University Libraries, 2019). The study investigated if researchers generated their own metadata or the library generated metadata for use. Study findings indicated that the library generated metadata for use. One interviewee explained that:

> *"We generate our own metadata for consistency in the definition of terms, clarity in relationships, security, and privacy mandates, ease in management, and discovery data" (RL5)*

4.2.4 Storage and Preservation of Research Datasets

A storage and preservation mechanism ought to be reliable, guarantee safety, permit access and observe all the RDM best practices (Josiline Chigwada, 2017). The study also investigated the short term and long term data storage practices of the UON Library.

The findings show that the library preserved data in the form of publications, webpages and the digital repository on long term. However, there were no clear structures for short term. One respondent stated that:

> *"The University has a number of journals through which we publish the data while the rest is stored in the repository." (RL6)*

4.2.5 Access, Sharing and Reusing Research Data

Research data access, sharing and reuse permit the development of new knowledge through building on previous research, discoveries and open scientific enquiry. Siyao, Whong, Martin-Yeboah, and Namamonde (2017) aver that it is an effective vehicle of information exchange and a way of maximizing on discoverable data as users explore diverse lines of inquiry, and more resourcefully conduct comprehensive research. Consequently, the study investigated the access, sharing and reusing of data at the UON Library.

4.2.5.1 Research Data Access

The investigation requested interviewees to reveal if research data was freely accessible or restricted. Study findings showed that data was both freely accessible and restricted depending on the kind of data itself, the use and the user. One respondent explained that:

> *"We have research data that is freely accessible. We also have data that is restricted. It depends on the kind of data, the use, intellectual property rights, and interests and rights of our researchers." (RL2).*

4.2.5.2 Sharing and Reusing Research Data

Moreover, the study investigated if research data was freely sharable and reusable. The findings indicated that the data was partly sharable and reusable depending on aspects such as the kind of data, the use, the user, sponsorship and ownership rights.

However, further inquiry suggested that data was rarely shared and reused because researchers were reluctant to share their own data. One respondent observed that:

> *"Research data sharing and reuse is dependent on a number of factors such as the use, the kind of data, ownership rights, funder and whom the data is being shared with." (RL2)*

4.3 Human Resources (Skills and Training Requirements for RDM Activities)

RDM encompasses infrastructure, tools, and services that support the research data lifecycle. Academic libraries should put in place infrastructure and services that efficiently support RDM (Cox Andrew, 2019). The DCC Curation lifecycle model, which underpins the study, sets out the roles and responsibilities at every data curation stage. The model blends with the CCMF to support skills, and training requirements for RDM.

4.3.1 Types of Skills Possessed in Regard to RDM

The study also asked respondents to disclose if they possessed skills in RDM activities such as data curation, policy formulation, technological infrastructure, and collaborative partnerships.

The findings evidenced generalized deficiencies of both information and technical data skills in all the examined activities. The respondents acknowledged skillset as a significant challenge to RDM implementation and mentioned the need for continued capacity building and resourcing. Furthermore, they indicated that it was equally challenging to find personnel with an appropriate mix of skillset in areas such as research methods, disciplinary knowledge, data curation, and technical data services. One respondent mentioned that:

> *"One major hurdle for us has been human resourcing, particularly in technical data aspects. We have inadequacies in areas such as metadata construction, data auditing copyright and licensing." (RL3)*

4.3.2 RDM Training Needs

The interviewees were further requested to reveal their immediate training needs regarding RDM activities such as metadata creation, data curation, Information technology tools and equipment, legal and ethical considerations, developing a DMP, grant proposal writing, collaborative partnerships and security and storage of research data.

The majority of the respondents revealed that they needed training in all the mentioned areas. The respondents stated that RDM was still new to them. They utilized experiences acquired over time while working with open access services to build capabilities by reskilling and re-assigning existing roles. There was also a need for continued training and refresher courses in RDM trends and technologies. One respondent commented that:

> *"RDM is a new area that is evolving rapidly. it is yet to be harmonized into the library and information science curriculum. There is need for continued training and capacity building." (RL6)*

4.3.3 Research Data Services and Quality

Regarding the RDS offered and the measures they put in place to enhance and maintain quality, the findings show that the library was primarily engaged in information data services. The interviewees mentioned that they were offering a number of consultative data services such as a data management plan (DMP) support service, data training (data analysis training with SPSS, matrix laboratory (MATLAB) and R, referencing and citation services, originality checker service, SciVal and publishing service guides) resource recommendation and web resource guides. One respondent stated that:

> *"We are primarily engaged in consultative data services. They are usually lighter, softer and reference-based, making it easier for us to implement them. They also align well with traditional reference services." (RL3)*

Similarly, the interviewees disclosed that they were offering a number of technical data services. They defined the services as the "hands-on" or involving specialized skillset and knowledge. The services included run a digital repository, advisory services on technical aspects of long-term data preservation, copyright and intellectual property rights and generation of metadata for datasets. One respondent advanced that:

> *"We are offering some technical data services such as advisory services on technical aspects of long term data storage, preservation and archiving, copyright and intellectual property rights and creation of metadata for datasets." (RL4).*

On how they maintained and advanced the quality of RDS, the interviewees disclosed that the library was ISO certified and observed the highest standards of quality as stipulated by the International Organization for Standardization (ISO). One respondent declared that:

> *"We observe the highest standards as specified by the International Organization for Standardization (ISO)." (RL2)*

4.4 Level of Technological Infrastructure Readiness

Technological infrastructure enables RDM activities. It permits ease in accessibility and management of research outputs, creating more appropriate ways of propagating research data, consequently promoting knowledge integration (Schultz, 2017). The variable is supported by the DCC Curation lifecycle model and the CCMF. The DCC Curation lifecycle model builds standards and frameworks, creating room for technological changes and ensuring the transition from each phase while the CCMF ensures all technical capabilities at each stage are executed adeptly.

4.4.1 RDM Technology and Infrastructure Readiness

Regarding technological infrastructure readiness for RDM activities, most interviewees acknowledged the need for computing facilities, software solutions, storage media, Data backup tools, and security solutions. According to the interviewees, technological infrastructure was a critical subject that hindered the full implementation of RDM activities. They described technical infrastructure as a fundamental component and driver of RDM activities. The interviewees mentioned the need for a powerful and robust technological infrastructure that can effectively support the whole data curation lifecycle, including storage, preservation, security, access, sharing, and reuse of research data. One respondent cited old equipment with low storage capacities and deficiencies in preservation, backups, and access, coupled with a considerable user population, as critical challenges to RDM implementation. The respondent commented that:

> *"There is need for new technologies to warrant efficiency in preservation, back up, and discovery of research data." (RL2).*

Asked about the security solutions implemented to protect research datasets from unwarranted access, the interviewees disclosed that the library had both networked and physical security systems to protect the research resources from unwarranted access. Similarly, data was backed up in the cloud and a central server, with physical documents being put in secure places with strict surveillance. The use of passwords was also encouraged. However, most interviewees cited the need for modern and advanced solutions to guarantee the utmost rigour. One respondent mentioned that:

> *"We have physical and networked security systems to protect data from unauthorized access. However, there is need for current solutions to guarantee maximum security." (RL5).*

4.5 Collaborative Partnerships Affecting Research Data Management Activities

Collaborative partnerships advance access, sharing and reuse of research data. Collaborative partnerships between the library and departments or disciplines within the University, at the national and international level, allow and promote the implementation of shared policies, technical frameworks, documentation, workflows and research cost cuts augmenting access, sharing, and reusing data.

The DCC Curation lifecycle model underpinned the variable by shaping access, sharing, and reusing data. The CCMF focused on capabilities in collaborations between the library and departments or schools within the University, nationally and internationally. Consequently, the investigation examined the collaborative partnerships between the UON Library and other disciplines internally, nationally and internationally.

4.5.1 Research Librarians' Experiences with Internal, External and International RDM Collaborative Partnerships

Study findings revealed that the UON Library was still in the early stages of establishing RDM collaborative partnerships internally, externally, and internationally. According to most interviewees, the library was still developing policies and evaluating the best possible areas of engagement with prospective collaborators. Nonetheless, the interviewees noted a number of collaborations within the university, externally, and internationally. For instance, the interviewees revealed that the library was partnering with academic departments, the ICT department, the research office, and the legal office to mold a rich RDS package for the research community. There were collaborations with the library and information science department to build the skills and training capability. The library was also collaborating with the research office to provide consultative services in the research process (disciplinary knowledge and training researchers on data analysis tools such as SPSS, R, and MATLAB). Externally, the library had partnered with other Kenyan academic libraries to form a consortium, KLISC (The Kenya Libraries and Information Services Consortium), to facilitate discovery and sharing of research resources. KLISC organizes workshops, seminars, and training conferences for members on emergent areas such as RDM. All academic libraries in Kenya are members of KLISC. One respondent stated that:

> *"We are currently collaborating with the research office, the ICT office, and the academic departments internally to fashion RDS that meet our users' needs. We are also partnering with external bodies such as KLISC and other Universities to build a wholesome package that is geared toward user needs." (RL1)*

Internationally, the library was collaborating with other organizations and institutions such as the World Confederation of Open Access Repositories (COAR), the Electronic Information for Libraries (EIFL), and the International Network for Availability of Scientific Publications (INASP) to offer RDS such as guides to the

research process, online consultative data services, and resource recommendations.

Asked about the benefits of collaborative partnerships, the interviewees generally mentioned resource recommendations that permitted navigation of external repositories and associated resources; advancing access, sharing, reusing data; access to a broader pool of funding, new research, and establishing research networks, and research cost cuts. One respondent commented that:

> *"Collaborative partnerships permit resource recommendation which enables navigation of external repositories advancing research resource sharing." (RL3).*

4.6 Challenges and Suggestions

4.6.1 Challenges Facing RDM Activities

Furthermore, the investigation asked respondents to mention what they thought were the challenges facing RDM implementation. The Interviewees generally mentioned below:

"Inadequate technological infrastructure"
"Lack of laws and policies"
"Lack of relevant skills and training"
"Inadequacies in collaborations"
"Insufficient funding"
"Unwillingness of researchers to share their data"
"Unethical data practices"

4.6.2 Suggestions to Address the Challenges

The interviewees were further requested to make suggestions on how best the challenges in 4.6.1 would be addressed. They generally cited:

"Hire, train, or reskill staff with RDM knowledge and skills"
"Make available the needed ICT infrastructure"
"Formulate robust and comprehensive RDM laws and policies"
"Create an enabling environment for RDM collaborations"
"Prioritize and increase RDM funding"
"Promote RDM advocacy support and training"
"Encourage researchers to share their data through incentives"
"Restructure the library's staff positions to data librarian, data curator, data steward, etc.",

5 Interpretation of Findings

5.1 Policy Guidelines

Despite the library having policies on research data, quality assurance, and Intellectual Property, study findings found no explicit policies to guide each phase of data curation and capabilities. These findings contradict those by Nhendodzashe and Pasipamire (2017), who acknowledge that policies are the cornerstone of RDM. They guide RDM operationalization. The policies also build stakeholders' confidence by ensuring research data retention, storage, and availability for use, reuse, sharing, or disposal in line with the ethical, legal, statutory, and donor requirements (Chigwada et al., 2017). Section 2 of Australia's Code for the Responsible Conduct of Research states that policies are necessary to guide ownership, storage, and preservation of research data, retention after project completion, and access convenience (Australian National Data Service, 2016). Therefore, in consultation with critical stakeholders, the management should develop and fully implement explicit RDM policies. These policies will guide each phase of data curation and capabilities, such as knowledge and skills, technological infrastructure, and collaborative partnerships.

5.2 Data Curation Activities

Study findings show that data was curated for value addition, discovery, sharing and reusing. The findings further divulged that interviewees seldom utilized the appraisal policy. UK Research and innovation ovation (2016) advanced that most personnel did not use the appraisal and selection policies when appraising data owing to deficiencies in knowledge and skillset. The fact that research librarians rarely utilized the appraisal policy due to their competencies suggests the need to equip research librarians with the knowledge and skills to appraise research data. Kim (2018) suggested the development of a standardized framework for appraising research data.

Second, the study findings established that the library generated metadata for probable purposes of privacy, consistency, simplicity and clarity in relationships (Tenopir et al., 2017). However, one major challenge of uniform metadata is the inability to display all the metadata, which may affect discovery, sharing, and reuse. Ricardo, João, João, and Cristina (2015) explain that metadata must be created at the point of data capture to guarantee richness.

Third, the findings denote that the library did not have clear structures on short term data storage. Chigwada et al. (2017) pointed out that it was critical to ensure standardized data storage in both the short term and long term. Therefore, the library should set up a coordinating unit to check standardization, quality, and safety of data on both short term and long. The UK Data Service (2020) advised that guaranteeing safe storage media for data necessitated consideration of the physical security, network security, computer systems, and files' security.

The findings indicated that key factors that affected sharing and reusing data included the ownership rights, sponsorship, etc. However, further inquiries suggested that data was rarely shared and reused because researchers were reluctant to avail their own data. Kurata, Matsubayashi, and Mine (2017) advance that researchers are usually unwilling to share data because of data abuse through unethical practices, deficiencies in supportive policies, technologies, and fear of losing competitive advantage. Consequently, it is necessary to address policy, privacy, and practicality issues for useful and practical sharing and reuse of data. Tenopir et al. (2020) encourage research data advocacy and training to advance access, sharing, and reuse of data.

5.3 Knowledge and Skills for RDM Activities

The study established that the research librarians had deficiencies in RDM skills and training. They needed training in data curation activities, metadata, Information technology tools and equipment, grant proposal writing, security and storage of data, DMP, legal policies, ethical considerations, and collaborative partnerships. According to Cox et al. (2017), training is considered a key component of RDM because it lays the ground for its implementation, management, and advocacy. Therefore, the library should organize RDM training through professional development working groups, in-house workshops, conferences, and short courses suited to work-based learning.

5.3.1 Research Data Services

Though the library offered a number of both information and technical data services, the study established that the services were unstructured, depicting that the library was still at the initial RDM implementation stages. These findings are corroborated by Dér (2015) who asserts that RDS in academic libraries of emerging economies are unstructured and immature. The study attributes the poor and rudimental services to lack of supportive policies and technological infrastructure, human resources, and insufficient funding. Flores et al. (2015) recommend establishing a division within the academic library to guide RDS implementation by assembling all the needed RDM capabilities.

5.4 Technological Infrastructure Readiness

Study findings evidenced deficiencies in the items of technological infrastructure capability. Deficiencies were apparent in computing facilities, software solutions, storage media, data backups and security solutions. For this reason, vast and robust technologies are necessary to support the whole research data lifecycle effectively. Similarly, the computing facilities and tools should have the capacity to support the storage and backup, preservation, and security of research data, including other RDM computing capabilities, such as interoperability, reliability, accessibility, and request responsiveness. According to the Helsinki University Library (2020) RDM computing technologies must be robust and effectual in data management and planning, description and documentation, security and ethical subjects, storage and backups, accessibility, and research data preservation.

5.5 Collaborative Partnerships Influencing RDM Activities

The study established deficiencies in RDM collaborative partnerships. In this regard, the library should move to foster collaborative partnerships internally, externally, and internationally. Internally, the library should engage stakeholders to develop policies to guide collaborations. The library should also explore other collaborations, including shared policies and documentation, workflows, and cost cuts with departments such as the information technology department, the research office, the legal office, and the academic departments. At the national level, the library, in consultation with the relevant authorities, should champion an RDM collaborative framework through legislation and policy to coordinate and guide joint activities. The framework should drive the creation of opportunities for engagement and mutual benefits, such as funding, scholarships, knowledge exchange, and job creation internally, at the national and international levels.

6 Conclusions

Academic libraries are moving from custodians to creators of knowledge resources through scholarly communication. This has particularly been typified in academic libraries of developed countries. However, study findings demonstrate that academic libraries in emerging economies are either lacking or still in the infant stages of RDM implementation. For instance, RDM activities at the UON Library faced challenges in all the examined capabilities.

Recommendation: The study recommends an organizational change, to develop an RDM unit within the UON Library to guide RDM implementation by assembling all the needed capabilities *(Policy guidelines, knowledge and skills, technological infrastructure, and collaborative partnerships)* to enable efficient managing, sharing, and reusing research data.

Implications: The study adds to existing literature from an emerging economy perspective. There is very little research on RDM in emerging economies, especially in the Kenyan setting. Similarly, the study has potentials of influencing implementation and efficient managing, sharing and reuse of data in Kenyan and other academic libraries of emerging economies

Future research: The study investigated the research librarians' experiences of RDM activities in an emerging economy. Future studies could involve other key RDM stakeholders such as the researchers, legal office, IT office and the research office to permit decisive and far-reaching conclusions.

References

Anduvare, E. M. (2019). *eResearch Support: An Exploratory Study of Private University Libraries in Nairobi County, Kenya*. (Doctoral Dissertation, University of KwaZulu-Natal). Retrieved from https://pdfs.semanticscholar.org/12db/fde09819dda35be1b99eba72d71d5fbdfbe1.pdf

Australian National Data Service. (2016). *Australian code for responsible research*. Retrieved from https://www.nhmrc.gov.au/guidelines-publications/r39

Block, E., & Erskine, L. (2012). Interviewing by telephone: Specific considerations, opportunities, and challenges. *The International Journal of Qualitative Methods, 11*(4), 428-445.

Braun, V. a. C., V. . (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*. *3* (2), 77-101.

Chigwada, J., Chiparausha, B., & Kasiroori, J. (2017). Research data management in research institutions in Zimbabwe. *Data Science Journal,* 16, 1–9.

Chigwada, J. P., Hwalima, T., & Kwangwa, N. (2019). A proposed framework for research data management services in research institutions in Zimbabwe. *Research Data Access and Management in Modern Libraries* (pp. 29–53). Hershey, Pennsylvania: IGI Global. doi:10.4018/978-1-5225-8437-7.ch002

Connell, J., Carlton, J., Grundy, A., Taylor Buck, E., Keetharuth, A. D., Ricketts, T., . . . Brazier, J. (2018). The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). *Qual Life Res, 27*(7), 1893-1902. doi:10.1007/s11136-018-1847-y

Cox, A., 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), 2182-2200. doi:https://doi.org/10.1002/asi.23781

Cox, A., & Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science, 46*, 299 - 316.

Crouch, M., & McKenzie, H. (2006). The logic of small samples in interview-based qualitative research. *Social Science Information, 45*(4), 483-499. doi:10.1177/0539018406069584

Dalhousie University Libraries. (2019). *Institutional research data management strategy*. Retrieved from https://dal.ca.libguides.com/rdm/repositories

Dér, Á. (2015). *Exploring the academic libraries' readiness for research data management: cases from Hungary and Estonia* (Master's thesis, Akershus University College of Applied Sciences). Retrieved from http://hdl.handle.net/10642/3367.

Dworkin, S. (2012). Sample size policy for qualitative studies using in-depth interviews. *Springer, 41*, 1319–1320.

Etikan, I., & Bala, K. (2017). Sampling and sampling methods. *Biom Biostat Int J., 5*(6), 215-217.

Flores, J. R., Brodeur, J. J., Daniels, M. G., Nichools, N., & Turnator, E. (2015). Libraries and the research data management landscape. *Council on library and Information resources*, 82 - 102.

Helsinki University Library. (2020). Research data managemennt. Retrieved from https://www.helsinki.fi/en/research/services-for-researchers-and-research-policy/research-data-management

Higgins, S. (2011). Digital curation: The emergence of a new discipline. *Int. J. Digit. Curation, 6*, 78-88.

Higgins, S. (2012). The lifecycle of data management. In G. Pryor (Ed.), *Managing Research Data* (pp. 17-46). London, United Kingdom: Facet.

Ingram, C. (2019). *How and why you should manage your research data: A guide for researchers: An introduction to engaging with research data management processes*. Retrieved from https://www.jisc.ac.uk/guides/how-and-why-you-should-manage-your-research-data

Janghorban, R., Roudsari, R. L., & Taghipour, A. (2014). Skype interviewing: The new generation of online synchronous interview in qualitative research. *International Journal of Qualitative Studies on Health and Well-being, 9*(1), 1–3. doi:10.3402/qhw.v9.24152

Josiline Chigwada, B. C. a. J. K. (2017). Research data management in research institutions in Zimbabwe. *Data Science Journal, 16*(31), 1-9.

Kim, K. (2018). *Issues in the Appraisal and Selection of Geospatial Data: A 2013 Report*. Retrieved from https://osf.io/qy47c/

Koltay, T. (2019). Data curation in academic libraries as part of the digital revolution. *Zagadnienia Informacji Naukowej - Studia Informacyjne, 57*, 28-36. doi:10.36702/zin.12

Kurata, K., Matsubayashi, M., & Mine, S. (2017). Identifying the complex position of research data and data sharing among researchers in natural science. *SAGE Open, 7*(3), 1–13.

Lyon, L., Ball, A., Duke, M., & Day, M. (2012). Developing a Community Capability Model Framework for data-intensive research. In R. Moore, K. Ashley, & S. Ross (Eds.), *iPres 2012: Proceedings of the 9th International Conference on Preservation of Digital Objects* (pp. 9-16). Received from https://ipres.ischool.utoronto.ca/sites/ipres.ischool.utoronto.ca/files/iPres%202012%20Conference%20Proceedings%20Final.pdf#page=19.

Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland Journal of Higher Education, 9*(3), 1–14.

Masinde, J. M., Wambiri, D. M., & Chen, J. (2020). Gender and cognitive factors influencing information seeking of graduate students at Kenyatta University Library. *South African Journal of Information Management, 22*(1), 1–10. doi:10.4102/sajim.v22i1.1154

Nacosti. (2020). *National Research Priorities 2018-2022*. Nairobi, Kenya: National Commision for Science, Technology and Innovation.

Ng'eno, E. J. (2018). *Research data management in Kenya's agricultural research institutes* (Doctoral dissertation, University of KwaZulu-Natal). Retrieved from http://hdl.handle.net/10413/15908

Nhendodzashe, N., & Pasipamire, N. (2017). Research data management services: Are academic libraries in Zimbabwe ready? The case of University of Zimbabwe library. *Proceedings of* IFLA WLIC 2017, Wrocław, Poland. Retrived from http://library.ifla.org/id/eprint/1728.

Ohaji, I. K. (2016). *Research data management: An exploration of the data librarian role in New Zealand research organisations*. (Doctoral dissertation, Victoria University of Wellington, Nieuw-Zeeland). Retrieved from http://hdl.handle.net/10063/5657.

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

Ray, J. M. (Ed.). (2013). *Research data management: Practical strategies for information professionals*. Purdue: Purdue University Press.

Ricardo, A., João, A., João, R., & Cristina, R. (2015). A comparative study of platforms for research data management: Interoperability, metadata capabilities and integration potential. *Advances in Intelligent Systems and Computing, 353*, 101-111. doi:10.1007/978-3-319-16486-1_10

Siyao, P. O., Whong, F. M., Martin-Yeboah, E., & Namamonde, A. (2017). Academic libraries in four Sub-Saharan Africa countries and their role in propagating open science. *IFLA Journal, 43*, 242 - 255.

Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., ... & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PloS one, 15*(3), 1–26. doi: 10.1371/journal.pone.0229003 M

Tenopir, C., Talja, S., Horstmann, W., Late, E., Hughes, D., Schmidt, B., . . . Allard, S. (2017). Research data services in European academic research libraries. *Liber Quarterly, 27*(1), 23–44.

Tessier, S. (2012). From field notes, to transcripts, to tape recordings: Evolution or combination. *The International Journal of Qualitative Methods, 11*(4), 446-460.

UK Data Service. (2020). *Managing and sharing research data: A guide to good practice*. Retrieved from https://www.ukdataservice.ac.uk/manage-data/store/storage.aspx

UK Research and innovation ovation. (2016). *RCUK policy and code of conduct on the governance of good research conduct*. Retrieved from https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-RCUKPolicyGuidelinesGovernanceOfGoodResearchConduct.pdf.

University of Nairobi. (2015). UoN is now ISO 9001:2015 certified. Retrieved from https://a.uonbi.ac.ke/category/tags/iso

University of Nairobi. (2020). *Student admissions*. Retrieved from https://a.uonbi.ac.ke/admission

U.S.News & World Report. (2020). *2020 Best global universities in Africa*. Retrieved from https://www.usnews.com/education/best-global-universities/africa

Woeber, C. (2017). *Towards best practice in research data management in the humanities*. Retrieved from http://researcharchive.vuw.ac.nz/handle/10063/6620

Yoon, A., & Schultz, T. (2017). Research data management services in academic libraries in the US: A content analysis of libraries' websites. *College & Research Libraries, 78*(7), 920–933.