Research Data Management Challenges in Kenya: The Case of Private Universities in Nairobi County.

By
Dr. Everlyn M. Anduvare
University of KwaZulu-Natal
Email: 217076206@stu.ukzn.ac.za
and
Prof Stephen M. Mutula
University of KwaZulu-Natal
Email: mutulas@ukzn.ac.za

Abstract

This research paper is a spinoff from a Doctoral degree study that was carried out at the University of KwaZulu-Natal between 2017-2019. The aim of the study was to establish the role private university libraries in Nairobi, Kenya play in supporting eResearch and the challenges thereof that librarians and researchers face in the process of managing data. The study employed both qualitative and quantitative epistemological approaches with semi structured interviews and survey questionnaires to collect data from a population consisting of university librarians, faculty members and doctoral students respectively. The population was sampled purposively. The qualitative and quantitative data sets were analysed using SPSS and content analysis respectively. The findings revealed several challenges, which included among others the lack of strategies and policies to guide research data management support, the lack of integrated RDM policies, a research process that was fragmented, and limited ICT policies and infrastructures. The institutionalisation of RDM in the private universities in Kenya is therefore urgent and imperative. The findings have policy, practical and theoretical implications for the effective RDM in Kenyan private universities in order to enhance scientific and scholarly communications. While the focus of the study limits generalisation of the findings, other universities may gain insights on RDM challenges within university settings.

Keywords

Research Data Management; RDM challenges; University Libraries; Kenya.

Introduction

Research data management (RDM) involves organising data from its entry to the research cycle through to disseminating and the archival of valuable results (Whyte & Tedds, 2011:1). RDM is, therefore, a critical part of the research process that aims at enhancing and making this process as efficient as possible (University of Leicester, 2017). It involves the planning, creation, documenting, organising, improving analysis procedures, securing, storing, backing up, providing access, and effective sharing of data to enhance publishing, citations and reusing of data for new projects (University of Leicester, 2017; Briney, 2015:17).

RDM is a moderately new term within the research arena having been established in the mid-2000s (Briney, 2015:13). The growth and development of RDM has mainly been as a result of funders and publishers mandates requiring that researchers ought to submit their generated raw data that has been used to report their findings, in open access data repositories to enable the public to access, browse, share, re-use and even validate reported research (Borghi, Abrams, Lowenberg, Simms & Chodacki, 2018:2; Tripathi, Shukla & Sonker, 2017:417; Ahlfeldt & Johnsson, 2015:12). Furthermore, there has been an increase in data sharing culture among researchers and an ongoing shift in policies that require not only open access to scientific publications, but also open access to research data leading to discussions on both national and international level about the significance of making research data publicly available and accessible, particularly from publicly financed research (Ahlfeldt & Johnsson, 2015:12).

According to Briney (2015:13), "such mandates gained momentum in the UK with the 2011 Common Principles on Data Policy from Research Councils UK (Research Councils UK 2011) and in the United States with the National Science Foundation's data management plan requirement in 2011 (NSF 2013)." Two key funders in the United States, that is, the National Science Foundation (NSF) and the National Institutes of Health (NIH) are provided as examples of organisations that require data sharing plans in order to provide funding (Zotus, 2017:291). This trend has caught up on a global scale. For instance, Kahn, Higgs, Davidson and Jones (2014:296) state that the National Research Foundation (NRF) which hosts the South African Data Archive (SADA) advocates for proper RDM especially to researchers who are receiving research funding. According to Chiware and Mathe (2015:3), in March 2015, the NRF in South

Africa “released a statement on open access for data retention, mandating that their funded research publications and supporting data be deposited in an accredited open access repository”. Thus, it is becoming evident that researchers must deposit their data in open access repositories necessitating increased support in RDM practices.

Significant benefits arising from RDM and open access data have been reported. Tripathi, Shukla and Sonker (2017:417) assert that properly stored data results into easy access, browsing, consultations, usage and building upon it in future for academic, research and scientific resolves. Additionally, data sharing enables researchers to reanalyse, re-evaluate and revalidate research findings enabling them to add their viewpoints which can enhance the creation of new knowledge. Ahlfeldt and Johnsson (2015:12) note that researchers can easily build on previous research results, thus, enhancing quality; collaborations can be nurtured to increase efficiency in research; reinventing the wheel can be curbed; acceleration of innovations and, increased transparency in the research process as there is citizen and society involvement.

According to Macquarie University (2016:8); Heidorn (2011:664); Henty, Weaver, Bradbury and Porter (2008:1), effective management of data enhances validation of the accuracy of results which can only be enhanced through direct access to original data; reproduction of data; progression of solutions; sharing of data; re-use of published data especially for distinct research problems; enhanced collaborations and communications among researchers; and, the possibility of unearthing unique data that cannot be replicated. Henty *et al*. (2008:1) add that, the expensive nature of data collection necessitates effort to manage it in order to circumvent duplication. Given the benefits and increasing funding mandates, RDM thus becomes essential in universities**.**

RDM in universities and the role of libraries

The management of research data has now posed a challenge for organisations. Universities are increasingly experiencing vast production of diverse born-digital data (Pinfield, Cox & Smith, 2014). For this data to be made available and useful, there is a need to format, document and organise it in a way that will enable it to be examined and reused (Borghi *et al*, 2018:2). The dramatic changes in the research landscape demands that researchers in universities change how

they work with and document their data, and as well, they are required to ensure accuracy, completeness and authenticity in their data (Ahlfeldt & Johnsson, 2015:18; Baykoucheva, 2015:72). As a result, researchers need support to appropriately manage their data. Lyon (2012:127) assert that libraries "have been positioned around a long-established publication process tailored to deliver the peer-reviewed scholarly article or monograph." Given the increasing mandates from funding agencies, university libraries have a greater responsibility to support researchers in the management of data and more increasingly in creation of data management plans (DMPs).

A study by Brown, Wolski and Richardson (2015:225) indicated that university libraries are confronted with the new roles for supporting research data. Ray (2014:6-7) reported that libraries had begun to pay attention to supporting data management in order to enable preservation and re-use of data and as well, to enable researchers to find their own data after the initial use. According to Borghi *et al*. (2018:2), library-based data management support to researchers has majorly been focused on "data management planning, metadata and documentation, data organization, storage and backup procedures, and long term preservation." Libraries could take more involvement in educating students and researchers on metadata creation but this could pose a challenge as most datasets have very few of these (Baykoucheva, 2015:81). A more focused approach for RDM could be for libraries to provide support across the data lifecycle. Shearer and Argaez (2010:3) state that to enhance long-term preservation, data ought to be created and maintained consistently and this would involve vigorous management of data in its entire lifecycle. Ahlfeldt and Johnsson (2015:15) emphasise this by stating that:

> *A data life cycle model is a process to describe the different stages and transformations that data will undergo from its creation to its final sharing and preservation. Using a data life cycle model can provide a useful framework to present and communicate the different stages of data, in order to deliver support for RDM in an organization. The process of research data management is often complex and it involves coordination between people, agencies and resources*.

However, Borghi *et al* (2018:2) and Briney (2015:17) allude to the fact that the focus can not only be on the data lifecycle but rather, given the wide range of practices in data management, there is a need to take steps in planning for RDM before the start of a research project or earlier

in a research process; during the project; and, after the completion. It is clear that there are enormous activities to be considered that are critical to the successful management of data in universities with university libraries being viewed as critical in these endeavours. According to Cox, Kennan, Lyon and Pinfield (2017), while institutional support for RDM has been drawn from libraries' previous involvement with digital services and open access endeavours, RDM has presented major challenges for library managers. It is evident that libraries have begun to offer RDM support but the nature and extent of their role remains blurred.

The purpose of this research paper is to present findings that the researcher sought in relation to problems of data management, organisation, dissemination and preservation that existed in six private chartered universities namely: Africa International University (AIU); Africa Nazarene University (ANU); Catholic University of Eastern Africa (CUEA); Daystar University; Pan Africa Christian University (PAC); United States International University (USIU). These universities were purposively selected because they met the requirements of the study which included: private chartered universities based in Nairobi County and offering PhD programs as the study targeted faculty members and PhD students (researchers). The study sought to explore private university libraries as they are reported to head other universities in adopting technologies (Herbling, 2012; Kavulya, 2003:156; Nganga, 2012; Otando; 2012:4), which is a critical enabler of RDM support services within an eResearch context. Table 1 highlights the status of the selected universities in terms of the schools or faculties in place, PhD programs on offer and the total population of the PhD students (308), faculty members (622), university librarians (6), reference librarians (13), and, the IR Managers (7). Furthermore, the availability of an RDM policy has also been indicated reflecting that only one university had one.

Table 1: Universities selected for the study

Universities	Schools/ Faculties	PhD programs	Population of researchers and librarians					RDM policy
			PhD Students	Faculty Members	University	Reference Librarians	IR Managers	
AIU	-School of Business and Economics -School of Education, Arts and Social Sciences -School of Theology	Doctor of Ministry Program PhD in Biblical Studies PhD in Business Administration & Management PhD in Education PhD in Intercultural Studies PhD in InterReligious Studies PhD in Leadership & Governance PhD in Systematic Theology PhD in Theology & Development PhD in Translation Studies PhD Theology and Culture	40	60	1	1	1	✗
ANU	-Business School -Law School -School of Humanities & Social Sciences -School of Religion and Christian Ministry -School of Science and Technology	PhD in Religion Doctor of Ministry.	35	50	1	1	1	✓
CUEA	-Faculty of Arts & Social Sciences -Faculty of Business -Faculty of Education -Faculty of Law -Faculty of Science -Faculty of Theology -School of Business	Doctor of Business Administration Doctor of Philosophy Counselling Psychology Doctor of Philosophy in Curriculum Studies and Instruction Doctor of Philosophy in Education Doctor of Philosophy in Education Planning and Administration Doctor of Philosophy in Religious Studies Doctor of Philosophy I Philosophy Doctor of Philosophy in Theology Doctorate in Sacred Theology PhD in Mathematics	141	237	1	4	1	✗
Daystar	-School of Arts & Humanities	PhD Communication PhD Clinical Psychology	20	120	1	1	1	✗

	-School of Business & Economics -School of Communication -School of Human & Social Sciences -School of Science, Engineering & Health -School of Law							
PAC	-Graduate School -School of Humanities & Social Sciences -School of Leadership, Business & Technology -School of Theology	PhD in Marriage & Family Therapy PhD in Organizational Leadership	25	35	1	1	1	✗
USIU	-Chandaria School of Business -School of Communication, Cinematic, and Creative Arts -School of Graduate Studies, Research and Extension -School of Humanities & Social Sciences -School of Pharmacy and Health Sciences -School of Science and Technology	Doctor of Business Administration Doctor of Psychology Doctor of Philosophy in International Relations	129	120	1	5	2	✗
			380	**622**	**6**	**13**	**7**	

Literature Review

The research question and literature for RDM challenges was guided by the element of data as derived from the eResearch Capability Model (eRCM) that was adopted from the Victoria University of Wellington (VUW). “Data is the management of all research inputs and outputs that are in a digital format. This includes the collection, curation, analysis, and provenance (metadata) of both basic data and information produced by research” (Whakamuri, Whakaaro & Aro, 2014:13). At VUW, it was found that due to lack of an organisational policy, researchers chose what to do with their data. The increasing pressure from the government to researchers to make their data accessible, and publishers asking for data to support work that researchers want to publish created a need for data management. It is based on the VUW study and the resulting

report by Whakamuri, Whakaaro and Aro (2014) that the current research question on RDM challenges was underpinned, particularly to seek gaps in RDM and, therefore, propose the way forward for effective RDM implementation.

Despite the depicted benefits, managing research data presents a range of challenges within university settings. Pinfield, Cox & Smith (2014:3) assert that RDM encompass a wide range of technical, cultural, managerial, legal and policy challenges. According to Yu (2017:793), academic libraries view the provision of research data services as a great additional service for its clientele but find the lack of formalised RDM infrastructure and policies, inadequate training for staff, unpreparedness and funding form a challenge to self-starter university wide RDM support. As a result, RDM-related services continue to be determined by funding agencies mandates.

Empirical studies have uncovered some challenges in relation to RDM. A recent study by Piracha and Ameen (2019) assessed the policy framework and planning in relation to RDM among 30 highly ranked universities by the higher education commission in Pakistan. The study established a lack of knowledge and awareness about RDM among library heads, lack of willingness, motivation and coordination by researchers, insufficient professional skills for RDM support, poor infrastructure and networking. Faniel and Connaway (2018) conducted interviews from 36 academic library professionals in the United States of America (USA) to establish librarians' perspectives on factors that influence RDM. It was established that technical resources, human resources, researchers' perceptions about the library, leadership support and communication, coordination, and collaboration influenced RDM activities. On the other hand, a survey carried out by Cox *et al*. (2017) carried out in higher education libraries in Australia, Canada, Germany, Ireland, the Netherlands, New Zealand, and the UK indicated that libraries had taken leadership roles to provide RDM support in universities but there was a greater focus on advocacy, policy development, advisory and consultancy services. Technical advancements to support RDM were found to be still wanting. Other concerns included inconsistency in terms of available skills to support RDM, resourcing, a lack of collaborations with other support services, and, challenges in involving key stakeholders like researchers and top management.

Lack of awareness and training among librarians and researchers have been presented as a hindrance to effective RDM. Borghi *et al*. (2018:3) finds a communication gap existing between researchers and library-based data service providers with a lack of user-friendly guides to enable researchers to advance their RDM practices. They propose that RDM should be integrated as part of a researcher's day to day activities but possible barriers such as language, terminologies and priorities among various research communities ought to be observed. Heidorn (2011:668) noted insufficient training for researchers towards long-term data access and preservation while indicating that libraries have the opportunity to actively engage in assisting researchers, failure to which, they may turn to other institutions to be offered the needed data management support. This can especially occur to researchers who require data management plans in order to secure grants with funding agencies. Baykoucheva (2015:80) reported that some academic institutions have considered the introduction of data management training for graduate students as mandatory with libraries taking the initiative to offer this training through workshops or integration in existing library instructions.

Appropriate skills and competencies for academic librarians will help enhance RDM support. Baykoucheva (2015:81); Cox and Pinfield (2013:301) identify the lack of technical knowledge, domain-specific expertise, and limited research experience as potential barriers to librarians taking up a critical role in RDM support. Lyon (2012:132) posits the need for librarians to have "a working knowledge of the research practices and workflows, an understanding of the specific technical standards, metadata schema and vocabularies routinely used in practice, an awareness of the national and international data centres where research data in that domain are deposited, and a good grasp of the data publication requirements of the leading scholarly journals". According to Heidorn (2011:667), having all the skills required to represent all the information and descriptions for data can be problematic for academic libraries because data extensively varies requiring a range of schemes to create appropriate metadata. This may require a close working relationship with data creators in order to understand their data and a need to work with other institutions in order to identify appropriate standards and practices for various datasets. Baykoucheva (2015) discussed the issue of data standards and noted that it is vital to have well established standards in order to appropriately describe data content and format. However, while there are established conventions for citing published papers, acceptable uniform standards for

research data are lacking. This poses a challenge to libraries as it is difficult to create new data standards. Furthermore, while some libraries have been able to adopt the use of institutional repositories to manage data, bibliographic metadata for datasets may vary.

Adopting effective data management requires that researchers and librarians make changes in how they handle data and this may be difficult since there is a cultural aspect that may hinder both individuals and organisations. Morgan, Duffield and Hall (2017:302) revealed that even in situations where the benefits of RDM are clearly known and appreciated, changing how people do things is challenging. For instance, researchers may not be willing to invest more time in the processes required to ensure that their data is well managed. Consequently, this calls for continuous engagement as the pace of change and adoption to RDM is not instantaneous. Deninson, Kethers and McPhee (2007:9) also expressed the same concern by asserting that researchers regard themselves as time poor and would want services that are interoperable with their usual work practices and technologies. Therefore, there is a tendency to avoid non-core tasks unless they are proven to have considerable benefits. In relation to librarians, Cox and Pinfield (2013:300) point to the increased staff time required on librarians to provide this service in the midst of already over-stretched library services. Supporting RDM may push libraries to downgrade other services and a top this, there is still instability in terms of infrastructure, policy, and governance which leaves the library at a stalemate in terms of positioning itself to support RDM.

It is apparent that advances in technology have hastened the amount of data produced, accessibility, analysis and, data protection (Mackie & Bradburn, 2000:2). In spite of this, technological challenges have arose. Briney (2015:14) indicates that digital files are fragile, thus, notes, problems such as corruption of storage devices, losing files, and obsolescence may easily be experienced. According to Baykoucheva (2015:73), preservation and storage present big challenges. He reports that surveys of researchers have indicated that their research data is most often on spreadsheets which limits manipulation and furthermore, the data is stored on computers and external hard drives without having backup. In addition, “a whole range of other activities commonly associated with datasets, such as reformatting them for analysis in various software packages, shipping them between sites, processing them for potential reuse, and carrying out

various preservation actions upon them" as highlighted by Cox and Pinfield (2013:299) would need attention. Technological infrastructure is, therefore, a challenge in itself and requires careful planning to avoid loss of data.

Ethical issues have also come into play as a challenge in RDM. Data about humans may raise privacy concerns while some data may be classified by a nation for security concerns, thus, requiring a library to apply appropriate access controls (Heidorn, 2011:668). "Data may be sensitive, containing personal information for example, and so needs to be managed with appropriate security measures in place" (Cox & Pinfield, 2013:299). For data that has to be stored externally, there have been concerns with regard to the level of trust that can be placed on external agencies to be in charge of the long-term preservation of data (Lewis, 2010:11). Trust in the technology used in terms of being reliable and stable is also of concern to researchers. They want to be able to trust the organisation that is managing their data, and as well, that the research community will not "misuse, alter, or steal the data" hence, libraries managing these data have to ensure that there is sufficient security in order to build trust in the systems and infrastructures being used in RDM (Denison, Kethers & McPhee, 2007:9). Briney (2015) notes that while researchers are not expected to be security experts, they have a role to play hence need to have the basics of security, ensure protection of data that has been entrusted to them and, ensure that data is always stored securely in an environment that is controlled.

Luce (2008) noted that adequate and sustainable funding is crucial in RDM as posed in the question below:

> *The cost of owning and managing data, hardware, and software is very high. How do we offset and share multi-institutional infrastructure investments? Because it takes a community to meet these challenges, how many research libraries need to work together to meet specific eResearch needs, and how do we collaborate in new, more effective ways*?

It is clear from this literature that there are vast challenges for RDM that would impact the entire university community. These challenges would require collaborative effort both at the national and international levels. According to Levine (2014:129), there exists a gap between aspiration

and reality as there are complexities in terms of making data available and usable in that all questions are yet to be figured out.

Research Methodology

The present study sought to find out the challenges relating to RDM in private university libraries in Nairobi County, Kenya. Six private chartered universities consisting of Africa International University, Africa Nazarene University, the Catholic University of Eastern Africa, Daystar University, Pan Africa Christian University, and the United States International University were selected for the study. The study employed both quantitative and qualitative epistemological approaches. Self-administered questionnaires were used to collect data from a population consisting of 306 PhD students, 462 Faculty members, 13 Reference Librarians, and 7 Institutional Repository (IR) Managers which had been achieved using survey monkey sample size calculator at a confidence level of 95% and a margin of error of 5%. The survey questionnaires integrated both closed and open-ended questions. Three sets of questionnaires were developed for i) PhD students and Faculty members who were provided with closed-ended questions on RDM challenges and furthermore, were asked to specify any other challenges they faced; ii) Reference Librarians whose open-ended question sought to establish challenges in the provision of RDM services, and iii) IR Managers whose question on RDM was open-ended and sought to establish what RDM challenges the libraries experience.

Semi-structured interviews were used to collect data about RDM challenges and barriers faced by the libraries from 6 University Librarians. A response rate ranging between 71%-92% was achieved. The population was sampled purposively targeting librarians as providers of eResearch support, as well as Faculty and Doctorate students because they are most likely to be involved actively in research. The quantitative and qualitative data sets were analysed using SPSS and content analysis respectively.

Research Findings

The findings are based on data collected in 2018 for a doctoral thesis. The findings from the PHD students and Faculty members indicated that privacy and confidentiality of research data was a challenge to 328 (53.1%) respondents, 301 (48.7%) indicated that they have challenges in

creating metadata, 299 (48.4%) are facing difficulties when it comes to locating of datasets and, 254 (41.1%) find data storage as a challenge. An average 52% of the respondents did not confirm if they were affected by these aspects. Furthermore, while respondents were given a chance to provide additional RDM challenges, none was provided. Table 2 shows the results.

Table 2: RDM challenges faced by PhD students and Faculty members (n=618)

RDM challenges	Selected (%)	Not selected (%)
Privacy and confidentiality issues associated with research data	328 (53.1)	290 (46.9)
Creating metadata	301 (48.7)	317 (51.3)
Locating datasets	299 (48.4)	319 (51.6)
Storage of data	254 (41.1)	364 (58.9)
Total % of cases	**1182 (191.3)**	**1290 (208.7)**

*Multiple responses possible

Using open-ended survey, Reference Librarians were asked to indicate the challenges they face when providing research data management services. Table 3 presents the findings.

Table 3: Challenges faced by Reference Librarians when providing RDM services (n=11)

University	Respondent	Responses
A	RL1	"We have not engaged in formal RDM services"
A	RL2	"Insufficient access"
A	RL3	"Some databases are a bit complicated and need much knowhow, the internet is guaranteed and the computer software and hard disk are not current for fast processing or downloading of work"
A	RL4	"Lack of administrative and academic support from the institutions. Lack of finances - cannot attend trainings off-campus, economic strains of institution who's going to pay for archiving and access, adopting new technologies etc"
B	RL5	"Inadequate staff and skills, lack of willingness from the researchers to share".
C	RL6	"Dealing with research students who don't understand what technologies they need to use; Finding time to work consistently with research centre thus creating a gap on any new concepts"
C	RL7	"Stereotypes, work overload, one may not deliver in time; lack of skill and knowledge; librarians could lack time to keep abreast of new tech; there are no policies around RDM; we may not see it as our work"
C	RL8	"Training, [it is a] new field - institutions of higher learning should

		develop a curriculum"
C	RL9	"Network failure"
D	RL10	"A lot of consultation back and forth in order to establish suitable RDM policies acceptable across the universities; poor attendance to RDM training session offered to researchers"
F	RL12	"It is a full-time job" and "equipment tools are inadequate".

** RL = Reference Librarian*

The IR Managers were asked to indicate the challenges that the library experiences while managing data. Three of them cited the following:

"*Copyright issues, plagiarism, fear of data being copied [and] awareness*" (IR2).

"*Poor attendance to RDM training sessions offered. A lot of consultations back and forth so as to establish suitable RDM policies acceptable across the universities*" (IR4).

"*Information overload*" (IR5).

The University Librarians were also interviewed and asked to outline any research data management challenges and barriers that their libraries face. The majority indicated not to have RDM but cited potential challenges as highlighted in Table 4.

Table 4: RDM challenges outlined by University Librarians (n=6)

University	**University Librarians**	**RDM challenges**
A	UL1	"Developing the right skill set especially in RDM so that we can be able to provide adequate services" "Adaptation of e-publishing even within the university. Adaptation of OA publishing" "Lack of awareness among researchers and students where they can publish their work quality areas or quality publishers. Sometimes you see areas they have published in their works are really journals that are not peer-reviewed" "…what we need is how to create awareness, how to sensitize, how to upgrade on knowledge of how to go about the eResearch"
B	UL2	"I think availing of the data for research is one of the challenges" "Lack of understanding what it is all about and why do I have to give my own data, what for? Maybe someone will be worried that if I give it then someone is going to use it" "The issue of space for data storage and we'd need to look for external servers for high level securing of data" "Then of course budget".
C	UL3	"I think researchers are protective of their data. I do not see how they would want to keep it in a library" "Once somebody has got the data and analysed it, most of the time it

		is not stored anywhere. Even mine if you ask me I don't know where it went. So there is that challenge of thinking that it is not important once you have already used it"
D	UL4	"I think the issue of attitude" "Probably lack of support, I mean we'll even have to lobby the university management and I would imagine some platforms that would call for money and the library budgets are so slim" "There is a gap in terms of skills. Your typical librarian may not be able to hack some concepts. I believe that there are some desires to competencies so for the typical librarian there will be a need for capacity building"
E	UL5	"Need for more awareness and advocacy"
F	UL6	"...I think maybe technological in terms of capacities unless we build technological capacities that will be able to store such data, but not only to store but also to make it more available when required..."

**UL = University Librarians*

Discussion

The findings in Table 2 indicate that 53.1% of the PhD students and Faculty members faced privacy and confidentiality data dilemmas which are ethical-related challenges associated with RDM. Studies by Cox and Pinfield (2013:299); Heidorn (2011:668); Lewis (2010:11); Denison, Kethers & McPhee (2007:9) indicate ethics as a challenge in RDM citing privacy issues of participants, security, sensitivity of data and trust. Consequently, ethical issues have to be observed at all the stages from data collection to preservation and re-use if researchers' trust in RDM has to be earned. The issue of ethics was also highlighted by a University Librarian (UL2) who said that, "*I think availing of the data for research is one of the challenges.*" UL2 reiterated, a "*lack of understanding what it is all about and why do I have to give my own data, what for*? *Maybe someone will be worried that if [they] give it then someone is going to use it*".

The PhD students and Faculty members also indicated to have challenges in metadata creation (301, 48.7%), locating datasets (299, 48.4%) and, data storage (254, 41.1%) as shown on Table 2. While the research study established that some libraries supported researchers in metadata creation, it also revealed that majority of them were not aware of that service in the particular libraries. This could have been attributed to lack of awareness and also, the lack of a policy as this service was provided by individual librarians in an ad-hoc manner. Baykoucheva (2015:81) advises that libraries take a more proactive role in educating students and researchers on

metadata creation. With regard to locating datasets, 48% of the respondents indicated this was a challenge while the rest (52%) could not even determine this as a service provision. Furthermore, the study established a gap in data storage. These challenges can be attributed to the lack of data storage facilities in the libraries as expressed by the IR Managers. In general, more than 46.9% of the respondents could not establish RDM challenges possible due to the lack of conscious RDM practices both by the researchers and their libraries.

From the library perspective, the University Librarians indicated not to have formal RDM in the libraries. Regardless, the study established that some minimal RDM support such as data entry and analysis were provided by individual IR Managers and Reference Librarians. A Reference Librarian confirmed the lack of formal RDM by saying that *"we have not engaged in formal RDM services"* (RL1). Nevertheless, Reference Librarians from five universities (A, B, C, D and F respectively) identified challenges to RDM to include: insufficient access to data, poor technological infrastructure, lack of university support, lack of funding, inadequate skills for librarians, unwillingness to share data by researchers, lack of awareness among researchers, lack of RDM policies and lack of curriculum on RDM. On the other hand, two of the IR Managers from universities B and D respectively identified challenges to managing data as: ethical issues, lack of trust, awareness and interest from researchers and lack of policies (See Table 3). University Librarians on the other hand cited likely RDM challenges as: the lack of knowledge about RDM and appropriate skill set among librarians, lack of awareness among researchers, inaccessibility of data, lack of storage space, lack of funding and institutional support, trust, attitude and inappropriate technological infrastructure (See Table 4).

The findings from the researchers and library staff are in consistent with the RDM challenges raised in the literature reviewed in this study. For instance, insufficient time for data management (Pinfield, 2013:300; Deninson, Kethers & McPhee, 2007:9); lack of training (Borghi *et al*., 2018:3; Yu, 2017:793; Heidorn, 2011:668); inadequate library support, ethical and storage concerns (Baykoucheva, 2015:81; Briney, 2015; Cox & Pinfield, 2013; Heidorn, 2011; Lewis, 2010:11) and technological challenges (Baykoucheva, 2015:73; Briney, 2015; Denison, Kethers & McPhee, 2007:9). In overall, the findings unearthed several RDM challenges in the universities studied that included: creation of metadata; locating datasets; data storage; lack of

RDM strategies and policies; insufficient access to data; poor technological infrastructure; lack of funding and institutional support; inadequate skill set among librarians; unwillingness to share data by researchers; lack of awareness and knowledge of RDM; lack of curriculum on RDM; lack of trust; lack of interest from researchers; inaccessibility of data; attitude and, inappropriate technological infrastructure. Evidently, this study has established that researchers are met with a wide range of challenges that are impeding RDM practices at the universities. Consequently, the university libraries are faced with the challenge of RDM majorly due to lack of strategies.

Conclusions and recommendations

Research data management is increasingly becoming crucial in universities with libraries being tasked to provide RDM support. Private universities in Nairobi, Kenya are yet to have formalised RDM support services for the university community, hence, the university libraries did not have established RDM support services. This was clear in the responses provided with an average 52% of the researchers not indicating if they were supported while none of the 618(100%) of the researchers provided an answer when they were provided with an open end question to indicate RDM challenges. University Librarians and Reference librarians confirmed this status. Despite this, minimal support for some RDM aspects were established including support in data entry and analysis but this was insufficient compared to the vast data practices that cut across the research data lifecycle. The findings revealed several challenges, which included among others the lack of strategies and policies to guide research data management support, the lack of integrated RDM policies, a research process that was fragmented, and limited ICT policies and infrastructures, ethical dilemmas and, lack of awareness and training. In view of these findings, it can be concluded that the management of RDM at the six private universities will remain an obstacle to eResearch if they are not addressed. The study proposes firstly that librarians be made aware and trained on all aspects of RDM to enable their understanding of RDM and thereafter, enhance their ability to set up appropriate strategies and policies, RDM support services and, training and support for researchers. The study finds that the institutionalisation of RDM in the private universities in Kenya is urgent and imperative.

References

Ahlfeldt, J. & Johnsson, M. 2015. *Research libraries and research data management within the humanities and social sciences*. Available at: https://portal.research.lu.se/portal/files/6286782/5050466.pdf (Accessed 31 May 2019).

Baykoucheva, S. 2015. *Managing scientific information and research data*. Amsterdam: Chandos Publishing.

Borghi, J., Abrams, S., Lowenberg, D., Simms, S. & Chodacki, J. 2018. Support your data: a research data management guide for researchers. *Research Ideas and Outcomes*, 4, p. e26439.

Briney, K. 2015. *Data management for researchers: organize, maintain and share your data for research success*. Exeter: Pelagic Publishing, UK.

Brown, R.A., Wolski, M. & Richardson, J. 2015. Developing new skills for research support librarians. *The Australian Library Journal*, 64(3):224-234.

Chiware, E & Mathe, Z. 2015. Academic libraries' role in research data management services: a South African perspective. *South African Journal of Library and Information Science*, 81(2):1-10.

Cox, A.M. & Pinfield, S. 2013. Research data management and libraries: current activities and future priorities. Journal of Librarianship and Information Science, 46(4): 299-316.

Cox, A.M., Kennan, M.A., Lyon, L. & Pinfield, S. 2017. Developments in research data management in academic libraries: towards an understanding of research data service maturity. *Journal of the Association for Information Science and Technology*, 68(9).

Denison, T., Kethers, S. & McPhee, N. 2007. Managing the soft issues in eResearch: a role for libraries? *Australian Academic & Research Libraries*, 38(1):1-14.

Faniel, I.M. & Connaway, L.S. 2018. Librarians' perspective on the factors influencing research data management programs. *College & research Libraries*, 79(1):100.

Heidorn, P.B. 2011. The emerging role of libraries in data curation and e-science. *Journal of Library Administration*, 51(7-8):662-672.

Henty, M., Weaver, B., Bradbury, S. & Porter, S. 2008. *Investigating data management practices in Australian universities. Report of Australian Partnership for Sustainable Repositories Online Research Collections Australia Support Network*. Available at: http://www.apsr.edu.au/orca/investigating_data_management.pdf (Accessed 21 May 2019).

Herbling, D. 2012. *Strathmore leads E Africa varsity in ICT ranking*. Available at: http://www.businessdailyafrica.com/news/Strathmore-leads-E-Africa-varsities--in-ICT-ranking---/539546-1611554-1a5gcez/index.html (Accessed 29 April 2017).

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.

Kavulya, J.M. 2003. *University libraries in Kenya: A study of their practices and performance* [Dissertation]. Available at: http://edoc.hu-berlin.de/dissertationen/kavulya-joseph-muema-2004-02-19/PDF/Kavulya.pdf (Accessed 30 March 2017).

Levine, M. 2014. Copyright, open data, and the availability-usability gap: challenges, opportunities, and approaches for libraries. In: Ray, J.M. (ed.). 2014. *Research data management: practical strategies for information professionals*. West Lafayette, Indiana: Purdue University.

Lewis, M. 2010. Libraries and the management of research data. In: McKnight, S. (ed.) *Envisioning future academic library services*. London: Facet Publishing.

Luce, R.E. 2008. A new value equation challenge: The emergence of eResearch and roles for research libraries. *CLIR*. Available at: https://www.clir.org/pubs/reports/pub142/luce.html (Accessed 28 April 2017).

Lyon, L. 2012. The informatics transform: re-engineering libraries for the data decade. *The International Journal of Digital Curation*, 7(1):126-138.

Mackie, C. & Bradburn, N. 2002. *Improving access to and confidentiality of research data: report of a workshop*. Washington, DC: National Academy Press.

Macquarie University. 2016. *Data Science and eResearch Platform Strategy*. Available at: https://staff.mq.edu.au/research/strategy-priorities-and-initiatives/data-science-and-eresearch/Data-Science-and-eResearch-Platform-STRATEGY.pdf (Accessed 21 May 2019).

Morgan, A., Duffield, N. & Hall, L.W. 2017. Research data management support: sharing our experiences. *Journal of the Australian Library and Information Association*, 66(3):299-305.

Nganga, G. 2014. *PhD to be the compulsory qualification for lecturers*. Available at: http://www.universityworldnews.com/article.php?story=20141030132504527 (Accessed 04 September 2018).

Otando, R.M. 2012. Buiding institutional repositories in KLISC member institutions in Kenya: Current status and emerging challenges. In: *INASP Newsletter*, issue 47. Available at: http://www.inasp.info/uploads/filer_public/2013/04/09/inasp_newsletter_47.pdf (Accessed 29 April 2017).

Pinfield, S., Cox, A.M. & Smith, J. 2014. Research data management and libraries: relationships, activities, drivers and influences. *Plos ONE*, 9(12):1-28.

Piracha, H. A. & Ameen, K. 2019. Policy and planning of research data management in university libraries of Pakistan. *Collection and Curation*, 38(2):39-44.

Ray, J.M. 2014. Introduction to research data management. In: Ray, J.M. (ed.). *Research data management: practical strategies for information professionals.* West Lafayette, Indiana: Purdue University Press.

Shearer, K. & Argaez, D. 2010. *Addressing the research data group: a review of novel services for libraries*. Ottawa: Canadian Association of Research Libraries.

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.

University of Leicester. 2017. *What is research data management?* Available at: https://www2.le.ac.uk/services/research-data/rdm/what-is-rdm (Accessed 28 May 2017).

Zozus, M. 2017. *The data book: collection and management of research data*. Boca Raton: CRC Press.

Whakamuri, E.K.A., Whakaaro, R. & Me Aro, M.T. 2014. *eResearch at Victoria: building our research with ICT.* Available at: http://www.victoria.ac.nz/its/staff-services/eresearch-capabilities/eResearch-Report-Final.pdf (Accessed 13 March 2017).

Whyte, A. & Tedds, J. 2011. *Making the Case for Research Data Management. DCC Briefing Papers. Edinburgh: Digital Curation Centre*. Available online: http://www.dcc.ac.uk/resources/briefing-papers (Accessed 28 May 2017).

Yu, H.H. 2017. The role of academic libraries in research data services (RDS) provision: opportunities and challenges. *The Electronic Library*, 35(4): 626-649.