Research data management literacy amongst lecturers at Strathmore University, Kenya

Fredrick Odhiambo Adika
Strathmore University, Nairobi, Kenya, and
Tom Kwanya
Department of Information and Knowledge Management, The Technical University of Kenya, Nairobi, Kenya

Received 11 March 2020
Revised 4 May 2020
Accepted 5 May 2020

Abstract

Purpose – The purpose of this study was to analyse the skills required by lecturers to be able to support research data management effectively; assess the research data management literacy levels amongst lecturers at Strathmore University; and suggest how research data management capacity can be strengthened to mitigate the knowledge gaps identified.

Design/methodology/approach – This study was conducted as a mixed methods research. Explanatory sequential mixed methods approach was used to collect, analyse and interpret quantitative and qualitative data from lecturers at Strathmore University in Nairobi, Kenya. Quantitative data was collected using questionnaires while qualitative data was collected through focus group discussions. Quantitative data was analysed using SPSS while qualitative data was analysed thematically.

Findings – The findings of this study indicate varied levels of research data management literacy amongst lecturers at Strathmore University. Lecturers understand the need of having literacy skills in managing research data. They also participate in data creation, collection, processing, validation, dissemination, sharing and archiving. This is a clear indication of good research data management. However, the study also revealed gaps in research data management skills amongst the lecturers in areas such as sharing of research data on open access journals, data legislation and securing research data.

Research limitations/implications – The study has been conducted in one university in Kenya. However, the findings have been contextualised in the global landscape through suitable references.

Practical implications – The findings of this study may be used to attract the attention of lecturers and librarians to research data management. The findings may also be used to develop institutional policies on research data management at Strathmore University and beyond. The suggested ways of research data capacity strengthening can be adopted or adapted by other universities to enhance research data management.

Originality/value – This is an original study.

Keywords Research data management, Research data management literacy, Strathmore university, Kenya, Africa

Paper type Research paper

1. Introduction

Research data management (RDM) is an umbrella term used to describe activities related to the creation, organisation, structuring, naming, backing up, storage, conservation and sharing of data as well as all actions that guarantee research data security. It aims to ensure reliable verification of results and permits new and innovative research built on existing information (Whyte and Tedds, 2011). RDM consists of different activities and processes associated with data creation, storage, security, preservation, retrieval, reuse and sharing taking into account technical capabilities, ethical considerations, legal issues, human resource capability and government frameworks (Ng'eno and Mutula, 2018). Research data, as one part of the scientific output must be understood in a broad sense as the recorded factual material commonly accepted in the scientific community as necessary to validate research findings. Research data are complex, dynamic, living and easier to describe than to

Library Management

0143-5124
DOI 10.1108/LM-03-2020-0043

define objects with characteristics changing along with the research process (Schöpfel *et al.*, 2018). The term data refers to facts and statistics collected together for reference and analysis. From an information science perspective, data can be defined more contextually in the scope of research to mean that which is collected, observed or created for purposes of analysis to produce original research results (Briney, 2015). Research data is any information that has been collected, observed, generated or created to validate original research findings (Ray, 2014).

RDM is a method that enables the integration, curation and interoperability of data created during the scientific process. RDM involves the production, access, verification, persistent storage and reuse of research data with the help of adequate and easy-to-use tools in virtual research infrastructures. Data constitutes the essential part of the curation cycle that comprises the conceptualisation, creation or reception, appraisal, selection, ingestion, preservation, storage, access, use and reuse and transformation of research data (Cox and Verbaan, 2018). Research data literacy refers to the human competence to locate, analyse, organise, present and evaluate research data in its context (Herzog, 2016).

Academic staff in universities work in increasingly complex research environments where advances in technology and research methodologies result in the gathering and analysis of large amounts of data. Proper management (organisation, protection, preservation, sharing) of this research data is essential for productivity, securing grant funding, enabling collaboration and ensuring the future use of data (Briney, 2015). The professionals working in data science, data librarianship and data management are a "new breed" for whom the knowledge and skills requirements are just emerging (Davenport and Patil, 2012; Provost and Fawcett, 2013). Organisations need data librarians and managers to enable better use, management, curation, preservation of data and to explore data reuse, aggregation and sharing. This need is evident in universities and research organisations as well as in business and government organisations (AGIMO, 2013; Corrall *et al.*, 2013).

There are many benefits associated with effective RDM practices. These include the advancement of scholarly research premised on a reliable and complete record of previous research. Creation of an RDM plan that allows for the re-use of publicly funded data is a necessary condition in many granting agency funding requests. This is aimed at protecting the enormous financial and time investments by mitigating data loss and avoiding the need for duplication of efforts to recreate lost data, validate existing research and create new knowledge by accessing and building on the work of others (Briney, 2015).

2. Literature review

Studies on research behaviour, competencies and attitudes related to RDM practices provide useful background information and context for our study. The themes covered in this literature review include research data production, characteristics of research data, research data sharing, storage and curation of research data and RDM skills. The authors also use this section to analyse gaps in literature which necessitated this study.

2.1 Research data production

Data is the currency of academic research (Koopman, 2015). According to Denny *et al.* (2015), data is described by researchers as the lifeblood of their work because it is inextricably connected to their research outputs, which are linked to publications, which in turn are linked to future funding. Jahnke and Asher (2012a, b) explain that readily available digital devices have made data creation easier. This has led to the production of vast volumes of research data. Denny *et al.* (2015) explain that the abundance of research data that exists today has enormous potential to unlock future advances in science while Chawinga and Zinn (2019)

assert that this data is the engine which propels scientific progress. From the foregoing arguments, the increased production of data is essentially good for research. However, the existing high volumes and varieties of data in the infosphere have created immense challenges around research data collection, curation, preservation and re-use. Indeed, Koopman (2015) reports that most of the existing research data has been largely invisible and do not contribute effectively development. As Kahn *et al.* (2014) argue, the need for data-driven research is increasing across most disciplines globally. Koopman (2015) suggests that research stakeholders can exploit emerging developments in information and communication technologies to improve not only data production but also its uptake by other researchers and users.

2.2 Characteristics of research data

According to a study conducted by Patterton *et al.* (2018) in a number of research institutes in Sub-Saharan Africa found that experienced as well as emerging researchers made use of many different data types, both groups totalling more than 15 types. The most common data formats across both groups were spreadsheets, image files and textual data formats. The study also revealed that a wide range of data volumes were used within the institute, with datasets ranging from less than 1 GB to datasets bigger than 100 TB.

In Kenya, a study by Kibe *et al.* (2020) analysed the characteristics of data produced by the Technical University of Kenya and Strathmore University. They concluded that staff in the two universities, both based in Nairobi, produced what can be described as big data since the data was produced in large volumes, diverse formats (variety) and fast (velocity). The volume of data produced in the two institutions ranged between several gigabytes and terabytes. In terms of variety, the data was both structured and unstructured and generated from e-mail, photos, videos and audio, social media, MS Office applications, cell phones, financial transactions, website content, blogs, gaming and related applications, web logs, click stream and GIS utilities. In terms of velocity, the study found that the two universities process big data using real-time, periodic, batch and near-time approaches. Real-time processing technology captures, processes and responds to big data as the events generating that data are happening in the real world. It deals with a continuous stream of inputs and has strict deadlines for completing the tasks. It is important to point out, however, that this study did not focus on research data but all categories produced in the two institutions. In spite of this limitation, it provides a glimpse at the characteristics of data, including research data, in the two institutions.

2.3 Sharing of research data

The bulk of the RDM literature reviewed was on research data sharing. This is an indication of greater interest on this RDM theme perhaps because of the difficulties encountered in realising effective sharing of research data. This is in spite of the understanding that researchers who share their data stand to benefit from higher citations and visibility (Van Noorden, 2014). It seems that many researchers are not convinced about the benefits of sharing their data (Koopman, 2015). Indeed, Chawinga and Zinn (2019) point out that despite resolute efforts to promote data sharing, relatively little data is shared. According to Koopman (2015), data sets of sufficient quantity and quality to answer research questions can take a researcher a lifetime to accumulate. She argues that this may explain why some researchers hesitate to share their data openly without the assurance that their work will be adequately acknowledged. Denny *et al.* (2015) explain that the potential for shared data to be misused, misunderstood and produce false conclusions that threaten the integrity of the primary research is also a deterrent to data sharing. In spite of these concerns, Anane-Sarpong *et al.* (2018) argue that research data sharing is inevitable and urge adequate preparedness.

Denny *et al.* (2015) argue that the need to protect data for its publication value is a key deterrent to releasing data as some researchers are worried that data would lose their value once placed in the public domain. Chawinga and Zinn (2019) identified the factors which frustrate data-sharing efforts to include lack of time and data misappropriation at the individual level; data-sharing training, absence of compensation and unfavourable internal policies at the institutional level; and weak policies, ethical and legal norms, lack of data infrastructure and interoperability issues at the international level. Denny *et al.* (2015) add that some researchers also feel like sharing research data may result in some form of neo-colonialist behaviour where the raw materials are taken out of the country to benefit foreigners. According to Anane-Sarpong *et al.* (2018), impediments to effective research data sharing include risks faced by under-resourced scientists and institutions which are slower in translating data produced into new knowledge; the absence of a harmonised guideline and structures to help address the risks and institute fairness in data-sharing rewards; and inadequate confidence in available protective safeguards including guidelines.

According to Patterton *et al.* (2018), there is a high willingness to share data with peers. This can be used to build confidence in researchers to promote seamless data sharing within teams before venturing outside. Chawinga and Zinn (2019) suggest a number of strategies which can be used to enhance research data sharing. These include recognising researchers who share data through data citations, acknowledgement and incentives; investing in infrastructure, conducting training and advocacy programmes; and formulating stringent and fair policies for data sharing can enhance it. Anane-Sarpong *et al.* (2018) suggest that data commodification via fee-for-use arrangements may be a possible solution to funding shortfalls in under-resourced contexts. They also assert that data-sharing deliberations need to shift from the focus on access to data to considering the whole gamut of people and processes that make data possible.

Ng'eno and Mutula (2018) argue that while the rest of the world have embraced research data management, Africa is lagging behind. However, Patterton *et al.* (2018) argue that researcher RDM behaviour is similar across the globe. Nonetheless, according to Pisani *et al.* (2016), research data sharing in Africa is constrained by the fear of loss of control once data is shared; sub-optimal gains to those who create and manage data; undue advantages to more technologically resourced contexts because of technological imbalances and skillsets in their favour; and technical issues including data quality, interoperability, and risks of misinterpretation due to unfamiliarity with data-originating contexts.

Denny *et al.* (2015) conducted a study in South Africa which revealed that data-sharing practices amongst researchers as either "ad hoc" decisions (post study) and informal practices of exchange between colleagues and interested persons. They also found that data sharing was a consequence of formal procedures, enforced by institutional policy in the form of contractual agreements between the principal investigator, their home institution and the funding body. In Kenya, Ng'eno and Mutula (2018) explain that most institutions have limited capacity, resources and facilities for collection, analysis, use and reuse and sharing of migration data hence making access and use of migration data difficult.

2.4 Storage, curation and preservation of research data

Koopman (2015) explains that most researchers do not store their data on university archives or libraries but instead prefer to store their own data through other means (special collections). Consequently, as Patterton *et al.* (2018) opine, researchers are concerned about possible data loss, accidental data deletion, data corruption, encryption problems and data loss due to equipment theft. Jahnke and Asher (2012a, b) explain that the concerns are exacerbated by the fact that technology moves so quickly that there is a real possibility that data collected today will not be readable or even accessible in the future. Indeed, there is

global awareness that digital data are in danger of being lost (Computer History Museum, 2011).

It appears that unless data archiving is mandatory, researchers are slow to archiving or making their data available (Koopman, 2015). Research funders are influential agents of long-term research data preservation (Doorn and Tjalsma, 2007). Most research funders now demand that research data be archived in an approve repository with protocols which enable its access and use (Wellcome Trust, 2010). Patterton *et al.* (2018) report that South Africa's National Research Fund (NRF) requires applicants for research funding to show proof of how data generated from NRF-funded research would be made publicly available. NRF (2015) requires research publications from funded studies to be deposited on accredited open access repository. NRF also requires the publications to have a Digital Object Identifier (DOI) to facilitate future citation and referencing. Kahn *et al.* (2014) also report similar trends with research funders in many countries such as the UK, the USA and Australia where the funders are increasingly eager that those in receipt of public funding undertake good RDM as a means of improving data quality and facilitating re-use. Anane-Sarpong *et al.* (2018) assert that research funders have a significant say on whether or how data is shared. Patterton, Bothma and Van Deventer (2018) state that having to adhere to funder requirements may be seen as a catalyst for organisations to establish RDM services and infrastructure and to appoint designated RDM personnel. International journals such as *Nature* have now made it mandatory for authors to archive data underlying their articles (Nature, 2014). Some peer reviewers also require research data for verification of the findings as a means of preventing research fraud (Koopman, 2015).

2.5 RDM skills

Studies on RDM skills are limited and seem to focus more on librarians than researchers. Shearer and Schmidt (2016) identify the skills to include knowledge of repositories, data manipulation, data discovery mechanisms, funders' policies and requirements, data centres, data publication requirements of journals, sharing and access, data citation and referencing, metadata standard and schemas amongst others. According to Fary and Owen (2013) as well as Creamer *et al.* (2012) RDM skills include storage, data migration, networking, legal, financial, security, metadata creation and assignment, scholarly data communications and preservation. Kennan (2016) focused on data professionals and identified essential RDM skills to include interpersonal skills, data specific knowledge and skills and metadata.

According to Plotkin (2014), essential RDM skills include ability to identify and analyse data. Berson and Dubov (2011) opine that effective RDM also requires interpersonal skills and behavioural characteristics such as communication, negotiation and competency building skills. According to Lewis (2010) training and advocacy-related communication skills are important for RDM. Earley and Henderson (2017) identify that researchers should understand manipulation of data, intellectual property rights, metadata standards and schemas, data formats, domain ontologies, identifiers, data citation, data licensing discovery tools, database design types and structures, data linking and data integration techniques, data repository and storage platforms. Working effectively with research data also requires an ability to select and appraise datasets, undertake digital preservation activities and apply forensic procedures in digital curation.

Marsh (2012) as well as Coburn and Turner (2012) explain that most universities boast of huge investments in a variety of interventions to help advance lecturers' capacity to engage in data management. However, most research on these interventions is theoretical and incomplete providing little information on what constitutes effective capacity building and the circumstances under which it occurs. In fact, the study by Patterton *et al.* (2018) revealed

that the majority (88%) of researchers had never received any RDM training at the time of data collection.

2.6 Gaps in literature

A number of gaps have been identified in the literature reviewed. One, as already pointed out, most of the studies have focused on data sharing. Assuming that research data cycle begins with data production, then this focus does not paint the complete picture of RDM. Without adequate RDM skills amongst data producers, there would be limited data to share. Two, most of the studies on RDM have been biased towards librarians and other information management professionals. Again, this leaves out the primary players in the research ecosystem. Lecturers play an important part in the producing, sharing and consuming research data. Three, RDM literature on Kenya is limited. Furthermore, the few studies also exhibit the limitations outlined above. Four, most of the studies have been conducted in public universities or institutions. The emphasis has been generating return on investment of public funds. There has been limited concern with private universities. The need to bridge the gaps in literature motivated this study which addresses RDM skills amongst lecturers in a private university in Kenya.

3. Statement of research problem

RDM is a critical part of the research process which is concerned with the organisation, entry, processing, dissemination and archiving of valuable research data. Research output from private universities in Kenya is low. This is evidenced partly by low Webometrics ranking of these universities over the years. Strathmore University is the highest ranked private university in Kenya sitting in position six in the January 2020 Webometrics [1] ranking nationally. The university has maintained a top-10 position nationally in the past five years. Given that lecturers are the major producers of research data in universities, the thesis of this paper is that Strathmore University is performing well in Webometrics ranking because the lecturers in the university have good research data management skills. However, the RDM literacy level of these lecturers has not been assessed or documented in any scientific literature. This study seeks to bridge this gap by identifying and analysing the skills required by lecturers to be able to support RDM effectively; assessing the RDM literacy levels amongst lecturers at Strathmore University, Kenya; and suggesting how RDM capacity of Strathmore University lecturers can be strengthened. The research questions were: What are the data types produced by lecturers at Strathmore University, Kenya? What volume of research data does each lecturer produce? How is the data produced stored? What are the research data sharing attitudes? What are the RDM literacy levels amongst the lecturers? How do lecturers find research data? How is the research data organised and stored? How is the research data secured and shared?

4. Methodology

The research design was a case study of Strathmore University which is situated in Nairobi, Kenya. The university was originally established as Strathmore College in 1961 to offer advanced level of high school education to men. The institution was later given a letter of interim authority by the Government of Kenya to operate as a university in 2002 and offer undergraduate programmes in Commerce and Information Technology. This was followed in 2008 by a full charter giving it legal authority to operate as a university (Strathmore University, 2020). The university currently offers undergraduate and postgraduate degree programmes in business, Information Technology, management, law, tourism and hospitality, applied sciences, as well as humanities and social sciences.

The case study design provided for detailed investigation into the research problem and provided sufficient data to address the objectives of the study. As explained earlier, Strathmore University was preferred because of its top Webometrics ranking amongst private universities in Kenya. The university also has good physical and policy research infrastructure. For instance, it has a policy on research whereby research symposiums are organised on a monthly basis. There is also research incubation where lecturers work together with librarians in ensuring the research output meets the required standards for publication and dissemination. Furthermore, there is lecturer–librarian collaboration in information and data literacy instruction thereby increasing information literacy levels and enhancing competency in information search and research.

This study used explanatory sequential mixed research approach to address the research objectives. This approach involved first collecting quantitative data using questionnaires and thereafter collecting qualitative data using focus group discussions. Mixed research approach enables researchers to make sense of or interpret both qualitative and quantitative phenomena in terms of the meanings people bring to theme as-is. According to Kothari (2014), mixed research is applicable to phenomena relating to or involving quantity and quality.

Information obtained from the Human Resource department indicated that there were 150 lecturers at Strathmore University at the time of data collection. This formed the target population of the study. The sample size was determined from a table published by Israel (1992). From the table, the sample size for a population of 150 is 61 at a confidence level of 95%. The actual respondents were selected through random sampling. Accordingly, 61 respondents were selected randomly from the lecturers who came to use the library during data collection. Each of the selected respondents was given a self-administered questionnaire to fill. The respondents were requested to leave the filled questionnaires in the library. This was the first phase of the explanatory sequential mixed methods approach involving the collection of quantitative data. It was then followed by the collection of qualitative data through focus group discussions. After analysing the quantitative data, it was found that 14 respondents used the library weekly. These lecturers were selected purposively to participate in the focus group discussions. According to Creswell and Creswell (2018), purposive sampling is used when the researchers believe that they can obtain a representative sample through sound judgment. Two focus group discussions consisting of seven participants were held. The number of participants was meant to encourage full participation in the discussions. The purpose of the focus group discussions was to validate and enrich the data collected through questionnaires. Quantitative data was analysed using SPSS while qualitative data was analysed thematically.

5. Findings and discussions

Of the 61 questionnaires issued, 50 (82%) were filled and returned. This response was considered to be very good. According to Meyer and Stockmann (2013), over 60% response rate is considered as good for data analysis and reporting. From the response, 30 (60%) indicated they were male, while 20 (40%) were female. This gender distribution concurs with other studies which generally indicate that there are more male than female lecturers in Kenyan universities with a ratio of about 4:1 (Munene, 2002; Nyaigotti-Chacha, 2004; Sifuna and Chege, 2006; Raburu, 2010). This skewed gender representation in academia seems to be relatively similar even in developed countries. For instance (Booth *et al.*, 2000) conducted a study which revealed that women formed only 28% of permanent academic positions in the United Kingdom. A similar trend has been confirmed by other studies such as (Mitroussi and Mitroussi, 2009), (Tenreyro, 2017) as well as (Santos and Dang, 2019).

The respondents were asked to indicate their age brackets. From the responses, 2 (4%) indicated their ages to be between 20 and 30 years; 15 (30%) 31–40 years; 25 (50%) between 41

and 50 years; and 8 (16%) above 50 years. From this, it can be deduced that the majority (50%) of the lecturers at Strathmore University are within the 41–50 years of age bracket while 30% were between 31 and 40 years of age. This indicates that the bulk (80%) of the lecturers were between 31 and 50 years of age. A study conducted by Munene (2002) at the University of Nairobi revealed that 42.9% of the lecturers at the institution were 31–40 years of age while those aged 41–50 years were 37.1%. Thus, the study by Munene (2002) also found that 80% of the lecturers at the University of Nairobi were aged between 31 and 50 years. Obwogi (2013) analysed the age distribution of academic staff in eight public and private universities in Kenya. The universities were University of Nairobi, Kenyatta University, Jomo Kenyatta University of Agriculture and Technology, Egerton University, United States International University, Kabarak University, KCA University and Maasai Mara University. He found that the majority (29.6%) of the lecturers were aged between 31 and 36 years. This was followed by those aged between 37 and 42 years at 22.4%. Those aged between 49 and 54 years were 16.3% while those aged between 43 and 48 years were 12.2%. The findings revealed that 80.5% of the lecturers in the studied universities were aged between 31 and 54 years. Similar general statistics are witnessed in developed countries. However, there is a relatively larger population of older lecturers in the United States, for instance, due to the removal of mandatory retirement age limits (Ghaffarzadegan and Xu, 2018).

5.1 Current highest degree

The researchers sought to find out the current highest academic levels of the lecturers (participants). From the response, 35 (70%) indicated that they are Master's degree holders while 15 (30%) indicated that they are PhD holders. This indicates that the majority of lecturers are Master's degree holders. These findings generally concur with Munene (2002) who found that 57.3% of lecturers at the University of Nairobi had Master's degree as the highest educational qualification. The same general trend was observed for the number of lecturers with doctoral degrees where Strathmore posted 30% while the University of Nairobi had 40%. From the focus group discussions, it emerged that the variance in the highest educational levels of staff can be attributed to the fact that Strathmore is a relatively young university chartered only in 2008 while the University of Nairobi was established in the 1960s. The low number of staff with doctoral degrees at both universities indicates inadequate postgraduate training in Kenya. According to Mukhwana *et al.* (2016), postgraduate students formed only 11% of the student population in Kenyan universities with PhD level accounting for only 1.3%. Therefore, the low number of PhD candidates being trained in Kenya contributes to the low number of academic staff with PhD level of education in Kenyan universities since the few PhDs produced are shared by the universities with the government, private and development sectors of the economy.

5.2 Disciplines

The study sought to find out the areas of professional qualification of the lecturers by discipline. The respondents indicated their areas of discipline as 15 (30%) business management; 2 (4%) natural sciences; 10 (20%) humanities and social sciences; 10 (20%) information technology; 5 (10%) law; 5 (10%) applied sciences; and 3 (6%) engineering. The majority (50%) of the lecturers at Strathmore University are in the business management and information technology disciplines. The participants in the focus group discussions confirmed that these are the two flagship disciplines of the university since it launched undergraduate programmes in 2002. Therefore, the bulk of academic programmes in the university are in these and related disciplines. It is worth noting that these disciplines cover humanities and social sciences; natural and applied sciences; as well as engineering and technology. These are the core disciplines of academic pursuit in Kenya. Therefore, the RDM

skills required by academic staff at Strathmore University are the same skills which academics in other universities in Kenya need.

5.3 Type of research data

The respondents were asked to indicate the type of research data they produce in their day-to-day activities. The response was as follows: structured scientific and statistical data (SPSS, GIS) at 15 (30%); source code (Java, Scripting, C) data at 3 (6%); structured graphic data (CAD, CAM) at 1(2%); configuration data (parameter settings, logs) at 1(2%); Internet and web-based data (Web pages, emails) at 20 (40%); non-digital data (paper, films, slides, artefacts) at 8(16%); and images (JPEG, GIF, TIFF) data at 2 (4%). The majority (40%) of the respondents indicated that they produce Internet and web-based data followed by structured and scientific statistical data (30%). The participants in the focus group discussions explained that this is a pointer to the fact that whereas different types of data are produced, much of it is scientific and web-based data which is used more in academia. This finding is in tandem with emerging trends in scholarly communication in which Internet data is becoming more important. This is because the Internet, as a technology and communication platform, has facilitated better creation, processing, packaging, dissemination and storage of scholarly information in the recent past. Thus, it is playing a more significant role in research and scholarly communication than before. Given challenges associated with credibility, security and longevity of digital data, academics in Kenya, and elsewhere, need specialised skills to effectively manage research data which is generated, collected, processed or disseminated through the Internet (Gupta, 2017; Loescher *et al.*, 2011; Lyons *et al.*, 2005; Nel *et al.*, 2017).

5.4 Volume of research data generated

The respondents were asked to indicate the volume of data they generate in their work. Twenty (40%) indicated that they generate data in Megabytes (MB), 25 (50%) in Gigabytes (GB), 5 (10%) in Terabytes (TB). Thus, the majority (50%) of the respondents indicated that most of the data they produce are in volumes of Gigabytes (GB). Through focus group analysis, the majority of respondents still indicated that they produce data in Gigabytes. At the same time, a number of respondents appeared not to know the other data volumes such as Exabytes, Petabytes, Zettabytes and Yottabytes. The findings demonstrate the fact that research data is increasing in volume and variety. This trend can be attributed to the growing ubiquity of the use of digital techniques and tools in research. Consequently, it has become easier and cost-effective to create and process research data digitally. As digital tools become more ubiquitous in academic settings, the volume of research data is bound to increase (Benfield and Szlemko, 2006; Jahnke and Asher, 2012a, b; Peersman, 2014; Connelly *et al.*, 2016; Sivarajah *et al.*, 2017). With the growing volume of research data created, shared or stored, academics are likely to suffer from information overload. Therefore, they require adequate skills to select and curate useful data for current and anticipated research needs.

5.5 Storage of research data

The researchers sought to find out from the respondents how they store the data they generate. A total of 32 (64%) indicated that they store their data in their own devices; 10 (20%) in cloud services; 2 (4%) in central servers; and 6 (12%) in institutional repositories. From this, it can be deduced that the majority (64%) of the respondents prefer storing data in their own devices such as personal computers, flash and external hard disks. The participants of the focus group discussions acknowledged that this preferred storage mechanism limits the ability to share or access research data. The data is also vulnerable to risks such as loss when the device is lost or corruption due to unauthorised access of the device. These findings

concur with Kennan and Markauskaite (2015) who conducted a study in ten universities in New South Wales in Australia and concluded that researchers did not have comprehensive research data storage plans and therefore used their own devices which were neither safe nor secure. They further stated that at the time the study was conducted, there was no requirement by the government and other stakeholders regarding research data storage. The situation is steadily changing and researchers are now expected to elucidate their data storage strategies during ethical clearance. Indeed Piorun *et al.* (2012, p. 47) argue that academics should be able to "handle issues involved in securely storing research data in central databases, archives and/or repositories, backing it up, and managing access to your data". Therefore, lecturers need skills in not only developing but also executing comprehensive research data storage plans. The fact that the majority of lecturers in this study were using personal devices to store research data indicates a gap in their RDM skills which needs to be addressed. The gap can be addressed through training and development of institutional policies on research data storage.

5.6 Research sharing attitude

The respondents were asked to indicate their research data sharing attitude. A total of 18 (36%) indicated that their research data is available openly on request; 9 (18%) stated that their data is available only to their research teams; 12 (24%) indicated that their data has restricted access; 10 (20%) explained that their data is openly available to everyone; 1 (2%) indicated that the data is not available to anyone else. From this, it can be deduced that the majority of the respondents avail their research data only on request. The participants in the focus group discussions added that lecturers are generally reluctant to share their research data fearing the same could be misused. Sharing of research data is critical in the generation and discovery of new knowledge. Borgman (2012) argues that researchers should share their data so as to enable other researchers to reproduce or to verify research; make results of publicly funded research available to the public; enable others to ask new questions of extant data; and advance the state of research and innovation. Indeed, researchers, such as the lecturers who were the subjects of this current study, use other researchers' work to develop their own. Fecher *et al.* (2015) assert that in spite of widespread support from policy makers, funding agencies and scientific journals, academic researchers rarely make their research data available to others. The reasons why researchers do not share research data vary but they largely revolve around lack of expertise, resources or incentives. Borgman (2012) further opines that research data may not at all times exist in transferable forms. Furthermore, some data may not be sharable for ethical or epistemological reasons. Regardless of the reasons for it, reluctance to share research data is a global concern. Therefore, the findings of the current study reflect this trend. This underscores the need to address it. This current study argues that improving RDM skills will result in increased sharing and disclosure of research data generated in universities. This opinion concurs with Sayogo and Pardo (2013) that the level of RDM skills amongst researchers determines the degree to which they share their research data.

5.7 RDM literacy levels

The study endeavoured to find out the literacy level of respondents on RDM skills in various areas. The findings, also shown in Table 1, were validated by focus group discussion participants. They are presented and discussed hereunder.

5.7.1 Planning for research data. Respondents were asked to rate their literacy levels in planning for research data. They indicated their skills as poor 2(4%), good 38(76%), very good 10(20%). The majority of the respondents indicated that they have good literacy levels in planning for research data. This implies that the lecturers are able to develop and execute

S/N	Literacy area	Literacy level: Very poor	Poor	Good	Very good
Planning for research data					
1	Determining research data needs	0 (0%)	2 (4%)	38 (76%)	10 (20%)
2	Identifying types of research data needed	0 (0%)	3 (6%)	32 (64%)	15 (30%)
Finding research data					
3	Use of institutional repositories	2 (4%)	3 (6%)	30 (60%)	15 (30%)
4	Use of research databases	0 (0%)	3 (6%)	38 (76%)	9 (18%)
5	Use of search engines	0 (0%)	5 (10%)	33 (66%)	12 (24%)
6	Use of Boolean operators	8 (16%)	10 (20%)	25 (50%)	7 (14%)
7	Assessing credibility of web content	0 (0%)	5 (10%)	32 (64%)	13 (26%)
8	Citing and referencing data	0 (0%)	1 (2%)	36 (72%)	13 (26%)
9	Use of citation and referencing software	0 (0%)	3 (6%)	27 (54%)	20 (40%)
Organising research data					
10	Formatting research data	1 (2%)	4 (8%)	25 (50%)	20 (40%)
11	Classifying research data	1 (2%)	3 (6%)	34 (68%)	12 (24%)
12	File naming and versioning	1 (2%)	3 (6%)	25 (50%)	21 (42%)
13	Use of metadata	3 (6%)	11 (22%)	18 (36%)	18 (36%)
14	Documenting research data	1 (2%)	10 (20%)	29 (58%)	10 (20%)
15	Use of research data analysis software	8 (16%)	15 (30%)	18 (36%)	9 (18%)
Storing research data					
16	Use of institutional repository	3 (6%)	8 (16%)	30 (60%)	9 (18%)
17	Use of research databases	1 (2%)	10 (20%)	29 (58%)	10 (20%)
18	Backing up research data	1 (2%)	12 (24%)	22 (44%)	15 (30%)
19	Storing data in the clouds	2 (4%)	15 (30%)	20 (40%)	13 (26%)
20	Use of self-archiving systems	6 (12%)	15 (30%)	19 (38%)	10 (20%)
Security of research data					
21	Data preservation	10 (20%)	17 (34%)	20 (40%)	3 (6%)
22	Data security	2 (4%)	10 (20%)	26 (52%)	12 (24%)
23	Access authentication	8 (16%)	17 (34%)	20 (40%)	5 (10%)
24	Conditions of use	4 (8%)	17 (34%)	23 (46%)	6 (12%)
25	Data legislation	8 (16%)	25 (50%)	12 (24%)	5 (10%)
Sharing research data					
26	Use of social media (e.g. ResearchGate)	1 (2%)	14 (28%)	30 (60%)	5 (10%)
27	Open access journals	1 (2%)	23 (26%)	18 (36%)	8 (16%)
28	Identifying high impact channels	1 (2%)	20 (40%)	22 (44%)	7 (14%)
29	Identifying predatory journals	2 (4%)	36 (72%)	8 (16%)	4 (8%)
30	Use of online manuscript submission	6 (12%)	15 (30%)	20 (40%)	9 (18%)
31	Online peer reviewing	3 (6%)	11 (22%)	32 (64%)	4 (8%)
32	Creating research networks	2 (4%)	18 (36%)	27 (54%)	3 (6%)
33	Use of plagiarism checkers	0 (0%)	6 (12%)	30 (60%)	14 (28%)

Table 1. Lecturers' RDM literacy levels

plans on how manage their research data. This finding may be attributed to the fact that most lecturers have been trained in how to design and conduct academic research studies. Similarly, the good result may be attributed to the fact that most academics are required to obtain permits and ethical clearance to conduct research. Comprehensive plans of how to collect and generally handle research data are required before these permits or clearances can be awarded.

5.7.2 Identifying types of research data needed. The respondents were also asked to rate their literacy level in identifying research data needed. Three (6%) indicated poor skills, 32

(64%) indicated they have good skills while 15 (30%) had very good skills. The majority (64%) of the respondents indicated that they have good literacy levels in identifying research data. Again, this finding may be attributed to good research proficiency amongst lecturers which is developed through postgraduate courses in research methodology. It is therefore expected that lecturers having taken these courses should be proficient in research methods which include skills to identify research data required to answer research questions in specific research projects.

5.7.3 Finding research data. The respondents were asked to indicate their literacy levels in the use of institutional repositories. Two (4%) indicated very poor skills on the use of institutional repositories; 3 (6%) poor skills; 30 (60%) good skills; while 15 (30%) very good literacy levels in the use of institutional repositories. Thus, most of the respondents indicated that they have good literacy skills in the use of institutional repositories with very few saying they have poor skills on the same. On the use of research databases, the response was 3 (6%) poor skills, 38 (76%) good skills and 9 (18%) very good skills. Therefore, the majority (76%) of the respondents indicated that they have good skills in the use of research databases. On the use of search engines, the response was poor 5 (10%), good 33 (66%) and very good 12 (24%). The respondents were also asked to indicate their skills on the use of Boolean operators. The response was very poor 8 (16%), poor 10 (20%), good 25 (50%) and very good 7(14%). It can be deduced that the majority of the respondents have good skills on the use of search engines and Boolean operators. On the ability to assess the credibility of web content, the response was poor 5 (10%), good 32 (64%) and very good 13 (26%). From this, it can be deduced that the majority of the respondents have good skills on assessing the credibility of web content. The response for ability to cite and reference data was poor 1(2%), good 36 (72%) and very good 13 (26%). Therefore, the majority of the respondents on average have good skills in citing and referencing data. On the use of citation and referencing software, the response was poor 3 (6%), good 27 (54%) and very good 20 (40%). Most of the respondents indicated that they have good skills in the use of citations and referencing software. Some respondents mentioned the software as Zotero and Mendeley.

The findings reveal that the lecturers studied have good skills in searching for, finding and accessing research data using digital platforms, techniques and tools. This can be attributed to the fact that most lecturers at Strathmore University are generally young and tech-savvy. Similarly, there has been concerted effort by members of the Kenya Library Information Services Consortium (KLISC) in promoting the use of e-resources, particularly through institutional repositories. Indeed, studies by Chilimo (2015), Moseti (2016), Ratanya (2017) and Kakai *et al.* (2018) demonstrate that there is growing ICT literacy amongst academic library users in Kenya and East Africa in general. This explains why many academic library users are proficient in using electronic sources of data.

5.7.4 Organising research data. The respondents were asked to indicate their literacy level in organising data. On formatting research data, the response was very poor 1 (2%), poor 4 (8%), good 25 (50%) and very good 20 (40%). One (2%) respondent did not respond to this question. Thus, most respondents indicated that they have good skills in formatting research data while a few have poor skills on the same. In classifying research data, the responses were very poor 1 (2%), poor 3 (6%), good 34 (68%) and very good 12 (24%). Thus, most respondents indicated that they have good skills in classifying research data. On file naming and versioning, the response was very poor 1 (2%), poor 3 (6%), good 25 (50%) and very good 21 (42%). It is evident that a few of the respondents have very poor skills on file naming and versioning while the majority have good skills in the same. On the use of metadata, 3(6%) indicated they have very poor skills, poor 11 (22%), good 18 (36%) and very good 18 (36%). On average, therefore, most respondents have good literacy skills in the use of metadata. On documenting research data, the response was very poor 1 (2%), poor 10 (20%), good 29 (58%) and very good 10 (20%). Most respondents indicated that they have good skills on

documenting research data. It can be deduced that the majority understand the need of documenting research data. On the use of research data analysis software, the responses were: very poor 8 (16%), poor 15 (30%), good 18 (36%) and very good 9 (18%). Most respondents indicated that they have good skills on the use of research data analysis software. It was noteworthy, however, that a good number at (30% and 8% respectively) have poor and very poor skills on the same.

5.7.5 Storing research data. The respondents were asked to indicate their literacy skills in storing research data. On storing data in institutional repositories, the response was: very poor 3 (6%), poor 8 (16%), good 30 (60%) and very good 9 (18%). Most respondents indicated that they have good skills in storing their research data in institutional repositories. This corroborated the findings of studies by Chilimo (2015) and Moseti (2016). On the use of research databases, the response was: very poor 1 (2%), poor 10 (20%), good 29 (58%) and very good 10 (20%). The study established that the majority of the respondents have good skills in the use of research databases while 1 (2%) and 10 (20%) have very poor and poor literacy skills respectively on the use of research databases. This is a great improvement on the scenario in 2009 when Shabaya (2009) reported a low usage of research databases by lecturers in Kenya. This improvement is likely to have been a result of increased accessibility of research databases by lecturers through collectively procured e-resources (Brooks *et al.*, 2005). On backing up research data the response was: very poor 1(2%), poor 12 (24%), good 22 (44%) and very good 15 (30%). Most respondents indicated that they have good skills on backing up of research data. On storing data in the clouds, the response was: very poor 2 (4%), poor 15 (30%), good 20 (50%) and very good 13 (26%). In the use of archiving systems, the response was: very poor 6 (12%), poor 15 (30%), good 19 (38%), very good 10 (20%). The majority of the respondents indicated that they have good skills in storing data in the clouds as well as in the use of archiving systems. According to Omwansa *et al.* (2014), storage is the most popular use of cloud computing platforms in Kenya. This view is also supported by Wasike and Njoroge (2015), Muli and Kimuai (2015) and Micheni (2015).

5.7.6 Security of research data. The respondents were asked to indicate their literacy level in securing research data based on the following:

(1) Data preservation – On this, 10 (20%) indicated that they have very poor skills, poor 17 (34%), good 20 (40%) and very good 3 (6%). Whereas the majority (40%) indicated they have good skills in data preservation, a good number (34%) have poor skills on the same. These findings corroborate those of Onyancha (2016) as well as Ng'eno and Mutula (2018).

(2) Data security – The response was: very poor 2 (4%), poor 10 (20%), good 26 (52%) and very good 12 (24%). Therefore, the majority indicated that they had a good understanding of research data security. These skills include a good understanding of the rationale, techniques and tools for enhancing research data security (Ng'eno, 2018).

(3) Access authentication – The response was: very poor 8 (16%), poor 17 (34%), good 20 (40%) and very good 5 (10%). From the response, the majority have good skills on data security as well as access authentication response. These findings are anchored on rationale similar to Mulder *et al.* (2017) who assert that universities have put in place measures to facilitate data safety in their networks through effective setting and use of strong passwords.

(4) Setting the conditions of use of data – The response was: very poor 4 (8%), poor 17 (34%), good 23(46%) and very good 6 (12%). Most respondents indicated that they have good skills on setting conditions for use of data. With growing drive towards

open access publishing, more academics are now familiar with creative commons and other regimes for accessing or sharing research data (Lipton, 2020).

(5) Data legislation – The response was: very poor 8 (16%), poor 25 (50%), good 12 (24%) and very good 5 (10%). The majority of the respondents indicated that they have poor literacy skills on data legislation. This study concurs with Ng'eno (2018) that most research institutions lack adequate policies and legislation governing RDM. Consequently, lecturers are largely unaware of legislative provisions regarding RDM. Ng'eno (2018) further opines that where they existed, policies and regulations were outdated and that research institutions in Kenya lacked dedicated units to coordinate RDM functions.

5.7.7 Sharing research data. The respondents were asked to indicate their literacy levels in sharing of data through a number of platforms. On the use of social media (such as ResearchGate), the response was: very poor 1 (2%), poor 14 (28%), good 30 (60%), very good 5 (10%). The majority of the respondents indicated that they have good skills on the use of social media for sharing research data. This corroborates Kibugi (2013) as well as Abok and Kwanya (2016). On the use of open access platforms, the response was: very poor 1 (2%), poor 23 (46%), good 18 (36%), and very good 8 (16%). Most respondents indicated that they have poor literacy skills in open access. Given the growing popularity of open access as a means of increasing access to research data, there is need to determine the reasons for the low level of proficiency in using open access platforms to share research data. On identifying high impact publication channels, the response was: very poor 1 (2%), poor 20 (40%), good 22 (44%) and very good 7 (14%). Most respondents have good skills in identifying high impact channels of communication. On identifying predatory journals, the response was: very poor 2 (4%), poor 36 (72%), good 8 (16%) and very good 4 (8%). The majority of the respondents indicated that they have poor literacy skills on identifying predatory journals. This finding can be attributed to the fact that most predatory journals have lately become so mixed with credible journals that it is difficult for ordinary academics to distinguish them (Xia *et al.*, 2015; Mouton and Valentine, 2017; Kurt, 2018). On the use of manuscript submission interfaces, the response was: very poor 6 (12%), poor 15 (30%), good 20 (40%) and very good 9 (18%). Most respondents have good skills in using manuscript submission interfaces. On online peer reviewing, the response was: very poor 3 (6%), poor 11 (22%), good 32 (64%) and very good 4 (8%). On creating research networks, the response was: very poor 2 (4%), poor 18 (36%), good 27 (54%) and very good 3 (6%). On the use of plagiarism checkers, the response was: poor 6 (12%), good 30 (60%) and very good 14 (28%). Most respondents have good literacy skills in online peer reviewing, creating research networks and use plagiarism checkers.

5.8 Recommendations on how to strengthen RDM literacy levels

The participants in the focus group discussions were asked to suggest strategies which can be used to strengthen RDM literacy levels amongst lecturers at Strathmore University. Their suggestions are summarised as follows:

(1) The library and directorate of research should organise RDM training for all lecturers. All new lecturers joining the university should be trained on RDM as part of their induction. Regular RDM training should be conducted to refresh or update the skills of the lecturers as new technologies as well as research techniques, tools and protocols emerge.

(2) The university management should develop and implement a system which encourages and rewards lecturers who excel in RDM.

(3) Lecturers who excel in RDM should be empowered to act as RDM champions amongst their colleagues.

(4) RDM should be integrated in the academic curricula in the university particularly at the postgraduate levels.

(5) The library as well as the directorates of research and ICT should collaboratively develop infrastructure which supports effective RDM.

6. Conclusion

The findings of this study reveal varied levels of RDM literacy amongst lecturers at Strathmore University, Kenya. Overall, the lecturers have good skills in the majority of the RDM capacity areas. Thus, the lecturers are able to plan for, search, find, organise, store, secure and share research data competently. Therefore, they participate actively in data creation, collection, processing, validation, dissemination, sharing and archiving. This good level of RDM expertise partly explains the relatively good performance of the university in Webometrics ranking of Kenyan universities. However, the study also revealed gaps in the RDM literacy levels amongst the lecturers in areas such as sharing of research data on open access journals, data legislation and securing research data. These knowledge gaps should be mitigated through training.

7. Recommendations

Based on the findings, the study recommends that more emphasis should be put on training lecturers to understand the different types of research data that they generate in the course of their work. This will help them to know how to manage the data effectively. RDM training should also address issues relating to research data sharing attitudes as a means of enhancing access to research data generated or collected by lecturers in Strathmore University. Similarly, particular attention should be put on sensitising the lecturers on research sharing platforms such as institutional repositories, subscription to authoritative research databases, as well as citing and referencing research data. More attention should also be put on training lecturers on organisation, storing, securing as well as sharing of research data with a special focus on providing skills to determine authentic platforms to share data, for instance, identifying predatory journals and high impact channels.

8. Practical implications and originality/value

The findings of this study may be used to sensitise lecturers, librarians and university management about RDM. The findings may also provide evidence for decision making and policy development on matters pertaining to RDM at Strathmore University and other universities. Furthermore, the findings may be used by researchers to provide context and background information for future research on RDM in private universities locally, regionally and globally.

Note

1. http://www.webometrics.info/en/Africa/Kenya%20

References

Abok, V.A. and Kwanya, T. (2016), "Maximising the potential of social media to deliver academic library services to students: a case study of the Technical University of Kenya Library", *Inkanyiso Journal of Humanity and Social Sciences*, Vol. 8, pp. 147-155.

AGIMO (2013), *Big Data Strategy – Issues Paper 12*, Commonwealth of Australia, available at: https://www.dijitalakademi.gov.tr/wp-content/uploads/2013/08/Big-Data-Strategy1.pdf (accessed 24 May 2020).

Anane-Sarpong, E., Wangmo, T., Ward, C.L., Sankoh, O., Tanner, M. and Elger, B.S. (2018), "You cannot collect data using your own resources and put it on open access": perspectives from Africa about public health data sharing", *Developing World Bioethics*, Vol. 18, pp. 394-405.

Benfield, J.A. and Szlemko, W.J. (2006), "Internet-based data collection: promises and realities", *Journal of Research Practice*, Vol. 2, pp. 1-15.

Berson, A. and Dubov, L. (2011), *Master Data Management and Data Governance*, 2nd ed., McGraw-Hill, New York, NY.

Booth, A.L., Burton, J. and Mumford, K. (2000), "The position of women in UK academic economics", *Economic Journal*, Vol. 110, pp. 312-333.

Borgman, C.L. (2012), "The conundrum of sharing research data", *Journal of the American Society for Information Science and Technology*, Vol. 63, pp. 1059-1078.

Briney, K. (2015), *Data Management for Researchers: Organize, Maintain and Share Your Data for Research Success, Research Skills Series*, Pelagic Publishing, Exeter.

Brooks, S., Donovan, P. and Rumble, C. (2005), "Developing nations, the digital divide and research databases", *Serials Review*, Vol. 31, pp. 270-278.

Chawinga, W.D. and Zinn, S. (2019), "Global perspectives of research data sharing: a systematic literature review", *Library and Information Science Research*, Vol. 41, pp. 109-122.

Chilimo, W. (2015), "Green open access in Kenya: a review of the content, policies and usage of institutional repositories", *Mousaion*, Vol. 33, pp. 25-54.

Coburn, C.E. and Turner, E.O. (2012), "The practice of data use: an introduction", *American Journal of Education*, Vol. 118, pp. 99-111.

Computer History Museum (2011), *Digital dark age*, Computer History Museum, available at: https://www.computerhistory.org/revolution/memory-storage/8/325/2208 (accessed 24 May 2020).

Connelly, R., Playford, C.J., Gayle, V. and Dibben, C. (2016), "The role of administrative data in the big data revolution in social science research", *Social Science Research*, Vol. 59, pp. 1-12.

Corrall, S., Kennan, M.A. and Afzal, W. (2013), "Bibliometrics and research data management services: emerging trends in library support for research. Libr", *Trends*, Vol. 61, pp. 636-674, doi: 10.1353/lib.2013.0005.

Cox, A.M. and Verbaan, E. (2018), *Exploring Research Data Management*, Facet Publishing, London.

Creamer, A., Morales, M.E., Crespo, J., Kafel, D. and Martin, E.R. (2012), "An assessment of needed competencies to promote the data curation and management librarianship of health sciences and science and technology librarians in New England", *Journal of EScience Librarianship*, Vol. 1, p. 4.

Creswell, J.W. and Creswell, J.D. (2018), *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 5th ed., SAGE, Los Angeles.

Davenport, T.H. and Patil, D.J. (2012), "Data scientist: the sexiest job of the 21st century", *Harvard Business Review*, October 2012, pp. 70-76.

Denny, S.G., Silaigwana, B., Wassenaar, D., Bull, S. and Parker, M. (2015), "Developing ethical practices for public health research data sharing in South Africa: the views and experiences from a diverse sample of research stakeholders", *Journal of Empirical Research on Human Research, Ethics*, Vol. 10, pp. 290-301.

Doorn, P. and Tjalsma, H. (2007), "Introduction: archiving research data", *Archival Science*, Vol. 7, pp. 1-20.

Earley, S. and Henderson, D. (Eds) (2017), *DAMA-DMBOK: Data Management Body of Knowledge*, 2nd ed., Technics Publications, Basking Ridge, New Jersey, NJ.

Fary, M. and Owen, K. (2013), "Developing an institutional research data management plan service", *Advanced Core Technology Initiative*, pp. 1-24.

Fecher, B., Friesike, S. and Hebing, M. (2015), "What drives academic data sharing?", *PloS One*, Vol. 10, available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4340811/pdf/pone.0118053.pdf (accessed 20 May 2020).

Ghaffarzadegan, N. and Xu, R. (2018), "Late retirement, early careers, and the aging of US science and engineering professors", *PloS One*, Vol. 13, available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6306255/pdf/pone.0208411.pdf (accessed 24 May 2020).

Gupta, S. (2017), "Ethical issues in designing Internet-based research: recommendations for good practice", *Journal of Research Practice*, Vol. 13, p. 1.

Israel, G.D. (1992), "Determining sample size", available at: http://www.academia.edu/download/58812033/samplesize1.pdf (accessed 24 May 2020).

Herzog, D. (2016), *Data Literacy: A User's Guide*, SAGE Publications, Thousand Oaks, California.

Jahnke, L.M. and Asher, A. (2012a), "The problem of data: data management and curation practices among university researchers", *Council Library Information Resources*, pp. 3-31.

Jahnke, L.M. and Asher, A. (2012b), "The problem of data: data management and curation practices among university researchers", *The Problem of Data*, pp. 3-31, available at: https://www.clir.org/pubs/reports/pub154/problem-of-data/ (accessed 24 May 2020).

Kahn, M., Higgs, R., Davidson, J. and Jones, S. (2014), "Research data management in South Africa: how we shape up", *Australian Academic and Research Libraries*, Vol. 45, pp. 296-308.

Kakai, M., Musoke, M. and Okello, O.C. (2018), "Open access institutional repositories in universities in East Africa", *Information and Learning Science*, Vol. 119, pp. 667-681.

Kennan, M.A. (2016), "Data Management: knowledge and skills required in research, scientific and technical organisations", *International Federation of Library and Association Institutions IFLA*, available at: http://library.ifla.org/1466/1/221-kennan-en.pdf (accessed 24 May 2020).

Kennan, M.A. and Markauskaite, L. (2015), "Research data management practices: a snapshot in time", *International Journal of Digital Curation*, Vol. 10, pp. 69-95.

Kibe, L., Kwanya, T. and Owano, A. (2020), "Relationship between big data analytics and organisational performance of the Technical University of Kenya and Strathmore University in Kenya", *Global Knowledge Memory and Communication*, doi: 10.1108/GKMC-04-2019-0052.

Kibugi, S. (2013), "The use of social media in the dissemination of information in selected public and private university libraries in Kenya", *Journal of Appropriate Librarianship and Information Work South Africa*, Vol. 47, pp. 100-120.

Koopman, M.M. (2015), *Data Archiving, Management Initiatives and Expertise in the Biological Sciences Department*, University of Cape Town, Cape Town.

Kothari, C.R. (2014), *Research Methodology: Methods and Techniques*, New Age International, New Delhi.

Kurt, S. (2018), "Why do authors publish in predatory journals? Learn", *Learned Publishing*, Vol. 31, pp. 141-147.

Lewis, M. (2010), "Libraries and the management of research data", in McKnight, S. (Ed.), *Envisioning Future* Academic Library Services, Facet, pp. 145-168, doi: 10.29085/9781856048750.011.

Lipton, V. (2020), *The Current Policies of Research Funders and Publishers*, in *Open Scientific Data-Why Choosing and Reusing the RIGHT DATA Matters*, IntechOpen, available at: https://www.intechopen.com/books/open-scientific-data-why-choosing-and-reusing-the-right-data-matters/the-current-policies-of-research-funders-and-publishers.

Loescher, L.J., Hibler, E., Hiscox, H., Hla, H. and Harris, R.B. (2011), "Challenges of using the internet for behavioral research", *Comput. Inform. Nurs. CIN*, Vol. 29, p. 445.

Lyons, A.C., Cude, B., Lawrence, F.C. and Gutter, M. (2005), "Conducting research online: challenges facing researchers in family and consumer sciences", *Family and Consumer Sciences Research Journal*, Vol. 33, pp. 341-356.

Marsh, J.A. (2012), "Interventions promoting educators' use of data: research insights and gaps", *Teachers College Record*, Vol. 114, pp. 1-48.

Meyer, W. and Stockmann, R. (2013), *Functions, Methods and Concepts in Evaluation Research*, Palgrave Macmillan, New York, NY.

Micheni, E.M. (2015), "Using the technology organization environment framework for adoption and implementation of cloud computing in institutions of higher learning in Kenya", *International Journal of Engineering Science*, Vol. 4, pp. 37-43.

Mitroussi, A. and Mitroussi, K. (2009), "Female educational leadership in the UK and Greece", *Gender in Management: An International Journal*, Vol. 24, pp. 505-522.

Moseti, I. (2016), "Digital preservation and institutional repositories: case study of universities in Kenya", *Journal of South African Society of Archivists*, Vol. 49, pp. 137-154.

Mouton, J. and Valentine, A. (2017), "The extent of South African authored articles in predatory journals", *South African Journal of Science*, Vol. 113, pp. 1-9.

Mukhwana, E.J., Oure, S., Too, J. and Some, D.K. (2016), *State of Post Graduate Research and Training in Kenya*, Commission for University Education, Nairobi.

Mulder, N., Adebamowo, C.A. and Adebamowo, S.N. (2017), "Genomic research data generation, analysis and sharing - challenges in the African setting", *Data Science Journal*, Vol. 16, pp. 1-15.

Muli, E. and Kimuai, J. (2015), "Adoption of cloud computing for education in Kenyan universities: challenges and opportunities", *Int J Adv Res Comput Eng Technol IJARCET*, Vol. 4, pp. 2854-2860.

Munene, I.I. (2002), "University academics: demographic, role structure characteristics and attitudes towards merit and equity–a Kenyan case study", *Research in Post-Compulsory Education*, Vol. 7, pp. 247-272, Nature, 2014. Data policies. Nat. Res.

National Research Foundation, (2004), "Statement on open access to research publications from the National Research Foundation (NRF)-Funded Research", available at: https://www.nrf.ac.za/media-room/news/statement-open-access-research-publications-national-research-foundation-nrf-funded (accessed 24 May 2020).

Nature, (2014), "Scientific data: data policies", available at: http://www.nature.com/sdata/data-policies (accessed 24 May 2020).

Nel, E., Rich, E., Morojele, N., Harker, B., Petersen, N., Williams, P. and Parry, C. (2017), "Data collection challenges experienced while conducting the International alcohol control study (IAC) in Tshwane, South Africa. Drugs", *Drugs: Education, Prevention and Policy*, Vol. 24, pp. 376-383.

Ng'eno, E.J. (2018), *Research Data Management in Kenya's Agricultural Research Institutes (Doctoral Dissertation)*, University of KwaZulu-Natal, Pietermaritzburg.

Ng'eno, E. and Mutula, S. (2018), "Research Data Management (RDM) in agricultural research institutes: a literature review", *Inkanyiso Journal of Humanity and Social Sciences*, Vol. 10, pp. 28-50.

Nyaigotti-Chacha, C. (2004), "Reforming higher education in Kenya", Seminar Paper, Naivasha, Kenya, available at: https://pdfs.semanticscholar.org/fc9a/632255e461300b9aea588050e32ca4cdd8e5.pdf (accessed 24 May 2020).

Obwogi, J. (2013), *Factors that Affect Quality of Teaching Staff in Universities in Kenya (Doctoral Thesis)*, JKUAT, Nairobi.

Omwansa, T.K., Waema, T.M. and Omwenga, B. (2014), *Cloud Computing in Kenya (Doctoral Dissertation)*, University of Nairobi, Nairobi.

Onyancha, O.B. (2016), "Open research data in sub-Saharan Africa: a bibliometric study using the data citation index", *Publishing Research Quarterly*, Vol. 32, pp. 227-246.

Patterton, L., Bothma, T.J. and Van Deventer, M.J. (2018), "From planning to practice: an action plan for the implementation of research data management services in resource-constrained institutions", *South African Journal for Librarianship and Information Science*, Vol. 84, pp. 14-26.

Peersman, G. (2014), "Overview: data collection and analysis methods in impact evaluation", *Methodological Briefs: Impact Eval. Florence UNICEF*, Vol. 10.

Pisani, E., Aaby, P., Breugelmans, J.G., Carr, D., Groves, T., Helinski, M. and Mboup, S. (2016), "Beyond open data: realising the health benefits of sharing data", *Bmj*, Vol. 355, p. I5295.

Piorun, M., Kafel, D., Leger-Hornby, T., Najafi, S., Martin, E., Colombo, P. and LaPelle, N. (2012), "Teaching research data management: an undergraduate/graduate curriculum", *Journal of EScience Librarianship*, Vol. 1, pp. 46-50.

Plotkin, D. (2014), *Data Stewardship: An Actionable Guide to Effective Data Management and Data Governance*, Elsevier/Morgan Kaufman, Amsterdam, Boston.

Provost, F. and Fawcett, T. (2013), *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*, 1st ed., 2. release. ed., O'Reilly, Beijing.

Raburu, P. (2010), *Women Academics' Careers in Kenya (Doctoral Dissertation)*, Lancaster University, Lancaster.

Ratanya, F.C. (2017), "Institutional repository: access and use by academic staff at Egerton University, Kenya", *Library Management*, Vol. 38, pp. 276-284.

Ray, J.M. (Ed.) (2014), *Research Data Management: Practical Strategies for Information Professionals, Charleston Insights in Library, Archival, and Information Sciences*, Purdue University Press, West Lafayette, Indiana.

Santos, G. and Dang, V.P.S. (2019), "Gender and academic rank in the UK", *Sustainability*, Vol. 11, p. 3171.

Sayogo, D.S. and Pardo, T.A. (2013), "Exploring the determinants of scientific data sharing: understanding the motivation to publish research data", *Government Information Quarterly*, Vol. 30, pp. 19-31.

Schöpfel, J., Ferrant, C., André, F. and Fabre, R. (2018), "Research data management in the French national research center (CNRS)", *Data Technol. Appl.*

Shabaya, P. (2009), *The Changing Role of Information and Communication Technologies (ICTs) for Instruction in Higher Education Institutions in Kenya*, Strathmore ICT 2009 Conference paper, Nairobi.

Shearer, B. and Schmidt, K. (2016), *Librarians Competencies Profile for Research Data Management*, Foster, available at: https://www.coar-repositories.org/files/Competencies-for-RDM_June-2016.pdf (accessed 24 May 2020).

Sifuna, D.N. and Chege, F.N. (2006), *Girls' and Women's Education in Kenya: Gender Perspectives and Trends*, Nairobi UNESCO, Nairobi.

Sivarajah, U., Kamal, M.M., Irani, Z. and Weerakkody, V. (2017), "Critical analysis of Big Data challenges and analytical methods", *Journal of Business Research*, Vol. 70, pp. 263-286.

Strathmore University (2020), "Our key milestones", *Strathmore University*, available at: https://www.strathmore.edu/about-strathmore/ (accessed 5 March 20).

Tenreyro, S. (2017), "Royal economic society's report on the gender balance in UK economics departments and research institutes in 2016", *Royal Economic Society Women's Committee*, available at: http://personal.lse.ac.uk/tenreyro/women.pdf (accessed 24 May 2020).

Van Noorden, R. (2014), "Confusion over open-data rules", *Nature*, Vol. 515, p. 478.

Wasike, J.M. and Njoroge, L. (2015), *Opening Libraries to Cloud Computing: A Kenyan Perspective*, Libr. Hi Tech News, "Opening libraries to cloud computing: a Kenyan perspective", *Library Hi Tech News*, Vol. 3, pp. 21-24, available at: https://www.emerald.com/insight/content/doi/10.1108/LHTN-09-2014-0072/full/pdf?title=opening-libraries-to-cloud-computing-a-kenyan-perspective.

Wellcome Trust (2010), "Data management and sharing", *Wellcome Trust*, available at: https://wellcome.ac.uk/what-we-do/topics/data-and-artificial-intelligence (accessed 24 May 2020).

Whyte, A. and Tedds, J. (2011), "Making the case for research data management", in *DCC Briefing Papers*, Digital Curation Centre, Edinburgh.

Xia, J., Harmon, J.L., Connolly, K.G., Donnelly, R.M., Anderson, M.R. and Howard, H.A. (2015), "Who publishes in "predatory" journals?", *Journal of the Association for Information Science and Technology*, Vol. 66, pp. 1406-1417.

Corresponding author
Tom Kwanya can be contacted at: tkwanya@tukenya.ac.ke