ISSN 2161-3974
Full-Length Paper

Journal of eScience Librarianship 14 (1): e806
DOI: https://doi.org/10.7191/jeslib.806

Journal of eScience Librarianship
putting the pieces together: theory and practice

Research Data Management at a Medical Facility in Uganda: Practices, Awareness, Challenges, and Recommendation

Edward Mukiibi, Kyambogo University, Kampala, Uganda, emukiibi@kyu.ac.ug
Joyce Bukirwa, Makerere University, Kampala, Uganda

Abstract

The paper explored research data management practices at a medical research facility in Uganda. It focused on the researchers' perception about research data practices, awareness, and challenges. Mixed methods were applied in which thirty (30) respondents out of a population of sixty (60) research community members. The research community was comprised of both research teams and research support members. Whereas the research team respondents were selected randomly and subjected to the questionnaire, the four key informants were purposively selected from the research support members and subjected to the interview. The findings showed variations in perception, management, and understanding of research data practices. Identified challenges were inadequate legal framework, lack of institutionalised storage facilities, and limited competencies in writing Data Management Plans. The recommendations were: the formation of a unit for the development of research data management policy, support services, and the introduction of formal research data management skills training to equip the research community at the facility.

Received: November 3, 2023 **Accepted**: May 1, 2025 **Published**: July 16, 2025

Keywords: Research Data Management, Research Data, Data Practices, Researchers, Protocols, Clinical Trials

Citation: Mukiibi, Edward and Joyce Bukirwa. 2025. "Research Data Management at a Medical Facility in Uganda: Practices, Awareness, Challenges, and Recommendation." *Journal of eScience Librarianship* 14 (1): e806. https://doi.org/10.7191/jeslib.806.

Data Availability: The data set of the study was anonymized and is openly available under Creative Commons License Attribution 4.0 at Zenodo: https://doi.org/10.5281/zenodo.10069454.

Background

Health research is at the heart of human well-being, emphasized at the global level as goal number three of the United Nations Sustainable Development Goals (SDGs). The goal focuses the global health effort towards "ensuring healthy lives and promotion of well-being for all ages" (Ages 2017). Health and well-being are underpinned by health and biomedical research as prerequisites for improving the understanding of healthcare service practices and delivery. In addition to enhancing understanding of diseases, treatment, cure, and development of vaccines for disease prevention is a priority. The search for understanding of diseases, treatment, cure, and development of vaccines results in massive research data being generated. The application of Information Communication Technology (ICT) in health and biomedical science has further resulted in generation of data that if well managed could significantly contribute to informed decision making on individuals health, which could further result in better care.

The application of advanced research technologies in health and biomedical research with low-cost communication is further leading to numerous collaborative research tools and projects between the developed and developing nations (Dunie 2017). Whereas developing nations are sources of cheap research data, it is remitted electronically to developed nations for advanced analysis. This results in generation of knowledge crucial in health and biomedical science, but with no reference to individual contributors from the developing nations who tirelessly work to collect quality data. Funders continue to sponsor collaborative research particularly through health research institutes. The health research institutes have merged specializing in areas that create niches as centres of excellence. It is therefore not surprising that these specialised health research institutes in Uganda hold massive research data generated, used only once, and archived with embargos for years. The archived research data remains unknown to the research community outside the institutes, although it holds value that could be optimised through reuse, generating new knowledge, validating earlier research findings, and advance science that contribute towards saving and improving human lives. Unfortunately, the existing research data are hardly known, inaccessible due to embargos, requiring laborious processes to fulfill the ethical and regulatory requirement for access. This necessitated research data management (RDM) as a best practice in research to derive value and augment the research investment. Research data management practices are currently being advocated for by the international research community as a pertinent element in implementing the FAIR data principles and open science initiatives.

Research data management encompasses all processes and measures that ensure the created research data are well-organized, documented, findable, preserved, stored, backed up, accessible, available, sharable, and reusable. Research data management mitigates data loss, eliminates duplication of efforts, supports validation of findings, and facilitates the creation of new knowledge by building on the work of others (Briney 2015). Research data management practices also contribute towards increasing opportunities for accessibility, discoverability, preservation, transparency, and reproducibility which are key components of

Open Science (Wiley and Burnette 2019). In addition, RDM practices mitigate data deluge, and ad hoc data practices (Tenopir et al. 2011).

Uganda has been at the forefront of health and biomedical research since the early 1930s. However, Uganda as a country is yet to incorporate RDM practices in the research ecosystem, though its desirous of implementing FAIR data principles and open science initiatives as part of participation in the global research agenda. Uganda has had unfortunate incidences of disease outbreaks that have turned into global epidemics. During the 1990s at the peak of the Human Immuno Virus (HIV) and the Acquired Immunodeficiency Syndrome (AIDS) pandemic, both government and the private sector established several initiatives to address the challenges of the day. Legal policy framework requirements saw the establishment of Uganda AIDS Commission (1992); care and support, The AIDS Support Organization (1987); Uganda Women Effort to Save the Orphans (1986); and understanding of the medical evolution of the HIV/AIDS phenomenon, Joint Clinical Research Centre (JCRC) (1991). The HIV which causes AIDS had challenged the world and there was no evidence of medical care or scientific explanation to address the devastating effects of the diseases among the population. Thus, collaborations became common among countries with high HIV/AIDS incidences. The collaborative research provided the platforms to mobilise resources to address the challenges both locally and internationally.

It was during this period that, the JCRC was established to focus on HIVAIDS medical research. The centre has evolved into a robust research system with integrated infrastructure concurrently supporting over 20 clinical trials today. Ninety percent (90%) of the clinical trials are supported through collaborative networks of both local and international scientists, research funding agencies, pharmaceutical companies, research institutes, and universities. The Joint Clinical Research Centre is generating massive research data much of which has been preserved and archived for the last 30 years. However, despite the massive research data held at JCRC it remains unknown to the research community and access is highly restricted by funders' embargos and ethical clauses which render the data inaccessible, unusable, and underused contrary to the global requirement for implementing FAIR data principles and open science (Clarkson and Clarkson 2020; Wilkinson 2016). This necessitated the need to explore how researchers at JCRC perceive research data practices, its management and awareness, and identification of the challenges experienced in the course of managing research data.

The study objectives were to:

i. Explore how researchers perceive research data management practices at JCRC.

ii. Examine the extent to which the researchers were aware of research data management practices.

iii. Identify the challenges experienced by the researchers in managing research data at JCRC.

Literature Review

In health and biomedical science research, the data practices are part of the culture in the discipline as a community of practice. Applying standardized data practices is a culture embedded in the systematic implementation of research protocols (Celebi et al. 2020; Federer et al. 2015; Tam et al. 2014). However, variations in data practices exist due to differences in the discipline, laboratory, and individual researchers' cumulative experiences. This calls for a deeper understanding of the community-specific practices for effective integration at the global scale (Wynholds et al. 2012). Searle et al. (2015) also found researchers' data practices being determined by individual attitudes which influence the behaviors in data collection, analysis, sharing, and reuse. Further urging that understanding the researchers' perception was crucial to support the designing of appropriate measures for the successful implementation of RDM practices and services.

The importance of research data management was given further prominence when the White House, Office of Science and Technology Policy (OSTP) in 2013 release called for increased public access to the results of federally funded research (Holdren 2013). The original release was recently updated and set December 2025 as the deadline for imposing embargos of 12 months. It explicitly stated that policies should be updated to make publications and research funded by taxpayers publicly accessible, without embargo and to be free on public release; establish transparent procedures to ensure scientific and research integrity is maintained in public access policies; and coordinate with OSTP for equitable delivery of federally funded research results and data (Nelson 2022). The policy shift caused major funders and publishers to espouse open data as default. The International Council of Science Union, (2014) endorsed the Organization of Economic Co-operation and Development (OECD) principles and guidelines for access to research data from public funding referring to open access as should be easy, timely, user-friendly, and preferably internet based. The International Council of Science Union further observed that in low-developing countries the need for an unrestricted flow of scientific knowledge to support local research required intentional promotion and maintenance of trusted, sustainable data repositories which are yet to be established. This could be attributed to as impeding access to research data and consequently deterring data sharing and reuse which limits knowledge generation and innovation.

Several studies have asserted that RDM has become a responsibility of researchers among other roles (Tang and Hu 2019; Chiware and Mathe 2016). However, many are still taking on this role without the necessary supportive services and know-how (Birkbeck et al. 2022). The challenges are more acute in low-developing nations where RDM practices are yet to be adopted (Renwick et al. 2017; Van Tuyl and Michalek 2015). In addition, Patterton et al. (2018), who studied RDM in resource-constrained institutions in the Republic of South Africa, found limited awareness about the RDM concept among researchers. Furthermore, the authors even found variations in handling research data between higher education and health institutes noting that research is conducted more by individuals in higher education and by teams in health institutions a common situation in most of developing nations. This presents a gap in practice and services which calls for strategies to align the individual differences for effective application. Studies further indicated individual researchers to be poor stewards of data over the long term, creating unique ad hoc approaches to organise data that become inaccessible and at risk of loss (Vines et al. 2014). It was further observed that researchers were engaging in research data practices with theoretical enthusiasm which had not been translated into practices due to absence of the supportive services. Additionally, the research terrain in health research institutions is characterised by an emphasis on complying with the funder's requirement rather than strategies to optimise the use of the generated data.

Health and Biomedical science have become more data-intensive, and researchers are faced with data management challenges, and needs (Martine 2013). This is further heightened by the demand for FAIR data principles to improve knowledge sharing to address crucial health challenges at the global scale. The International organisations including, the Committee on Data of the International Science Council (CODATA), the International Science Council-World Data Systems (ISC-WDS), and the Research Data Alliance (RDA) are advocating for RDM as good research practice for Open Data and Open Science. The challenges impede RDM implementation which is more evident in low developing countries.

In Health and Biomedical research institutions, research data are highly valued, and controlled by diverse internal and external regulatory, ethical frameworks and protocols which shape its management (Knight 2015; Marutha 2018). This prompts researchers and institutions to develop ad hoc measures to manage research data which in most cases become irretrievable (Wallis et al. 2013). In Low Developing countries, researchers and institutions may not even be aware of the research data management practices but rather unconsciously implement the protocol in fulfillment of the funders' requirement which makes many unaware of RDM practices in reality.

Methods

The mixed methods applied both quantitative and qualitative approaches to data collection. The method adopted for the study was found to be ideal since the primer study focused more on generating baseline information about research data. The mixed methods approach addresses the weaknesses of one method with the strengths of the other method and in the process eliminated possible pitfalls of a single method. Both the questionnaire and interview questions were derived from the Data Asset Framework (DAF) (Jones et al. 2009) and the Digital Curation Centre Lifecycle Model Framework (DCC) (Higgins 2008). This eliminated ambiguity and ensured that only applicable questions were included.

The Joint Clinical Research Centre serves Uganda through five sites with its Headquarters at Kampala, Lubowa. The Headquarter has a total population of 140 staff composed of research teams and research support members. The study considered research and research support staff at the headquarters being the coordinating centre of research relate activities for the whole organisation. The number of staff involved in both research teams and research support vary based on the number of projects being implemented. At the time of collecting data for this study there were a total of 60 staff including medical doctors, clinical officers, medical nurses, medical counselors, pharmacists, laboratory technologists, radiographers, and sonographers, data officers and managers, system administrator, medical records officers, and quality control and quality improvement officers.

The questionnaire was pre-tested at two institutions with similar research culture. This was carried out at both the Infectious Diseases Institute and The AIDS Support Organisation. The purpose of pre-testing the data collection instrument was to ensure sound, consistent, and relevant data are collected during the actual answering of the questions posed. The questionnaire and interview guide were further improved by the researcher's supervisor. Thereafter the questionnaires were administered to 30 respondents who were randomly selected from among the research team at JCRC headquarters. In this study, the members of the research team included Medical Doctors, Clinical Officers, Study Nurses, Pharmacists, Medical Counselors, and Laboratory Technologists. In addition, a Research Assistant was recruited from among the headquarter staff to support the distribution of the questionnaires to the identified members. All respondents to the questionnaire were required to consent to participating in the study, receive the questionnaire, answer it, and be returned within 14 working days. In this study, the four research support staff were Data Manager, System Administrator, Medical Record Officer and Quality Control and Quality Improvement Officer. The four research support staff were purposively selected with the help of the JCRC Research Administrator and were interviewed by the Researcher. Overall consideration included only members who were accessible at JCRC headquarters since the data collection took place during the peak of COVID-19 epidemic when Uganda was under lockdown. Procuring the services of a Research Assistant from the staff, among other strategies, could have contributed to the 100% response rate attained for the questionnaire. In addition, interviews were held with four key informants purposively selected including: a Data Manager, System Administrator, Medical Records Officer, and a Quality Control and Quality Improvement Officer at JCRC headquarters. These four were a crucial part of the research support team since their mandate involved regular handling of research data and ensuring its quality and effective management at the centre.

Other strategies included all the study participants were communicated to by the Research Administrator informing them of the study. This came after receiving all the required ethical and administrative clearances. The research proposal had earlier been approved by Makerere University College of Health Science-Research Ethics Committee under registration no. 2021-141, and Uganda National Council of Science and Technology (UNCST). The protocol had also received clearance from the JCRC Research and Ethical Committee and the Chief Executive Officer. The researcher then approached the selected respondents with a personal email requesting them to take part in the study. Members of the research team were given a self-administered questionnaire to be filled within fourteen working days. The purposively selected key informants were also approached and requested for an appropriate date and time for conducting the interview. The interview lasted on average approximately one hour at the respondent's respective place of work. Overall, data for the study was collected between October 2021 and April 2022.

The questionnaire had seven sections with a total of 15 questions and sub-questions. The questions were mixed: some had closed questions with answers to select from and open questions with an option of "Others specify" seeking independent views from the respondents about the different aspects of RDM. In some questions the respondents were required to select more than one answer and all those that could apply in the circumstances. The interview guide similarly had six main and four sub-questions. Supplementary data was collected from documents extracted from the JCRC's website during the period. The objectives of the study guided the drafting of the questionnaire and interview questions. The results of the questionnaires were input into the RedCap database and applied Microsoft Excel 2023 to support the descriptive analysis. The results of interviews were transcribed using Microsoft Word 2023 and analysed thematically with the support of Atlas ti22. The data set of the study was anonymized and is openly available under Creative Commons License Attribution 4.0 at Zenodo: https://doi.org/10.5281/zenodo.10069454.

Results

The result of the study presents the perceptions of the respondents towards research data management practices, awareness and challenges encountered. Thirty respondents returned a fully filled questionnaire, among whom 12 (40%) had spent 1 to 5 years at the JCRC, 10 (33%) had spent between 6 to 10 years, 6 (20%) had spent 11 to 15 years and only 2 (7%) had spent less than one year at the JCRC. Most of the respondents had a good understanding of the JCRC and its operations given the experiences held and were deemed to be the right person to answer the questionnaires. 14 (47%) had been members of the research group, 10 (33%) had played a role in generating research data based on the specialization held, and 6 (20%) had been Research Assistants. The research team were involved in different research data lifecycle activities, generating, maintaining, preserving, and archiving research data. Respondents were responsible for ensuring quality and remission of research data to the next level to make complete data sets. As such the respondents had good working knowledge and were well-placed to talk about the research data management-related issues and processes at the JCRC. Both the research team and research support members had personal experience of having participated in the research data lifecycle at the JCRC at one point in time.

In Figure 1, 11 (36%) respondents identified existence of experimental data generated from the laboratory work. This is data generated out of bio-samples extracted from study participants and are analysed for specific results. Participants in the studies are selected from the triage process of all patients who attend the clinics or are admitted to the JCRC. Those who meet the pre-set criteria are sent to the medical counselors for further handling and recruitment in the study on consent.

According to one of the interviewees

> *"The results from the bio-sample are analysed in the laboratory based on the clinician's request as guided by the protocol being implemented. After generation, they are presented and integrated in other earlier created data sets for each participant."*

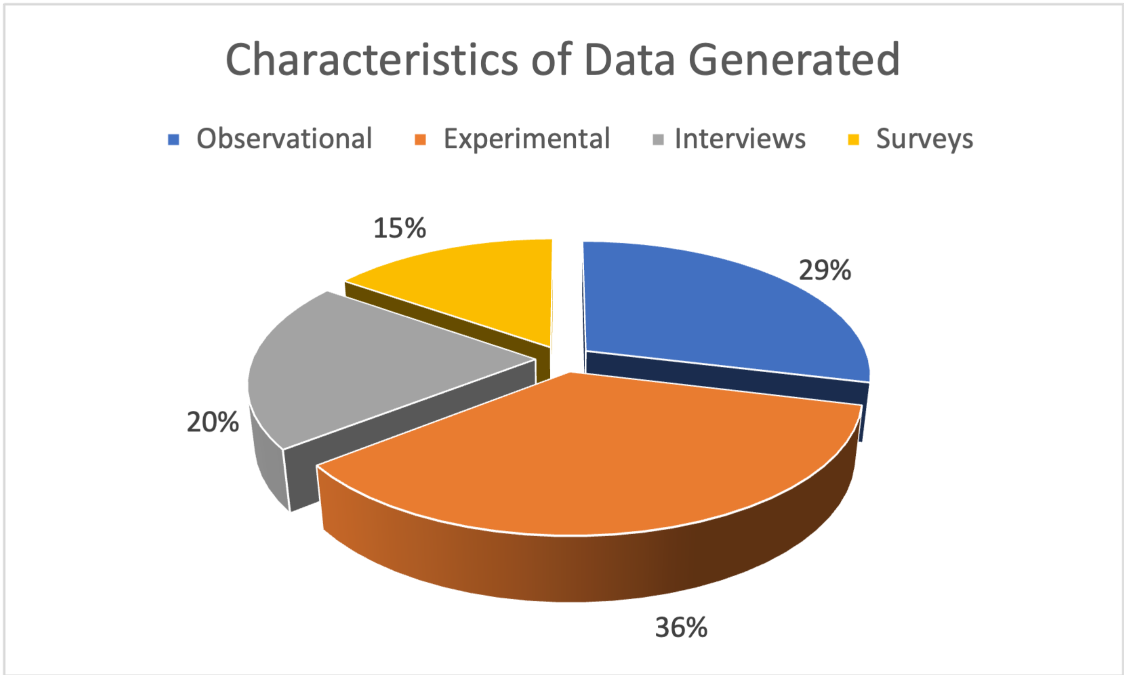

Figure 1: Key characteristics of data generated at the JCRC.
Note: The characteristics of the research data generated were derived from the research data collected as part of the study on research data management practices of selected research institutes in Uganda. Data was collected between October 2021 and April 2022.

Completed data sets are then input in pre-defined databases and remitted to the funders in a batch on a weekly basis until all the anticipated data is collected and submitted. Once the number of participants has been achieved, and all the required data extracted at specified intervals, the data sets are combined with other data sets from other sites for analysis. The characteristics of research data mostly generated at the JCRC confirmed earlier findings showing diverse categorisation and formats as in the study of Liu and Ding (2016).

According to the responses, the type of data generated presented in Figure 2 indicate diverse data sets attributed to use of different tools and equipment as guided by the protocols being implemented. This was further ascribed to by application of Information Technology in health as confirmed by studies of Fuhr (2019) and Manurung (2019). One of the interviewees further confirmed that

> *"Most common data types were medical and laboratory notes, text corpus, images, and databases generated using ICT tools."*

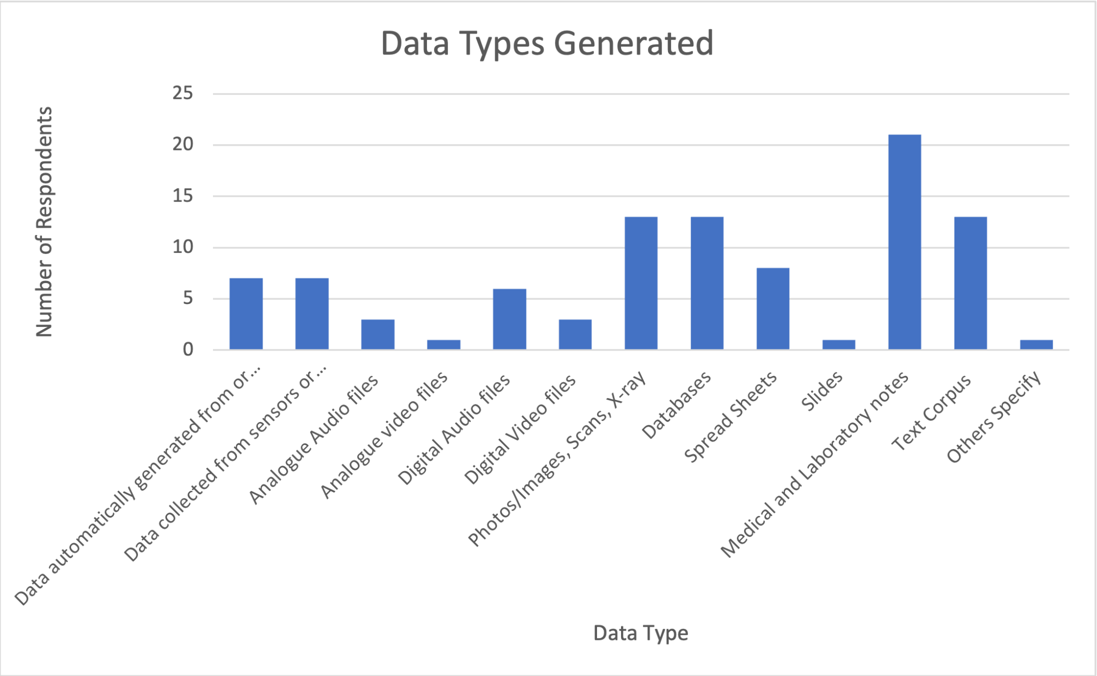

Figure 2: Data types generated from the research at JCRC.

Note: The data type generated was derived from the research data collected as part of the study on research data management practices of selected research institutes in Uganda. Data was collected between October 2021 and April 2022.

The protocols being implemented determines the type and nature of databases designed for the purpose. The System Administrator and the selected Data Manager for a specific project are responsible for designing the databases. Sometimes the funders may come with an already tailor-made database that is installed on the local servers by the System Administrator. The Data Manager and the respective Data Officers responsible for a specific project are required to have the necessary technical competencies to handle such databases. The databases are then synchronised with the key attributes defined by the protocol and formatted in form of Clinical Record Form. This allows research data to be generated in both digital and captured in physical formats too. The physical duplicate is like the digital copy and therefore managing two parallel processes in managing the research data. This has implications on research resources and challenge to management since the physical data are under the custody of the Medical Records Officers and the digital copies are under the management of the assigned Data Manager and the System Administrator. In addition, this creates unnecessary discrepancies since both the physical and digital research data are subjected to different policy requirements.

It was revealed that 21 (75%) of all the respondents didn't know the amount of physical or digital research data currently available at the JCRC. 3 (11%) estimated the digital data to be between 50 to 100 TB, 2 (7%) estimated it to be between 1 to 50 GB, 1 (3%) between 50 to 100 GB, and 1 (4%) over 100TB, only 2 (7%) never answered this question. Whereas it was difficult for respondents to know the amount of data generated in digital form, the physical data could be estimated based on the number of files periodically created and held at the Medical Records Registry. The Data Manager and the System Administrator estimated the amount of research data to be in the range of 50Tb to 100Tb. This could be considered as true given that the job mandate for the two officers involved handling data generation, processing, cleaning, and storage for all research projects at the JCRC. According to one of the interviewees

> *"The amount of data generated at the JCRC is in terabytes given the number of projects being implemented."*

The active physical data files are kept in and managed from the medical registry. Two registries exist for adults and pediatrics and the records are kept secure by the Medical Records Officer. Once the project ends or is terminated the files are closed, cleaned, documented, and sent to customized purpose-built spaces for long-term archiving. Whereas the archival files are kept in archival boxes under the oversight of the Medical Records Officer, the digital files are preserved and stored on secondary storage devices on-site or off-site under the oversight of the System Administrator. The funders own the research data and impose embargos for a period of time for exclusive use of the data. After the embargo period, access for possible reuse requires fulfillment of several conditions under the trust of the local Research Ethical Committee, the Research Director, and the Chief Executive Officers of the JCRC.

Respondents accorded the responsibility over research data for both ongoing and completed projects to the System Administrator as shown in Figure 3. The funders, and Principal Investigators at the JCRC, also have responsibility over research data. The Data Manager routinely ensures that all the collected data is cleaned, conflicts resolved or flagged off and completed data sets are sent to the funder through the database designed for the purpose. When funders have issues with the data set, these are highlighted and sent back to the Data Manager for redress. Once cleared, they are resent. In addition, funders employ the services of external auditors and monitors to ensure data quality following the protocol are generated and all the processes adhered to for trusted data.

Asked about when the research data for completed projects is moved to long-term storage, 10 (29%) respondents agreed that data is moved upon completion of each set of experiments. Whereas 9 (26%) respondents did not do anything, but data was automatically moved to archives, 8 (24%) of the respondents agreed that research data was moved to long-term storage immediately after the publication of a paper. Only 7 (21%) respondents agreed that the decision was left to the funders and the Research Director who in most cases doubles as the local Principal Investigator.

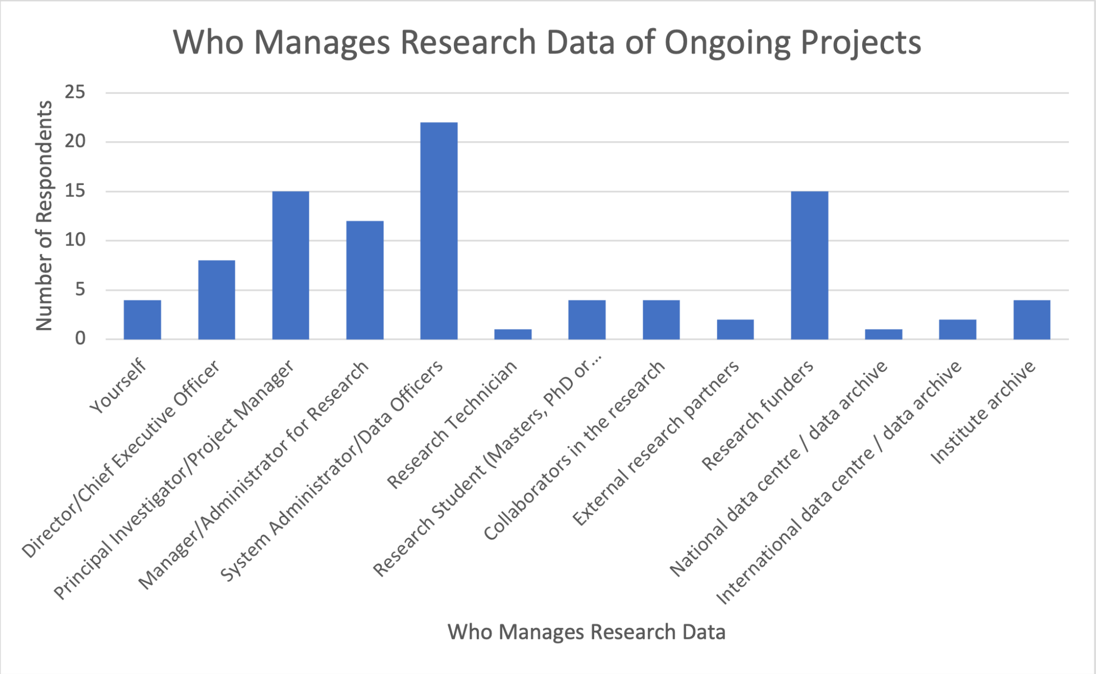

Figure 3: Responsibility for Research data at the JCRC.

Note: The responsibility over the research data for ongoing projects was derived from the research data collected as part of the study on research data management practices of selected research institutes in Uganda. Data was collected between October 2021 and April 2022.

Respondents agreed that physical files are archived at the JCRC. Archive was selected by 20 (67%) in Figure 4, and the digital data is input into different storage facilities, including the JCRC data bank as selected by 5 (17%), and with web-based open research repositories chosen by 3 (10%) currently available and accessible globally. Research data storage for the most recently completed research project were physical files selected by 22 (73%), hard disk drives by 11 (37%), and shared drives/servers by 10 (33%). Only 4 respondents (13%) had used an external drive; 3 (10%) had cloud-based free services and 2 (7%) had CD/DVD as storage devices.

Asked about the frequency of research data backup respondents 12 (43%) were doing it daily, 4 (14%) weekly, and 2 (7%) annually. This was not surprising since most computer systems automatically back up minute by minute to avoid data loss in case of any eventualities. However, 6 (21%) didn't know when their research data was backed up, and 1 (4%) never backed up since the system software had been programmed to back up automatically. The type of research data backed up according to the respondents were 26 (87%) all data collected, 2 (6%) only critical data to the project, and 2 (7%) didn't know. However, one of the interviewees pointed out that

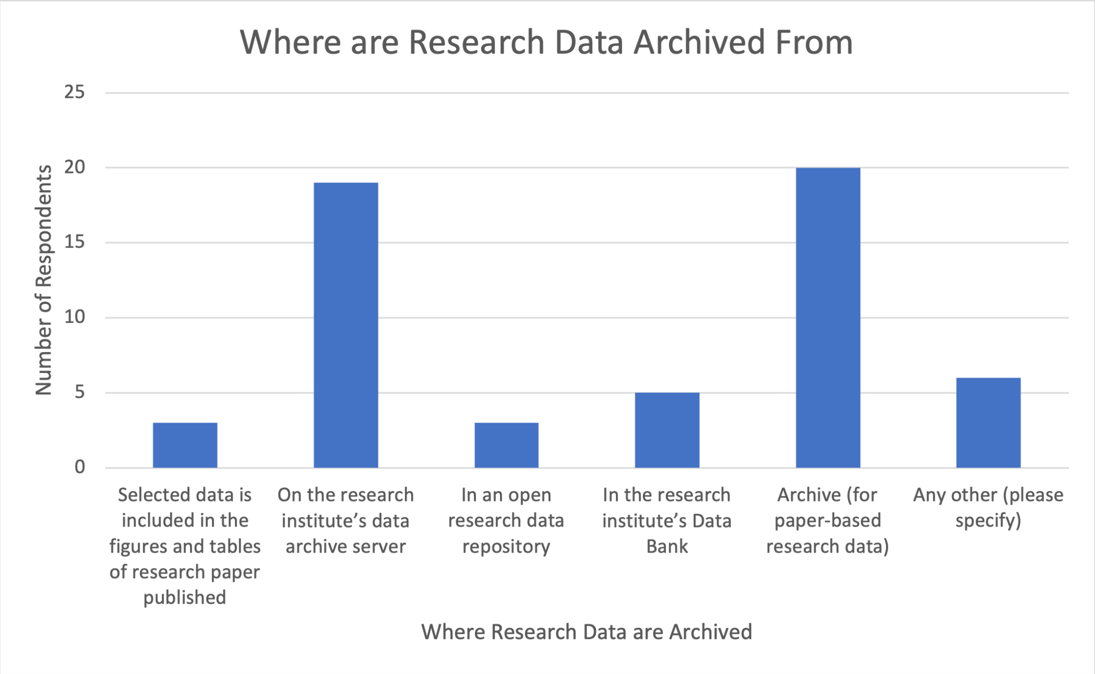

Figure 4: Location of research data archiving.

Note: The location where research data is archived was derived from the research data collected as part of the study on research data management practices of selected research institutes in Uganda. Data was collected between October 2021 and April 2022.

> *"The protocol sets the type of data to be collected, and the research teams are trained to collect only the relevant data for the purpose, so all the collected data is backed up minute by minute by the system to avoid any loss that may have resulted from power failures."*

Research data for completed projects at the JCRC could be discoverable but not accessible due to various reasons as presented in Figure 5. Some conditions should be fulfilled before the data is accessed. From the questionnaire and interviews, respondents agreed that the data could be accessible but under conditions particularly to non-associated researchers. One of the interviewees pointed out further that

> *"Access without following the acceptable due process is termed as 'breach of privacy or intrusion' which is a serious issue in medical research."*

Research teams are well aware of the limitations and always conscious about making research data discoverable particularly to non-associated researchers. Although the data discovery services were available, they required specialised knowledge according to 3 (10%) of the respondents. Nonetheless, research teams

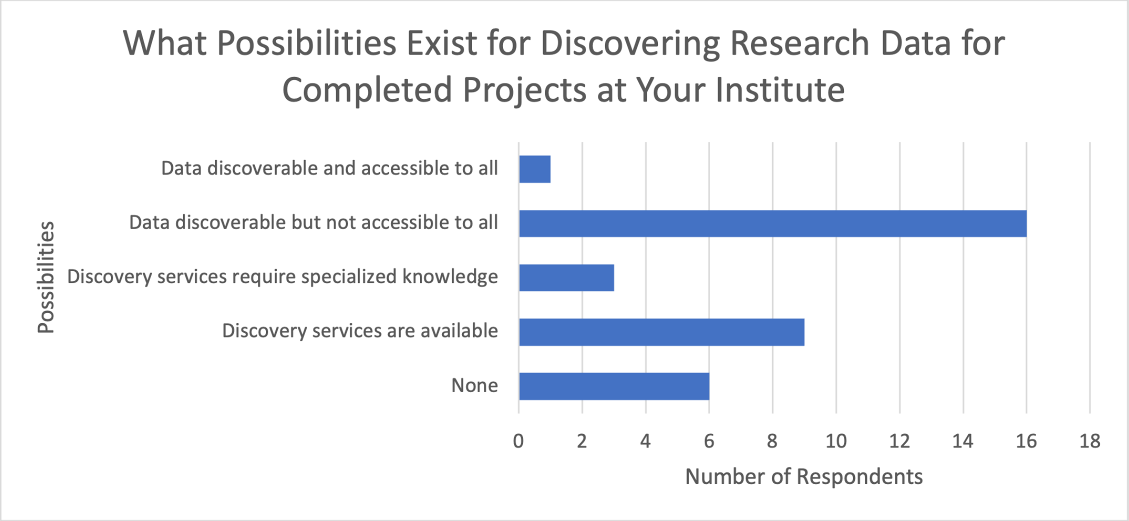

Figure 5: Possibilities of discovering research data for completed projects.
Note: The possibilities of discovering research data of completed projects derived from the research data collected as part of the study on research data management practices of selected research institutes in Uganda. Data was collected between October 2021 and April 2022.

were not aware of the existing tools for discovering research data of completed projects at the JCRC. This is indicative of little interest in existing data. One of the respondents pointed out that

> *"Before the COVID-19 outbreak people were eager to write because many could be sponsored to travel outside Uganda to conference to present the findings. But COVID-19 pandemic stopped all this, and people are no longer motivated to write since they don't see any benefit."*

Another respondent added that

> *"Since they are not in academics, even if you write a number of papers, you will never be recognized later, alone rewarded."*

Apart from funders, research data is shared with supervisors, the research team, and those who participate in data set creation. Research data is also shared with journal publishers as a new demand for publishing, and colleagues in the discipline as indicated in Figure 6. In collaborations, research data is shared using electronic mail to colleagues as agreed by 20 (60%) of the respondents. 9 (30%) of the respondents share research data via storage facilities, databases and only 1 (10%) were using portable storage devices such as CDs, DVDs, and memory sticks. Sharing within teams was considered okay but outside the team is a breach

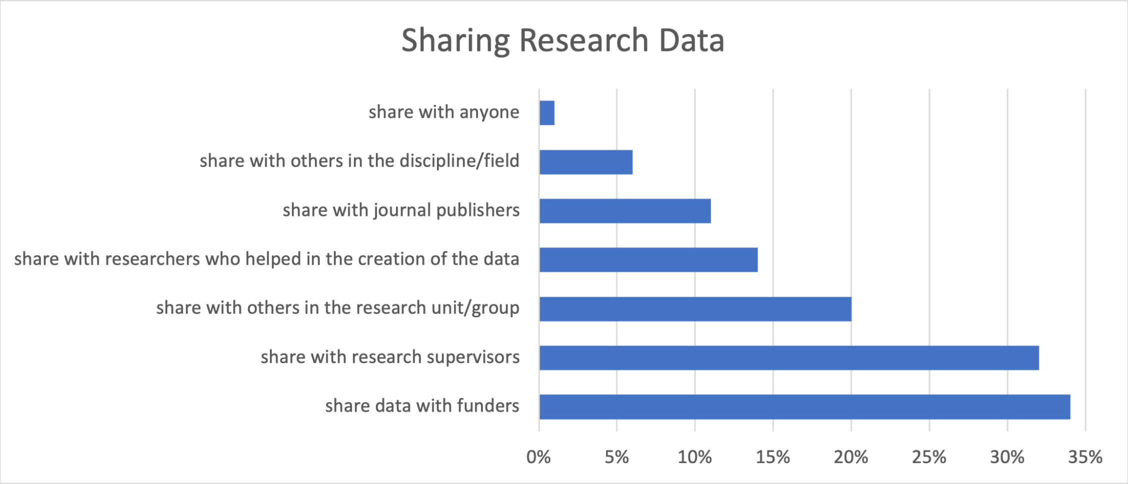

Figure 6: Research data sharing at the JCRC.

Note: The sharing of research data at the JCRC derived from the research data collected as part of the study on research data management practices of selected research institutes in Uganda. Data was collected between October 2021 and April 2022.

that pose serious repercussion to the study. Research teams had no experience in data reuse at the JCRC. One of the respondents pointed out that

> *"Data had embargos, was not openly accessible and required coming up with a substantive protocol with a justification for reusing existing research data. In addition, the protocol required ethical and administrative clearance and one of the local senior staff as a supervisor."*

This was restrictive and the research team members lacked the motivation to reuse the data they helped to generate after the embargo. On the other hand, the research community lacked the information about the research data available and the awareness that such existing data could be reused. There are very limited venues available for non-associated members to know current research projects and the probable data generated and available at the JCRC. A respondent pointed out

> *"The venues used to let the public know about the research projects at the JCRC were basically conference presentations which take place outside Uganda."*

It is believed that in the absence of local or international data publication journals, locally organised research dissemination conferences and open science days, the visibility of the research undertakings at the

JCRC remain unknown to the public. It was further noted that data reuse was anticipated to take place only after the embargo period and with clearance from the funders. Furthermore, the research teams lacked the analytical skills and were faced with the hurdle of accessing the whole project data collected from the diverse sites. Funders legally allow retention of data only collected on the site. This renders the available data to be statistically negligible for meaningful analysis. An interviewee further affirmed that

> *"The lack of skills and support services coupled with the stringent requirement to fulfill before one could access the available research data, in addition to fear of insignificant data deter interested non associated researchers from any attempt to reuse the research data available at the JCRC."*

The outstanding challenges in order of prevalence are presented.

Table 1: Challenges experienced in the course of executing research data management practices.

Challenge	Number of respondents	Response rate
Absence of training programs for skills development in RDM	8	27%
Lack of appropriate institutional storage system	6	22%
Security of research data in an increasingly digital environment	6	22%
Absence of harmonized RDM legal policy and regulation at institute level	6	21%
Low ability in writing Data Management Plan	6	21%
Lack of supportive services to researchers about RDM	6	20%

Note: The challenges experienced in the process of executing research data management practices at the JCRC were derived from the research data collected as part of the study on research data management practices of selected research institutes in Uganda. Data was collected between October 2021 and April 2022.

The challenges highlighted in Table 1 included: the absence of training programs for skills development, lack of an appropriate institutional storage system, security in an increasingly digital environment, absence of harmonized RDM legal, policy, and regulations at the institutional level, low ability in writing a Data Management Plan, and lack of RDM supportive services at the JCRC. Several recommendations have been derived from the study to integrate RDM practice in the research ecosystem at the JCRC as a measure to enhance its competitiveness in an increasingly global research world advocating for implementing FAIR data principles and open science initiatives.

Discussions

The Data Asset Framework provides information on the state of the research data by identifying its location, who is (or is not) responsible for data, as well as awareness of research data management practices (Cox and Williamson 2015). The Digital Curation Centre Model provided insight into the research data practices and provided a platform for identifying the existing challenges. The variables identified were data types/format, volumes of research data held, storage location, responsibility for research data, backup frequency, and awareness of legal/policy frameworks and challenges related to research data management at the JCRC.

The state of research data at the JCRC showed the formats to be mainly experimental at 21 (36%) respondents and observations 17 (29%). The data types were mostly medical and laboratory notes at 8 (25%) respondents, text corpus, databases and photos/images, scans, and X-rays were 4 (13%). This agrees with Martone's et al. (2018) assertion of the existence of an all-encompassing nature of research data covering both quantitative and qualitative types and aligned to the OECD (2007) definition of research data. In addition, much of the data is collected using standardized tools based on the procedures for compiling medical-related data (Health Management Information System: 2010).

Most of the team members, 21 (75%) were not aware of the amount of research data stored, and only 3 (11%) could estimate the digital data to be between 50Tb to 100Tb. The enormous research data born digital and also in physical formats are currently being created at the JCRC and this could be attributed to the e-science paradigm driven by data-intensive research which is bound to grow further beyond the current levels (Choi and Lee 2020; Hey et al. 2009). Although responsibility for research data is not necessarily with researchers per se it lies with the research institute (Engelhardt et al. n.d.). Petersen et al. (2020) urged that it lies with stakeholders with different expertise and experiences which create discrepancies in its management. However, funded projects render the responsibility to the Principal Investigators and the institute (Khodyakov et al. 2013; Tenopir et al. 2015). The person responsible for digital research data at the JCRC was the System Administrators and Data Managers for the different funded projects. These have the expertise in handling research data for both ongoing projects with 22 (73%) respondents and completed projects with 17 (57%). However, the two officers are answerable to the Principal Investigator/Research Director and the Chief Executive Officer who ultimately are responsible for all the research data generated and are accountable to the funders.

Archiving is part of the preservation of research data for long-term value, access and reuse (Yoon et al. 2017; Qin 2013). Archiving both the physical and digital research data is important but more crucial is the ability to retrieve it for reuse once needed. Twenty (67%) respondents believed that research data in physical formats are archived at the JCRC and the digital data at the JCRC's servers. However, archive management is ineffective resulting in exponential costs for maintenance of resources with no possible returns on investment. According to 22 (73%) respondents, most research data from recently completed projects remain in physical formats and only 11 (37%) respondents keep it on hard disc drive. The number of physical files is growing and with no substantive Archive, research data is bound to get lost. In addition, duplication takes place as every digital copy has a corresponding physical file in funded research which waste the research resources. This could be mitigated by establishing a Digital Archive for both physical files (through digitizing) and born-digital research data. This is bound to save JCRC from the acute space constraints and reduce costs incurred for preservation of the physical files for a long term.

Backup is a preservation requirement that ensures that research data generated and the different versions are tracked to identify the action areas accomplished by the different persons to give credence to the data sets (Pryor 2012; Verbaan and Cox 2014). Patterton et al. (2018), found lack of sufficient backup knowledge and experience in adding metadata in an earlier study in the Republic of South Africa. The current study found 12 (43%) respondents were aware of the need for daily data backup, 1 (4%) had never backed up and 26 (87%) believed that all collected data is automatically backed up. Once the data has been safely stored, it enters a period where it could be accessed for re-use. In addition, preservation actions associated with collecting, structuring, indexing, and cataloging data sets for users take place with an emphasis on data discoverability and accessibility to the public. In the study, 16 (53%) respondents noted that data is discoverable though not accessible to all. Nine (30%) respondents believed that discoverable services were available but required specialized knowledge. Fifteen (50%) respondents noted that research data was openly accessible on request after the embargo and with other conditions for both the team members and the public. Making research data discoverable by secondary users require protection of data quality, integrity and facilitating reuse which are dependent upon adoption of RDM practices (Borghi et al. 2018). Research data management has also been acknowledged to lead to well-managed research data that is discoverable for reuse increases reproducibility, and productivity, and optimize research resource usage (Beitz 2013).

Sharing of research data makes data work for people and society as a force for good (Mwangi et al. 2021). Tam et al. (2014) found variance in the sharing practices across disciplines and sub-disciplines. Further noting that sharing is affected by the exponential growth in research data and commended seamless sharing (Si et al. 2013; Zotoo and Liu 2019). This is the same advocated for by key international organizations CODATA, RDA, and WDS. Schöpfel et al. (2021), further urged that sharing of research data opens new perspectives for researchers to obtain the knowledge needed for their research. Therefore, policies and guidelines are needed to enforce data sharing and the policies should be developed in consultation with relevant stakeholders (Rani and Buckley 2012; Edwards et al. 2012). This is envisaged to curb "free-riders" who wait to harvest research data for reuse before the primary data collectors can do so (Nobes, 2018). Findings showed that data is commonly shared with funders 22 (34%), supervisors (21 (32%), and researchers who helped to create it 24 (30%). It was further noted that research data is shared with colleagues using emails 22 (59%). The precursor to data sharing is good data management and stewardship (Wiley 2020) and in health and biomedical sciences sharing is commonly built over time through collaborative networks and acknowledged co-authorship (Parker and Bull 2015; Federer et al. 2015a). Research data sharing is built on trusted relationship and it difficult for data owners to release their data to anybody.

Collecting quality research data requires resources, commitment and selflessness, many researchers in low developing countries due to the hurdles they go through to collect this data are reluctant to share it due to fear non associate researchers taking advantage at their expense.

In collaborative research, sharing is mandatory as it reduce wasteful duplication and prevent data losses (Borghi et al. 2018; Lowndes et al. 2017). Collaborative research was found to be important in JCRC context due to the provision of funds for improving the research infrastructure, developing local research competencies, and expediting knowledge generation (Ecuru et al. 2008; Kamya 2017). Neema (1999) noted that two-thirds of the research completed and ongoing were in health and biomedical science and only one-third were in the other fields due to the availability of funds. The Joint Clinical Research Centre has been competitive in this area and its research productivity for the period up to 2015 was in the fourth position in Africa (Uthman et al. 2015). This could further be attributed to collaborations and 23 (46%) of the respondents were engaged in international collaborations and consortia.

The Data Centre Curation Model (DCCM) (Higgins 2008), defines the process and policies that underpin; the creation, organization, preservation, storage, access, sharing, and reuse of research data which together inform the research data management practices at institute. Current research data management practices for funded projects are pre-defined by the protocols for generating quality data. The demand for compliance and quality data forced the JCRC to assign one team member as the Quality Assurance and Quality Control (QAQC) Officer responsible for keeping an eye on compliance issues before the external auditors and monitors could do so. This compliments the individual team member's effort to deliver quality research data. The quality of research data is accomplished through adherence to the protocol and the Standard Operating Procedures which are aligned to research data management practices (Kahn et al. 2014; Kim 2013). Funders also employed external monitors and auditors for periodical auditing of the due process. Discrepancies identified by QAQC, monitors, and auditors are addressed on an ongoing basis. Where adverse effects or diversions are identified, the whole project could be halted or terminated until effective measures are put in place for total compliance. Furthermore, the systems software is embedded with controls that alert Data Entrants, Data Officers, and Managers of any possible discrepancies that require attention to resolve before being flagged off.

The research team members showed little awareness of research data management practices but rather were driven by the need to comply with protocol requirement. Funded protocols are developed by researchers from the developed countries who respond to funders' call. Most of the funders require Data Management Plans as a pre-requisite for RDM for funding approval. The protocols embed RDM within the actual research workflow since they are compliant to the funders policy requirement. The data management practices are therefore embedded in the protocol and the database software. Thus, the research team members with no prior awareness implement the protocol while intuitively are carrying out RDM practices. The contradictions exist for non-funded research data generated from routine medical clinic works. During the routine medical clinics, a standardised CRF is used to collect data that could later be extracted for research work. However, the CRF are subject to medical records management processes which focus more on surveillance than preserving the research data for FAIR. Consequently, non-funded research data receive little attention and rather focuses on privacy and confidentiality. The CRF are used to extract and post the mandatory information required in the Health Management Information System for surveillance and improving healthcare management decisions (Ministry of Health Resource Centre 2010).

Several challenges were identified at the JCRC attributed to hindering effective research data management. The challenges included government funding gaps which leave JCRC to fend for itself through participating in funding calls for competitive research grants. Funded projects once won have been acknowledge for contributing towards improving the research infrastructure, procuring specialized equipment, and reagents, improving researchers' competencies, and other necessities for creating a competitive research environment (Patel 2016; Tenopir et al. 2011). However, receiving a research grant in a highly competitive environment disadvantages the JCRC. Since funders have many options at both the national and regional level with many who are interested to be considered. This gives funders leeway on most of the critical issues and as the negotiations focus on equipment and developing researchers' competencies, research data is usually neglected to the periphery. Given the limited awareness of the crucial importance of research data, negotiations rarely arise about it and in most cases research data is automatically surrendered to funders' ultimate control (Kennan and Markauskaite 2015). Funders decide the embargo and archiving periods at the JCRC's cost even after the project term. Funders also remotely control how the data they paid for to generate are managed, accessed during and after the embargo periods, and by whom. The JCRC that archived the data seemed to have limited control as the research team and support members who worked tirelessly to collect, organize, preserve, and archive the data may not know what happens but rather shift attention to new projects. This also comes at a time of increasing storage costs for physical data and shifting research interests and priorities (Jao et al. 2015).

The JCRC has no policy on research data management despite having a strategic objective that desires to increase the reuse of existing data. The absence of such a policy affects research data management practices and without a coordination unit, RDM development and support services become insurmountable (Chigwada et al. 2017; Ng'eno 2018; Chawinga 2019).

Conclusion

The current data practices at the JCRC combine both the manual and electronic processes from the time of data creation to archiving. There are two types of research data generated; the funded and the non-funded and exhibited discrepancies in which the two are managed. At the JCRC, research is conducted through teams and in funded projects, research data management practices are embedded in the protocol and intuitively implemented. The non-funded research data generated is managed based on the Ministry of Health records management manual. However, it focuses more on inputting data into the National Health

Information Management System for surveillance. Once the entries are made, not much interest is shown, and such data is left to oblivion.

Research data is generated through specialised team members, though most of the members were not aware of the research data management practices but rather focused on adhering to the protocol requirement. The need for protocol adherence is further enforced by Quality Control and Quality Improvement Officers, auditors, and monitors whose duty is to report any deviation for instant redress to ensure the protection of participants and the production of trustworthy and quality research data. The System Administrator and Data managers are responsible for research data for both ongoing and completed research projects. However, finding research data in physical form for completed projects was considered easy due to a well-managed Archive. The System Administrator and the Data Managers indicated competence to find digital research data ever conducted at the JCRC though there was no known tool to support this claim. Access to funded research data under institutional custody require an approved protocols after the embargo period and clearance from the funders. Issues of interoperability were not adequately interrogated though the IT infrastructure was aligned to both the national and international requirements which made research data workflow seamless across the different platforms at the JCRC. Research data reuse had not been explored despite its potential to increase the JCRC research productivity.

It is therefore recommended that a legal policy framework is needed for research data management. Such a framework recognizes research data as valuable resources with the potential to increase knowledge productivity at the JCRC. Furthermore, the framework provides the basis for future negotiations considering data ownership, open access, and access to related research data from all sites involved in the study. The submitted protocol for ethical clearance should also include a Data Management Plan which should indicate the type of data to be created, organisation, file naming, preservation, storage, access, sharing, and reuse. It should also indicate a repository where the data sets are deposited, and a Creative Commons License applied. In addition, it should provide how the data should be cited. It should be agreed that research data generated from individual research participants should be owned by the individual and co-owned by the JCRC now and in the future. This could contribute to protecting JCRC's future patents that may be derived from such data. The JCRC should digitise the physical files to reduce the storage space demands and work towards opening the available data to the public for reuse. This could unlock the knowledge currently held and increase research productivity at the JCRC. Lastly, there is a need to introduce research data management support services at the JCRC. The Research and Training Department should initiate a process of developing and coordinating RDM and the support services at the JCRC.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgements

This paper is part of the Ph.D. dissertation draft made possible through the generosity of Makerere University Library under MAK/SIDA bilateral support project 374 to Kyambogo University Library. I acknowledge this lifetime opportunity accorded to me in addition to all my respondents and support from management of Joint Clinical Research Centre.

Data Availability

The data set of the study was anonymized and is openly available under Creative Commons License Attribution 4.0 at Zenodo: https://doi.org/10.5281/zenodo.10069455.

References

Ages, A. 2017. "Undp Support to the Implementation of Sustainable Development Goal 3." *United Nations Development Programme*.

Beitz, Anthony. 2013. "Growing an Institution's Research Data Management Capability Through Strategic Investments in Infrastructure." Presented at *International Digital Curation Conference, Amsterdam*, January 2013. https://www.dcc.ac.uk/sites/default/files/documents/IDCC13presentations/1000BeitzIDCC2013.pdf.

Birkbeck, Gail, Tadhg Nagle, and David Sammon. 2022. "Challenges in research data management practices: a literature analysis." *Journal of Decision Systems* 31 (sup1): 153-167. https://doi.org/10.1080/12460125.2022.2074653.

Borghi, John, Stephen Abrams, Daniella Lowenberg, Stephanie Simms, and John Chodacki. 2018. "Support your data: A research data management guide for researchers." *Research Ideas and Outcomes* 4: e26439. https://doi.org/10.3897/rio.4.e26439.

Briney, Kristin. 2015. *Data management for researchers: Organize, maintain and share your data for research success*. Pelagic Publishing Ltd.

Celebi, Remzi, Joao Rebelo Moreira, Ahmed A. Hassan, Sandeep Ayyar, Lars Ridder, Tobias Kuhn, and Michel Dumontier. 2020. "Towards FAIR protocols and workflows: the OpenPREDICT use case." *PeerJ Computer Science* 21 (6): e281. https://doi.org/10.7717/peerj-cs.281.

Chigwada, Josiline, Blessing Chiparausha, and Justice Kasiroori. 2017. "Research data management in research institutions in Zimbabwe." *Data Science Journal* 16 (0): 31-31. https://doi.org/10.5334/dsj-2017-031.

Chiware, Elisha, and Zanele Mathe. 2015. "Academic libraries' role in research data management services: a South African perspective." *South African Journal of Libraries and Information Science* 81 (2): 1-10. https://doi.org/10.7553/81-2-1563.

Clarkson, Melissa D. 2020. "Research Data Management from the STEM Perspective: Reproducibility, Data Reuse, Data integration." *Institute for Biomedical Informatics Presentations* 1. https://uknowledge.uky.edu/bmi_present/1.

Cox, Andrew, and Laurian Williamson. 2015. "The 2014 DAF survey at the University of Sheffield." *International Journal of Digital Curation* 10 (1): 210-229. https://doi.org/10.2218/ijdc.v10i1.362.

Ecuru, J., N. Nawegulo, R. B. Lutalo, D. Kasule, E. Tujunirwe, and I. Akampurira. 2008. "Research in Uganda: Status and Implications for Public Policy." *Kampala, Uganda: UNCST*.

Edwards, Paul N., Steven J. Jackson, Melissa K. Chalmers, Geoffrey C. Bowker, Christine L. Borgman, David Ribes, Matt Burton, and Scout Calvert. 2013. "Knowledge infrastructures: Intellectual frameworks and research challenges." Deep Blue Documents. https://hdl.handle.net/2027.42/97552.

Engelhardt, Claudia, Harry Enke, Jochen Klar, Jens Ludwig, and Heike Neuroth. 2017. "Research Data Management Organiser: A tool to support the planning, implementation and organisation of research data management." In *iPRES 2017*, Kyoto, September 2017. https://phaidra.univie.ac.at/o:931097.

Federer, Lisa M., Ya-Ling Lu, Douglas J. Joubert, Judith Welsh, and Barbara Brandys. 2015. "Biomedical data sharing and reuse: attitudes and practices of clinical and scientific research staff." *PLoS ONE* 10 (6): e0129506. https://doi.org/10.1371/journal.pone.0129506.

Fuhr, Justin. 2018. ""How Do I Do That?" A Literature Review of Research Data Management Skill Gaps of Canadian Health Sciences Information Professionals." *Journal of the Canadian Health Libraries Association* 40 (2): 51-60. https://doi.org/10.29173/jchla29371.

Tansley, Stewart, and Kristin Michele Tolle. 2009. *The fourth paradigm: data-intensive scientific discovery*. Edited by Tony Hey. Vol. 1. Redmond, WA: Microsoft Research. https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery.

Higgins, Sarah. 2008. "The DCC curation lifecycle model." *International Journal of Digital Curation* 3 (1): 134-140. https://doi.org/10.2218/ijdc.v3i1.48.

Holdren, John P. 2013. "Memorandum for The Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research." Office of Science and Technology Policy. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

Jao, Irene, Francis Kombe, Salim Mwalukore, Susan Bull, Michael Parker, Dorcas Kamuya, Sassy Molyneux, and Vicki Marsh. 2015. "Involving Research Stakeholders in Developing Policy on Sharing Public Health Research Data in Kenya: Views on Fair Process for Informed Consent, Access Oversight, and Community Engagement." *Journal of Empirical Research on Human Research Ethics* 10 (3): 264-277. https://doi.org/10.1177/1556264615592385.

Jones, Sarah, Seamus Ross, and Raivo Ruusalepp. 2009. *Data Audit Framework Methodology*. Version 1.8. HATII at the University of Glasgow, May 2009. https://data-audit.eu/documents.html.

Kahn, Michelle, Richard Higgs, Joy Davidson, and Sarah Jones. 2014. "Research Data Management in South Africa: How We Shape Up." *Australian Academic and Research Libraries* 45 (4): 296-308. https://doi.org/10.1080/00048623.2014.951910.

Kamya, Moses R. 2017. "Getting fair and Equitable benefits from collaboration in research with respect to human materials and data." *Annual National Research Ethics Conference (ANREC)*, Kampala-Uganda, July 2017.

Kennan, Mary Anne, and Lina Markauskaite. 2015. "Research Data Management Practices: A Snapshot in Time." *International Journal of Digital Curation* 10 (2): 69-95. https://doi.org/10.2218/ijdc.v10i2.329.

Khodyakov, Dmitry, Susan Stockdale, Andrea Jones, Joseph Mango, Felica Jones, and Elizabeth Lizaola. 2013. “On Measuring Community Participation in Research.” *Health Education and Behavior* 40 (3): 346-354. https://doi.org/10.1177/1090198112459050.

Kim, Jeonghyun. 2013. “Data Sharing and Its Implications for Academic Libraries.” *New Library World* 114 (11): 494-506. https://doi.org/10.1108/NLW-06-2013-0051.

Knight, Gareth. 2015. “Building a Research Data Management Service for the London School of Hygiene and Tropical Medicine.” *Program: electronic library and information systems* 49 (4): 424-439. https://doi.org/10.1108/PROG-01-2015-0011.

Liu, Xia and Ning Ding. 2016. “Research data management in universities of central China: Practices at Wuhan University Library.” *The Electronic Library* 34 (5): 808-822. https://doi.org/10.1108/EL-04-2015-0063.

Lowndes, Julia S. Stewart, Benjamin D. Best, Courtney Scarborough, Jamie C. Afflerbach, Melanie R. Frazier, Casey C. O’Hara, Ning Jiang, and Benjamin S. Halpern. 2017. “Our Path to Better Science in Less Time Using Open Data Science Tools.” *Nature Ecology & Evolution* 1 (6): 0160. https://doi.org/10.1038/s41559-017-0160.

Martone, Maryann E.; Alexander; Garcia-Castro, and Gary R. VandenBos. 2018. “Data Sharing in Psychology.” *American Psychologist* 73 (2): 111-125. https://doi.org/10.1037/amp0000242.

Manurung, Rahel. 2019. “Towards successful research data management.” Master’s Thesis. Aalto University. https://urn.fi/URN:NBN:fi:aalto-202001191726.

Marutha, Ngoako Solomon. 2011. “Records Management in Support of Service Delivery in the Public Health Sector of the Limpopo Province in South Africa.” Doctoral Dissertation. University of South Africa. https://uir.unisa.ac.za/handle/10500/5737.

Mukiibi, Edward. 2023. “Research Data Management in Selected Health Research Institutions in Uganda.” Zenodo. https://doi.org/10.5281/zenodo.10069455.

Ministry of Health Resource Centre. 2010. The Health Management Information System: Volume 1-Health Unit and Community Procedure Manual. Ministry of Health, Kampala.

Mwangi, Kennedy W., Nyabuti Mainye, Daniel O. Ouso, Kevin Esoh, Angela W. Muraya, Charles K. Mwangi, Careen Naitore, Pauline Karega, Gilbert Kibet-Rono, Sebastian Musundi, Jennifer Mutisya, Elizabeth Mwangi, Cavin Mgawe, Silviane Miruka, Caleb K. Kibet, and OpenScienceKE Collaborators. 2021. “Open Science in Kenya: Where Are We?” *Frontiers in research metrics and analytics* 6: 669675. https://doi.org/10.3389/frma.2021.669675.

Neema, Stella. 1999. “Community participation in essential national health research process: Uganda’s experience.” Kampala: Makerere Institute of Social Research. http://www.cohred.org/downloads/545.pdf.

Nelson, Alondra. 2022. “Memorandum for The Heads of Executive Departments and Agencies: Ensuring Free, Immediate, and Equitable Access to Federally Funded Research.” Office of Science and Technology Policy. https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf.

Ng’eno, Emily Jeruto. 2018. “Research Data Management in Kenya’s Agricultural Research Institutes.” PhD Thesis. University of KwaZulu-Natal. http://ir.mu.ac.ke:8080/jspui/handle/123456789/3261.

Nobes, Andy. 2018. “New Platform Facilitates International Research Collaboration to Help Solve Our Biggest Global Challenges.” *The Scholarly Kitchen* (blog). *Society for Scholarly Publishing*. March 19. https://scholarlykitchen.sspnet.org/2018/03/19/guest-post-new-platform-facilitates-international-research-collaboration-help-solve-biggest-global-challenges.

OECD. 2007. “OECD Principles and Guidelines for Access to Research Data from Public Funding.” OECD Publishing, Paris. https://doi.org/10.1787/9789264034020-en-fr.

Parker, Michael, and Susan Bull. 2015. “Sharing Public Health Research Data: Toward the Development of Ethical Data-Sharing Practice in Low- and Middle-Income Settings.” *Journal of Empirical Research on Human Research Ethics* 10 (3): 217-224. https://doi.org/10.1177/1556264615593494.

Patel, Dimple. 2016. “Research Data Management: A Conceptual Framework.” *Library Review* 65 (4-5): 226-241. https://doi.org/10.1108/LR-01-2016-0001.

Patterton, Louise, Theo Bothma, and Martie Van Deventer. 2018. “From Planning to Practice: An Action Plan for the Implementation of Research Data Management Services in Resource-Constrained Institutions.” *South African Journal of Libraries and Information Science* 84 (2): 14-26. https://doi.org/10.7553/84-2-1761.

Petersen, Mareike, Bianca Pramann, Ralf Toepfer, Janna Neumann, Harry Enke, Jana Hoffmann, and Reiner Mauer. 2020. “Research Data Management - Current Status and Future Challenges for German Non-University Research Institutions.” *Research Ideas and Outcomes* 6: e55141. https://doi.org/10.3897/rio.6.e55141.

Pryor, Graham. 2012. *Managing Research Data*. Facet Publishing.

Qin, Jian. 2013. “Infrastructure, Standards, and Policies for Research Data Management.” In *Sharing of Scientific and Technical Resources in the Era of Big Data: The Proceedings of COINFO 2013*, China, October 2013. https://surface.syr.edu/istpub/164.

Rani, Manju, and Brian S. Buckley. 2012. “Systematic archiving and access to health research data: rationale, current status and way forward.” *Bulletin of the World Health Organization* 90 (12): 932-939. https://doi.org/10.2471/blt.12.105908.

Renwick, Shamin, Marsha Winter, and Michelle Gill. 2017. “Managing research data at an academic library in a developing country.” *IFLA Journal* 43 (1): 51-64. https://doi.org/10.1177/0340035216688703.

Schöpfel, Joachim, and Otmane Azeroual. 2021. “Rewarding Research Data Management.” In *Companion Proceedings of the Web Conference 2021*, Slovenia, April 2021. https://doi.org/10.1145/3442442.3451367.

Searle, Samantha, Malcolm Wolski, Natasha Simons, Joanna Richardson. 2015. “Librarians as partners in research data service development at Griffith University.” *Program: electronic library and information systems* 49 (4): 440-460. https://doi.org/10.1108/PROG-02-2015-0013.

Si, Li, Xiaozhe Zhuang, Wenming Xing, and Weining Guo. 2013. “The Cultivation of Scientific Data Specialists: Development of LIS Education Oriented to e-Science Service Requirements.” *Library Hi Tech* 31 (4): 700-724. https://doi.org/10.1108/LHT-06-2013-0070.

Tam, Winnie, Jenny Fry, and Steve Probets. 2014. “The Disciplinary Shaping of Research Data Management Practices.” In *iConference 2014 Proceedings*, Berlin, March 2014. https://doi.org/10.9776/14338.

Tang, Rong and Zhan Hu. 2019. "Providing Research Data Management (RDM) Services in Libraries: Preparedness, Roles, Challenges, and Training for RDM Practice." *Data and Information Management* 3 (2): 84-101. https://doi.org/10.2478/dim-2019-0009.

Tenopir, Carol, Suzie Allard, Kimberly Douglass, Arsev Umur Aydinoglu, Lei Wu, Eleanor Read, Maribeth Manoff, and Mike Frame. 2011. "Data Sharing by Scientists: Practices and Perceptions." *PLoS ONE* 6 (6): e21101. https://doi.org/10.1371/journal.pone.0021101.

Tenopir, Carol, Elizabeth D. Dalton, Suzie Allard, Mike Frame, Ivanka Pjesivac, Ben Birch, Danielle Pollock, and Kristina Dorsett. 2015. "Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide." *PLoS ONE* 10 (8): e0134826. https://doi.org/10.1371/journal.pone.0134826.

Uthman, Olalekan A., Charles Shey Wiysonge, Martin O. Ota, Mark Nicol, Gregory D. Hussey, Peter M. Ndumbe, and Bongani M. Mayosi. 2015. "Increasing the value of health research in the WHO African Region beyond 2015—reflecting on the past, celebrating the present and building the future: a bibliometric analysis." *BMJ Open* 5 (3): e006340. https://doi.org/10.1136/bmjopen-2014-006340.

Van Tuyl, Steve, and Gabrielle Michalek. 2015. "Assessing Research Data Management Practices of Faculty at Carnegie Mellon University." *Journal of Librarianship and Scholarly Communication* 3 (3): eP1258. https://doi.org/10.7710/2162-3309.1258.

Verbaan, Eddy, and Andrew M. Cox. 2014. "Occupational Sub-Cultures, Jurisdictional Struggle and Third Space: Theorising Professional Service Responses to Research Data Management." *The Journal of Academic Librarianship* 40 (3-4): 211-219. https://doi.org/10.1016/j.acalib.2014.02.008.

Vines, Timothy H., Arianne YK Albert, Rose L. Andrew, Florence Débarre, Dan G. Bock, Michelle T. Franklin, Kimberly J. Gilbert, Jean-Sébastien Moore, Sébastien Renaut, and Diana J. Rennison. 2014. "The Availability of Research Data Declines Rapidly with Article Age." *Current Biology* 24 (1): 94-97. https://doi.org/10.1016/j.cub.2013.11.014.

Wallis, Jillian C., Elizabeth Rolando, and Christine L. Borgman. 2013. "If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology." *PLoS ONE* 8 (7): e67332. https://doi.org/10.1371/journal.pone.0067332.

Wiley, Chris. 2020. "The Case for Data-sharing Policies and FAIR Sharing Principles: Analyzing Journals and Articles of Engineering and Medical Faculty." Presented at *2020 ASEE Virtual Annual Conference*, Virtual, June 2020. https://doi.org/10.18260/1-2--35295.

Wiley, Christie A., and Margaret H. Burnette. 2019. "Assessing Data Management Support Needs of Bioengineering and Biomedical Research Faculty." *Journal of eScience Librarianship* 8 (1): e1132. https://doi.org/10.7191/jeslib.2019.1132.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al. 2016. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 3 (1): 1-9. https://doi.org/10.1038/sdata.2016.18.

Wynholds, Laura A., Jillian C. Wallis, Christine L. Borgman, Ashley Sands, and Sharon Traweek. 2012. "Data, data use, and scientific inquiry: Two case studies of data practices." In *JCDL '12: Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries*, Washington DC, January 2012. https://doi.org/10.1145/2232817.2232822.

Yoon, Ayoung, Wei Jeng, Renata Curty, and Angela Murillo. 2017. "In between Data Sharing and Reuse: Shareability, Availability and Reusability in Diverse Contexts." *Proceedings of the Association for Information Science and Technology* 54 (1): 606-609. https://doi.org/10.1002/pra2.2017.14505401085.

Zotoo, Isidore Komla, and Guifeng Liu. 2019. "Research Data Management (RDM) Strategy for Academic Libraries in Ghana: Setting a National Development Agenda." *Open Access Library Journal* 6: e5383. https://doi.org/10.4236/oalib.1105383.