

Understanding the legal framework in Research Data Management: a study of selected academic libraries in Kenya

Vol. 9 No. 1

April 2024

Pauline Ruguru Njagi
Kenyatta University
rugurupauline@gmail.com

Gitau Njoroge
Kenyatta University
gitau.njoroge@ku.ac.ke

Ronald Bituka
Kenyatta University
ronaldbituka@gmail.com

Abstract

Rationale of Study – The study focused on the legal framework instrumental in protecting research participants' rights and interests. It also sought to identify potential legal risks and recommend strategies to mitigate them, promote ethical conduct, and enhance security for data reuse and sharing.

Methodology – The study used a quantitative research design to gather quantifiable data. The target population was academic librarians and executive library users from Kenyatta and Egerton University. A total of 35 participants participated in the study. Structured open-ended questionnaires were used to obtain responses from participants purposively selected.

Findings – The study revealed that academic institutions have policies. However, to maintain a healthy research ecosystem that facilitates accessibility, interdisciplinary collaboration, knowledge transmission, and data sharing, academic libraries need to implement research data management policies that match open science standards, such as promoting reuse in the context of adherence to intellectual property rights and security of sensitive data.

Implications – The study sheds light on issues and legal concerns addressed or should be addressed by research data management policies adoption or being prepared for adoption in academic libraries. Its issues are intellectual property, protection of data against security breaches, terms and conditions regarding data use, and protection of sensitive data confidentiality.

Originality – The paper shed light on research data management policy, precisely the issues which should be included in the policy, as well as legal concerns.

Keywords

Research data management, policy issues, policy legal concerns

Citation: Njagi, P.R., Njoroge, G. & Bituka, R. (2024). Understanding the legal framework in Research Data Management: a study of selected academic libraries in Kenya. *Regional Journal of Information and Knowledge Management*, 9(1),93-102.



Published by the
**Regional Institute of
Information and Knowledge
Management**

P.O. Box 24358 – 00100 –
Nairobi, Kenya

1 Introduction

Research data management is an increasingly salient issue as the open science movement encourages culture change around data management, sharing, and reuse by institutions of higher learning (Hess, 2019). Before a project starts, it is essential to think about all the different facets of data management, metadata creation, data preservation, and analysis to ensure that data is appropriately managed today and preserved for the future. RDM is a collection of various data lifecycle activities and procedures that take into account legal constraints, ethical principles, and governance frameworks while designing, creating, storing, protecting, retrieving, sharing, and reusing data (Hess, 2019). Xiaofeng and Xiang's (2013) study on research data management in China discovered that RDM is associated with several activities and processes within the data lifecycle, including design and creation of data, storage, security, preservation, retrieval, sharing, and reuse, which require proper governing framework and strategies to deal with legal and ethical aspects. As noted by Tenopir et al. (2014), in a study done in the US and Canada, the library's role has changed from conventional classifying and cataloguing to a new role in research data management, which requires thorough adaptation and preparation of policies, infrastructure, and revolutionization of the type of services offered. In Africa, the South Africa Data Archive (SADA) was given the mandate and responsibility to enact guidelines and data best practices instrumental in hosting large-scale data collected from national, regional, and international arenas, particularly in areas of humanities and social sciences (Kahl et al., 2015). Kinde et al. (2021) observed that higher education institutions in Ethiopia, despite funders and publishers' pressure on proper research data management, lacked policies and strategies to enhance research data management. A study by Ng'eno and Mutula (2018) on the agricultural research institute in Kenya observed that some developing states had adopted strategies to enhance RDM services. However, inadequacies were noted in ethical deliberations, the legal framework for data standards and procedures contributing to incomplete, inaccurate, and loss of data hindering essential activities such as wide data distribution and reuse.

The activities associated with research data management, such as creating datasets, reformatting for analysis in various software packages, transporting between sites, and processing for possible reuse, are undertakings that require policy to guide on the legal issues and governance of metadata. Having a policy to govern research data management practices provides a structured framework for ensuring compliance, promoting data quality

and integrity, improving data security, facilitating data sharing and access, optimizing resource allocation, fostering collaboration, and safeguarding the organization's reputation and trustworthiness. Since research data management is an area which evolves with open science, it is, therefore, essential to explore academic policies to ascertain the issues and legal concerns covered in readiness to adopt RDM practice or to improve the already implemented policy. Therefore, the study sought to explore the policy legal framework adopted or considered for adoption in preparation for effective research of data management practices.

The objective of the study was to establish the legal framework for the adoption of research data management practices.

2 Literature review

Policies are adopted or proposed statements that guide decisions, plan to tackle specific issues and achieve a rational outcome. Cox and Verbaan's (2018) semi-structured interview with librarians in the UK study enlightened that institutional guidance in research governance, preservation and managing, and rapid data contribute to the slow implementation of RDM policies in the UK compared to the United States. In 2011, the UK Research Council agreed to a common principle on access to research output policy to facilitate data sharing. Although libraries play a unique role in the research process, the UK library associations have influenced the enactment and implementation of the RDM policy (Cox & Verbaan, 2018). At the institutional level, as observed by Cox and Verbaan (2018), some constructs could impact RDM, such as present inclinations, particularities, adopted institution strategies, and broader state standpoints. Wiley's (2018) exploration of 850 journals for dissemination points out that though publishers play a critical role in the RDM process, the study revealed that only 132 had identifiable policies; hence, the existence of journal data policies positively correlates with data-sharing practices among authors. Further, Wiley (2018) states that to ensure additional data validation, preservation, and indexing to facilitate discoverability and reuse by other journal publishers, there is a need to revise data-handling policies, publications, and data-sharing requirements. In Australia, as put into perspective by Shelly and Jackson (2018), a literature review of Australian universities informs that there is a broader adoption of strategies in most institutions of higher learning for functional scholarly activities and open access initiatives in compliance with government funding body requirements.

Similarly, in European countries, as observed by Nhendodzashe and Pasipamire (2017), less than (40.9%) have not adopted policies to guide research data management. Consequently, the policy has been replaced by a new Australia Code for the Responsible Behavior of Research with eight principles for responsible research behaviour, including responsibilities for research institutions and transparency in reporting (Ng'eno & Mutula, 2022). Further, the policy stipulates that those facilities should have infrastructure built with security aspects to safeguard data from unauthorized manipulation. Equally, the policy guides that the researchers allow access and reference to raw data by interested parties and disseminate findings responsibly, accurately, and broadly by maintaining clear, precise, and comprehensive records (Ng'eno & Mutula, 2022). However, Shelly and Jackson (2018) observed that collaboration in the implementation of research data management within universities acts as a challenge to policies and procedures implementation, though the OECD members, including Australia, implemented the Declaration on Open Access to Research Data, a policy adopted by a Public Funding Organization. The Fitzsimmons et al. (2022) policy statement supports FAIR access, which allows research data created in Australia to be easily found, metadata to be available in repositories, and data to be unified by linking with international standards, format, and language that adhere to the FAIR principle (Ng'eno & Mutula, 2022). Although the FAIR policy enacted by the National Health Medical Research Council (2018) may improve study reliability and citation, Shelly and Jackson (2018) observe that it will be challenging to implement the policy in repositories developed using discipline-based practices and protocols on how data should be shared.

Further, adoption and compliance with the FAIR principle require an institution to have open guidelines rather than rules to cater to facets, including linking and reuse, often subjective to open interpretation. Consequently, Shelly and Jackson (2018) enlightened us that implementing and publishing policies on identifiers, metadata, licensing, and protocol can be a good way for repositories to adhere to the fair principle. According to Lefebvre et al. (2020), an explanatory case study in the Netherlands revealed that RDM guidelines adopted in preparation for the implementation of RDM services should stipulate a series of activities, including data collection, integration, preservation, sharing, and reuse, which influence the importance of a policy document to support the development of cases for information technology infrastructure funding, clarification of institutional position, and outline of stakeholder's roles and responsibilities.

According to Nhendodzashe and Pasipamire (2017), the study enlightened that the success of RDM activities, data gathering, preservation, and distribution is determined by the adopted policies. Further, the institutions seeking funding for RDM implementation use policies to spell out critical issues that promote RDM success and prevent loopholes that could adversely implicate research. Nhendodzashe and Pasipamire (2017) state that RDM policies are essential as they stipulate whether members can make a formal request to withdraw their shared research, compensation and safeguarding of research data to avoid manipulation by unauthorized personnel. However, a study by Nhendodzashe and Pasipamire (2017) identified issues that curtail RDM policy implementation, particularly in developing states, including policy treatment as a statement of aspiration as opposed to intent and changes in institutional culture. Further, Nhendodzashe and Pasipamire (2017) cite that institutions in developed states such as Australia often operate without policies to guide data management.

Policy, regulatory framework, and ethical issues are critical variables that enhance effective RDM practice Wong and Chan (2021). Further, Kinde et al.'s (2021) multi-stage sampling in Ethiopian higher education institutions informs of challenges often observed in policies to manage data, educate stakeholders and support open access activities. According to Birkbeck et al. (2022), literature analysis revealed that the solution to challenges associated with research data management is to develop infrastructure and services that conform to specific legal and policy frameworks and respond to storage, security, conservation, quality, compliance, and shared drives. Since the legal environment surrounding research data entails rules, principles, procedures, and plans to enhance data eminence in academic institutions, the adopted policy and legal framework should enhance the implementation of architecture that fosters the management and distribution of scholarly work (Andrikopoulou et al., 2021). Further, the regulatory framework needs inclusivity, comprehensiveness, and details by including critical facets such as patent protection, patents, moral rights, privacy, agreement, internal guiding principles, and procedures (Machimbidza et al., 2022).

Moreover, Chiarelli et al. (2022), a case study at the University of Oxford, informs that a functional RDM legal and policy framework should spell out purpose, scope, applicability, contributors' guidelines, copyright agreements, terms, metadata, data classification, security breaches, and intellectual property concerns. Additionally, the policy should spell out research data ownership to minimize disputes, thus promoting accountability. Equally, the participant's contribution to the construction, formation, certification, application,

preservation, and assertion of RDM policy led to rigorous research data in academic institutions. Ng'eno and Mutula (2022) stated that the RDM policy should stipulate players' roles and responsibilities, which leverage the researcher's needs and accommodate technological dynamics. Therefore, Ng'eno and Mutula (2022) outlined below various elements that should form part of the RDM legal and policy framework.

Proposed Research Data Management Policy	
Stakeholders	Roles and responsibility
Researchers	Provide raw data
Librarians	<ul style="list-style-type: none"> ▪ Facilitate data access and reuse ▪ Curation and preservation, both long/short term ▪ Metadata and data security ▪ Data open accessibility and availability ▪ Research data subject to legitimate protection ▪ A mechanism for storage, backup, and registration ▪ Training and support ▪ Copyright and intellectual property support ▪ Guidelines for record disposal
IT professionals	<ul style="list-style-type: none"> ▪ Research data security ▪ Training and support ▪ Sharing guidelines and restrictions ▪ Removal and data transport ▪ Preferred data licenses

Source: Ng'eno and Mutula (2022)

3 Theoretical framework

The study adopted the Technology, Economic, Legal, Organizational, and Schedule (TELOS) feasibility model coined by Hall (2007). It focuses on the legal aspect as it guides standards, compliance, and ethical conduct when dealing with research data. Research data management requires careful consideration of legal, policy, regulatory, and ethical issues to guide data service activities, including research data construction, evaluation, accounting, conservation, intellectual property, sharing, reuse, and removal (Ssegawa & Muzinda, 2021). The legal framework guides research data custodians on data types and formats that should be considered for archiving (Nhendodzashe & Pasipamire, 2017). Further, policy on research data management guides metadata and data repository for data file storage, use, and data sharing procedures. Morgan et al. (2017) identified various research data policies that institutions preparing to adopt RDM should consider, including curation of data, administration, publication, and distribution.

4 Methodology

The study used descriptive quantitative design to ascertain facts with the intention of ensuring dependability, objectivity, and generalizability of results. The study used closed-ended questionnaires to collect data from two university participants, who were selected purposefully. The study used content validity to establish the degree to which the measure represents the paradigm of interest. Cronbach Alpha test to measure the questionnaire's reliability was 0.785, indicating a moderate to good level of internal consistency among the elements in the tested scenario. The finding was presented using tables for interpretation and understanding.

5 Findings and discussions

5.1 Policy legal framework to govern research data management

A majority (70.97%) of the participants were of the opinion that a policy legal framework is essential in research data management practices. This is supported by Shelly and Jackson (2018), who state that there is a broader adoption of strategies for RDM policy implementation in most institutions of higher learning to support scholarly activities and open access initiatives in compliance with government funding body requirements. Concurring, Chiarelli et al. (2022) inform that a functional RDM legal and policy framework should spell out purpose, scope, applicability, contributors' guidelines, copyright agreements, terms, metadata, data classification, security breaches, and intellectual property concerns. Additionally, an RDM policy should help to spell out research data ownership to minimize disputes, thus promoting accountability. Equally, the participant's contribution to the construction, formation, certification, application, preservation, and assertion of RDM policy led to rigorous research data in academic institutions. The dissatisfaction expressed by 29.03% on RDM policies already implemented concurs with Nhendodzashe and Pasipamire's (2017) study that identified issues that curtail RDM policy implementation, particularly in developing states, including policy treatment as a statement of aspiration as opposed to intent and changes in institutional culture.

5.2 Research data management policy issues

A significant percentage of participants in various roles, including chief librarian K/E, agreed on the presence or planned to adopt a policy in readiness to research data management practices. All the participants in this category had 100% responses on issues which should be supported by the policy, including intellectual property rights of 5.9%,

protection of data against security breaches of 6.7%, terms and conditions regarding data use of 5.9%, and protection of sensitive data confidentiality. Similarly, Deputy Librarian K's response concurred with chief librarians on the importance of a policy, particularly issues that should be covered by the implemented or about to be adopted research data management policy. However, the deputy librarian's opinion is unknown as the response was not captured. The responses are supported by Nhendodzashe and Pasipamire (2017), that policies are essential as they stipulate whether members can make a formal request to withdraw their shared research, compensation and safeguarding of research data to avoid manipulation by unauthorized personnel. The section heads K/E as well 100% strongly supported the adoption of an RDM policy that covers intellectual property 17.6%, 23.5%, protection of data against security breach 20.0%, 26.7%, terms and conditions regarding data use 17.7%, 23.5%, and protection of sensitive data confidentiality 6.7%, 20.0%.

Table 1: Illustrating issues supported by policy

Category	Issues supported by policy							
	Intellectual property aspects		Protection of data against security breaches		Terms and conditions regarding data use		Protection of sensitive data confidentiality	
Chief Librarian K	1	5.9%	1	6.7%	1	5.9%	1	6.7%
Chief Librarian E	1	5.9%	1	6.7%	1	5.9%	1	6.7%
Deputy Librarian K	1	5.9%	1	6.7%	1	5.9%	1	6.7%
Deputy Librarian E	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Section Heads K	3	17.6%	3	20.0%	3	17.6%	1	6.7%
Section Heads E	4	23.5%	4	26.7%	4	23.5%	3	20.0%
Senior Library Assistant K	5	29.4%	2	13.3%	4	23.5%	3	20.0%
Senior Library Assistant E	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Research Directorate Staff E	1	5.9%	1	6.7%	2	11.8%	1	6.7%
Graduate School Staff E	1	5.9%	2	13.3%	1	5.9%	4	26.7%
Total	17	100.0%	15	100.0%	17	100.0%	15	100.0%

Consequently, the senior library assistant from E had a 0.0% response. However, in the same category, 100% of K strongly supported the adoption of a research data management policy that clearly stipulates issues including intellectual property at 29.4%, protection of data against security breaches at 13.3%, terms and conditions regarding data use at 23.5%, and protection of sensitive data confidentiality 20.0%. Also, the research directorate and graduate school category responses were received from E. The participants strongly supported an RDM policy that covers issues including intellectual property 5.9%, protection against security breaches 6.6%, 13.3%, terms and conditions regarding data use 11.8%, 5.9%, and protection of sensitive data confidentiality 6.7%, and 26.7% respectively.

Ssegawa and Muzinda (2021) support participants' perception of issues which should be covered by RDM policy by alluding that research data management requires careful consideration of legal, policy, regulatory, and ethical issues to guide data service activities, including research data construction, evaluation, accounting, conservation, intellectual property, sharing, reuse, and removal.

5.3 Research data management policy legal concerns

When asked about the legal concerns that should be addressed by the research data management policy in preparation for the practice adoption, there were varying perceptions of the issue. The chief librarian K/E and deputy librarian K disagreed on copyright and patent inclusion as a legal concern, but 100% supported data protection against breaches and confidentiality at 7.7%. Consequently, the deputy librarian in E strongly agreed that copyright could be an issue of legal concern. In contrast, data protection against security breaches and confidentiality and patents should not be included in the policy. The section head K/E strongly agreed on issues which should be considered as legal concerns in research data management policy, including copyright at 14.3%, 7.1%, data protection against security breaches and confidentiality at 23.1%, 15.4%.

Table 3: Illustrating legal concerns

Category	Legal Concerns						Total	
	Copyright		Data protection against security breaches and confidentiality		Patent			
	N	%	N	%	N	%	N	%
Chief Librarian K	0	0.0%	1	7.7%	0	0.0%	1	3.4%
Chief Librarian E	0	0.0%	1	7.7%	0	0.0%	1	3.4%
Deputy Librarian K	0	0.0%	1	7.7%	0	0.0%	1	3.4%
Deputy Librarian E	1	7.1%	0	0.0%	0	0.0%	1	3.4%
Section Heads K	2	14.3%	3	23.1%	0	0.0%	5	17.2%
Section Heads E	1	7.1%	2	15.4%	1	50.0%	4	13.8%
Senior Library Assistant K	4	28.6%	1	7.7%	0	0.0%	5	17.2%
Senior Library Assistant E	3	21.4%	1	7.7%	0	0.0%	4	13.8%
Research Directorate Staff E	2	14.3%	1	7.7%	0	0.0%	3	10.3%
Graduate School Staff E	1	7.1%	2	15.4%	1	50.0%	4	13.8%
Total	14	100.0%	13	100.0%	2	100.0%	29	100.0%

The regulatory framework needs inclusivity, comprehensiveness, and details by including critical facets such as patent protection, patents, moral rights, privacy, agreement, internal guiding principles, and procedures (Machimbidza et al., 2022).

Consequently, patents as a legal concern received varied responses from the section heads, with category K agreeing 50%, whereas E disagreed. The senior library K/E strongly agreed that copyright 28.6%, 21.4%, data protection against security breaches, and confidentiality 7.7% are legal concerns which should be considered in RDM policy, whereas patents received 0.0%. Research directors and graduate schools perceived that copyright 14.3%, 7.3%, data protection against breaches 7.7%, and 15.4% are significant legal issues of concern that should be included in RDM policy. Consequently, the categories were perceived differently on patent, with graduate school participants agreeing at 50%, whereas the research directorate staff disagreed. Institutions that have a legal framework are better positioned to ensure that data protection rules and ethical standards are followed. This reduces the potential for legal complications while also demonstrating a dedication to appropriate and accountable research data management. Abraham et al. (2019) support this by saying that data governance is the exercise of authority and control over the data management processes. It attempts to achieve a corporate-wide data agenda, maximize the value of an organization's data assets, and manage data-related risks.

6 Conclusion

The participants' recognition of the value of a research data management policy shows dedication to safeguarding and controlling the creative works produced during research projects. It is acknowledged that strong policies are required with regard to data security, terms of use, confidentiality, and intellectual property. The legal concerns, including copyright, data protection against security breaches and confidentiality, were supported for inclusion in the research data management policy. However, patents, as a legal concern, received low support. Therefore, research data management policy should be carefully considered in light of the issues and concerns identified for successful implementation and functional practice.

7 Recommendation

The study shed light on research data management policies adopted or being prepared for adoption in academic libraries. Essential issues identified, including intellectual property, protection of data against security breaches, terms and conditions regarding data use, and protection of sensitive data confidentiality, should be considered by librarians when preparing RDM policy. Also, it is essential to consider legal issues such as copyright, data protection against security breaches and confidentiality, and fairly patenting as a way of promoting flowless research data management practice. Consequently, the study

recommends further research on more issues which could be covered by research data management policy and legal concerns.

References

- Abraham, R., Schneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International journal of information management*, 49, 424-438.
- Andrikopoulou, A., Rowley, J., & Walton, G. (2021). Research Data Management (RDM) and the evolving identity of academic libraries and librarians: a literature review. *New Review of Academic Librarianship*, pp. 1-17.
- Birkbeck, G., Nagle, T., & Sammon, D. (2022). Challenges in research data management practices: a literature analysis. *Journal of Decision Systems*, 1-15.
- Chiarelli, A., Beagrie, N., Boon, L., Mallalieu, R., Johnson, R., May, A. W., & Wilson, R. (2022). To protect and to serve: developing a road map for research data management services. *Insights*, 35.
- Cox, A., & Verbaan, E. (2018). *Exploring research data management*. Facet publishing.
- Fitzsimmons, T., Jayasena, W., Holden, C. A., Dono, J., Hewett, P., Moore, J., & Sammour, T. (2022). Assessing the impact of the 2018 National Health and Medical Research Council polyp surveillance guidelines on compliance with surveillance intervals at two public hospitals. *ANZ Journal of Surgery*, 92(11), 2942-2948.
- Hess, S. F. (2019). Exploring research data management.
- Kahl, M., Chakraborty, S., Pompilio, M., & Costa, F. (2015). Constraints on the nature and evolution of the magma plumbing system of Mt. Etna volcano (1991-2008) from a combined thermodynamic and kinetic modelling of the compositional record of minerals. *Journal of Petrology*, 56(10), 2025-2068.
- Kinde, A. A., Addis, A. C., & Abebe, G. G. (2021). Research data management practice in higher education institutions in Ethiopia. *Public Services Quarterly*, 17(4), 213-230.
- Lefebvre, A., Bakhtiari, B., & Spruit, M. (2020). Exploring research data management planning challenges in practice. *It-Information Technology*, 62(1), 29-37.
- Machimbidza, T., Mutula, S., & Ndiweni, M. (2022). Exploring the feasibility of Research Data Management services in Zimbabwean academic libraries: the case of Zimbabwe School of Mines. *Library Philosophy and Practice*, 1-34.
- Morgan, A., Duffield, N., & Walkley Hall, L. (2017). Research data management support: sharing our experiences. *Journal of the Australian Library and Information Association*, 66(3), 299-305.
- Ng'eno, E., & Mutula, S. (2018). Research Data Management (RDM) in agricultural research institutes: A literature review. *Inkanyiso: Journal of Humanities and Social Sciences*, 10(1), 28-50.
- Ng'eno, E. J., & Mutula, S. M. (2022). Research data management in Kenya's agricultural research institutes. In *Handbook of Research on Academic Libraries as Partners in Data Science Ecosystems* (pp. 334-361). IGI Global.
- Nhendodzashe, N., & Pasipamire, N. (2017). Research data management services: Are academic libraries in Zimbabwe ready? The case of the University of Zimbabwe library.
- Shelly, M., & Jackson, M. (2018). Research data management compliance: Is there a bigger role for university libraries? *Journal of the Australian Library and Information Association*, 67(4), 394-410.
- Ssegawa, J. K., & Muzinda, M. (2021). Feasibility Assessment Framework (FAF): A Systematic and Objective Approach for Assessing the Viability of a Project. *Procedia Computer Science*, 181, 377-385.
- Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*, 36(2), 84-90.
- Wiley, C. (2018). Data sharing and engineering faculty: An analysis of selected publications. *Science & technology libraries*, 37(4), 409-419.
- Wong, G. K., & Chan, D. L. (2021). Designing library-based research data management services from the bottom up. In *Future Directions in Digital Information* (pp. 55-68). Chandos Publishing.
- Xiaofeng, M., & Xiang, C. (2013). Big data management: concepts, techniques, and challenges. *Journal of Computer Research and Development*, 50(1), 146.