

2023-10-31

The utilisation of open research data repositories for storing and sharing research data in higher learning institutions in Tanzania

Mosha, Neema

Emerald Publishing Limited

<https://doi.org/10.1108/LM-05-2023-0042>

Provided with love from The Nelson Mandela African Institution of Science and Technology

The utilisation of open research data repositories for storing and sharing research data in higher learning institutions in Tanzania

Neema Florence Mosha

*College of Graduate Studies, University of South Africa – Muckleneuk Campus,
Pretoria, South Africa, and*

Patrick Ngulube

*Department of Interdisciplinary Research and Postgraduate Studies,
University of South Africa, Pretoria, South Africa*

Abstract

Purpose – The study aims to investigate the utilisation of open research data repositories (RDRs) for storing and sharing research data in higher learning institutions (HLIs) in Tanzania.

Design/methodology/approach – A survey research design was employed to collect data from postgraduate students at the Nelson Mandela African Institution of Science and Technology (NM-AIST) in Arusha, Tanzania. The data were collected and analysed quantitatively and qualitatively. A census sampling technique was employed to select the sample size for this study. The quantitative data were analysed using the Statistical Package for the Social Sciences (SPSS), whilst the qualitative data were analysed thematically.

Findings – Less than half of the respondents were aware of and were using open RDRs, including Zenodo, DataVerse, Dryad, OMERO, GitHub and Mendeley data repositories. More than half of the respondents were not willing to share research data and cited a lack of ownership after storing their research data in most of the open RDRs and data security. HLIs need to conduct training on using trusted repositories and motivate postgraduate students to utilise open repositories (ORs). The challenges for underutilisation of open RDRs were a lack of policies governing the storage and sharing of research data and grant constraints.

Originality/value – Research data storage and sharing are of great interest to researchers in HLIs to inform them to implement open RDRs to support these researchers. Open RDRs increase visibility within HLIs and reduce research data loss, and research works will be cited and used publicly. This paper identifies the potential for additional studies focussed on this area.

Keywords Tanzania, Research data, Higher learning institutions, Research data sharing, Open research data repositories, Research data storage

Paper type Research paper

1. Introduction and theoretical background

The pressure of storing and sharing research data in open research data repositories (RDRs) has been placed on researchers by funding agencies, journal publishers, funding



agencies, higher learning institutions (HLIs), open science initiatives such as findable, accessible, interoperable and reusable (FAIR) and research data movements (Boyd, 2021; Hrynaskiewicz *et al.*, 2021; Uzwyshyn, 2016). Scientific discovery is accelerated by the storage and sharing of research data. It also improves the return on investment for researchers' institutions and communities (Donaldson and Koepke, 2022). To improve access to and reuse of such research data, it is necessary to store and exchange it utilising open platforms such as Open RDRs (Sicilia *et al.*, 2017 Bezuidenhout, 2019; Boyd, 2021; Donaldson and Koepke, 2022; Jeng and He, 2022; Mauthner and Parry, 2013; Tenopir *et al.*, 2015).

Open RDRs are amongst the online systems for storing and sharing research data (Besançon *et al.*, 2021; Chawinga and Zinn, 2019; Donaldson and Koepke, 2022; Jeng and He, 2022; Thoegersen and Borlund, 2021; Nielsen, 2011; Uzwyshyn, 2016). Holdren (2013, p. 5) defined RDRs as "large database infrastructures set up to manage, share, access, and archive researchers' datasets" when practicable and possible from a legal and ethical standpoint. Data centres, data archives and/or scientific databases are other names for open RDRs (Uzwyshyn, 2016). Open RDRs enhance the storage and sharing of a wide range of data types in a wide variety of formats (Wilkinson *et al.*, 2016) and play a vital role in the research life cycle and ensuring research data deposits remain FAIR (Boyd, 2021; Wilkinson *et al.*, 2016). Open RDRs enable researchers to store and share their research data using automatic persistent identifiers (e.g. persistent identifier (PID) and uniform resource locator (URL)), indexing and metadata to ensure access control possibilities (e.g. embargo and authentication) (Donaldson and Koepke, 2022; Thoegersen and Borlund, 2021; van Wyk Johann and van der Walt, 2020). Additionally, they support academic research cooperation, publishing and sharing, enabling HLIs to manage their study data and present them to larger research communities (Van Wyk Johann and Van der Walt, 2020).

The three main categories of open RDRs are:

- (1) Domain-specific or discipline repositories (DRs) which are used to store research data on the specific subject areas (Garijo *et al.*, 2022), for example, National Institutes of Health (NIH) data repositories for medical research data which are maintained by the Trans-NIH Biomedical Informatics Coordinating Committee (BMIC) (Gonzales *et al.*, 2022).
- (2) General-purpose, public open repositories (PR) or open repositories (ORs) which enable researchers to deposit and make their research data available regardless of disciplinary or institutional affiliation (Austin *et al.*, 2015). Good examples include Open Science Framework (OSF), Zenodo, Figshare, Dryad and Harvard Dataverse (Goben and Sandusky, 2020; Park, 2022; Scherer and Valen, 2019; Sicilia *et al.*, 2017; Van Wyk Johann and Van der Walt, 2020).
- (3) Institutional data repositories (IDRs) which are maintained and operated by a specific institution, these repositories are often university-based (Xafis and Labude, 2019), for instance, the University of Pretoria (UP) RDR (<https://researchdata.up.ac.za/>) which is used to facilitate data publishing, sharing and collaboration of academic research and allowing UP to manage their research data (Van Wyk Johann and Van der Walt, 2020).

For open RDRs to be utilised they must earn the trust of the research communities and demonstrate that they are reliable and capable of appropriately managing the data they hold (Lin *et al.*, 2020). That implies that HLIs should opt for trusted RDRs (Giaretta, 2007) that are indexed by the Registry of Research Data Repositories (re3data) (Boyd, 2021). Thus, HLIs need to ensure that research data collected by their institutional members such as students,

staff and researchers are stored and shared via open RDRs to enable access and reuse of such data to the public.

Due to their potential to provide very substantial and potent datasets, the majority of HLIs worldwide are starting to adopt and implement open RDRs (Thoegersen and Borlund, 2021; Van Wyk Johann and Van der Walt, 2020) through their libraries and Information and Communication Technologies (ICTs) departments. In this regard, there is a need for a close collaboration between libraries and ICTs departments for the implementation of open RDRs in HLIs. Despite the training and workshops guided by libraries and ICTs to ensure academic community members are well informed and are capable on using open RDRs for storing and sharing their research data (Gordon *et al.*, 2015; Van Wyk Johann and Van der Walt, 2020), other activities to ensure proper implementation of open RDRs such as metadata creation and software installation must be conducted in collaboration between libraries and ICTs departments within HLIs (Mosha and Ngulube, 2023; Van Wyk Johann and Van der Walt, 2020). Thus, this study used the Nelson Mandela African Institution of Science and Technology (NM-AIST) in Tanzania as a case study to explore how open RDRs might enhance storage and research data sharing in HLIs.

2. Significance of the study

The present study is significant because it addresses a pressing research issue of the need for researchers to store their research data in open RDRs with most of them available free of charge. That enables the preservation and curation of research data beyond the lifetime of a research project. The study also creates an opportunity for researchers and students in HLIs to gain new knowledge about storing and distributing research data for later use and produce new research results. In this scenario, reusing research data in open RDRs saves time and cost for researchers going to field for collecting data. Open RDRs facilitate research data publishing, sharing and collaboration of academic research, allowing HLIs to manage and in some cases showcase their research data within and outside their countries. In other words, this study raises awareness of open RDRs and their advantages to researchers and science.

3. Problem statement

Advances in technology have increased the types of digital data collected and analysed during research projects in many institutions, including HLIs (Thoegersen and Borlund, 2021). There is evidence that research data are saved on personal devices such as laptops, iPads, flash drives, hard drives and tablets, which makes it difficult to keep and share research data (Haixia *et al.*, 2017). The lack of open RDRs in many HLIs, the lack of policies and guidelines to improve the storage and sharing of research data in HLIs, the cost of hosting research data within open RDRs and a lack of knowledge and expertise about the storage and sharing of research data utilising open RDRs are some of the reasons why some HLIs do not properly save and share these data (Donaldson and Koepke, 2022; Hrynaszkiewicz *et al.*, 2021; Uzwyshyn, 2016; Xu *et al.*, 2022). However, little is known about how HLIs in Tanzania utilise open RDRs to store and share research data. Therefore, this study investigated the utilisation of open RDRs for storing and sharing enhanced research data in HLIs using the NM-AIST as a test case. The specific objectives of the study were to:

- (1) Determine the level of awareness and usage of open RDRs amongst postgraduate students.
- (2) Establish the extent to which postgraduate students were willing to share research data.

-
- (3) Assess the role of HILs in supporting research data storage and sharing.
 - (4) Outline the requirements for trusted open RDRs.
 - (5) Ascertain the inclusion of research data traceability information in the repository.
 - (6) Identify the benefits of storing and sharing research data in open RDRs.
 - (7) Find out the challenges that hinder the storage and sharing of research data in open RDRs.

4. Methods

4.1 Research design

This study employed a cross-sectional study design (Creswell and Creswell, 2018; Ponto, 2015). A pretested closed-ended questionnaire was used to collect research data from postgraduate students. Since it was COVID-19 period, the questionnaires were circulated online (using individual emails), with a rate of return of 58 (83) %.

4.2 Study area

The study was conducted at the NM-AIST, one of the public universities located in Arusha, Tanzania. The institution has four schools, namely the School of Computational and Communication Science Engineering (CoCSE), the School of Materials, Energy, Water and Environmental Sciences (MEWES), the School of Life Sciences and Bioengineering (LiSBE) and the School of Business and Humanities (BuSH). The School of BuSH was not included in this study because it was not offering any postgraduate degree during data collection.

4.3 Study population

The study investigated second-year postgraduate students at the NM-AIST who were in the final stages of research data collection, analysis and storage.

4.4 Sample size and sampling procedure

The study employed census sampling which provides an equal chance for all members of a given population to participate in the study. According to Fottrell and Byass (2008), census sampling refers to the quantitative research method, in which all the members of the population are enumerated. Furthermore, add that census sampling has higher participation of individuals and it has low chances of bias, it is also more applicable to a small number of the total population (Gilmore *et al.*, 2022). Thus, in this study the whole population that is second-year postgraduate students was included. The main reason for using census sampling was because the target population was small (70 postgraduate students). We also focussed on the second-year students since they were in the middle of their research activities, whilst first-year students were busy with their coursework.

4.5 Data collection

Data was collected using questionnaires that included both closed and open-ended question items. Responses from the open-ended questions (qualitative data) were used to supplement responses from the closed-ended questions (quantitative data). The main aim is to enable both quantitative and qualitative results to be compared and reported at the same time (Fetters *et al.*, 2013).

4.6 Data analysis

The Statistical Package for Social Sciences (SPSS), also known as IBM SPSS Statistics software package was used to analyse research data.

4.7 Ethical consideration

Ethical clearance was obtained from the University of South Africa (UNISA), whilst permission to collect data at the institution was obtained from the NM-AIST. On the other hand, participants were asked to sign an informed consent form. In addition, respondents were assured that the information collected will be treated confidentially and used purely for research work. Aliases or pseudonyms were used in data analysis to ensure the confidentiality and privacy of the study participants.

5. Results

5.1 Demographic information of respondents

A total of 58 (83%) postgraduate students responded to the questionnaire. Male respondents accounted for 38 (66%), whilst females accounted for 20 (34%). Many respondents (53%) were aged between 31 and 40 years of age. Table 1 presents the demographic information about respondents. There were no obvious demographic trends in the data.

5.2 Awareness and usage of open RDRs amongst researchers

More than half of the respondents 32 (55%) were not aware and were not using open RDRs for research data storage and sharing. Amongst them, a total of 20 (62.5%) out of 32 were storing their research data using their personal own devices such as laptop computers, iPad, flash disks, external hard drives and CDs/DVDs. Figure 1 provides personal research data storage and sharing facilities amongst respondents.

Less than half of the respondents 26 (45%) were aware of and using open RDRs for research data storage and sharing, whereas a total of 9 (35%) respondents were using Zenodo data repository to store and share their research data. Figure 2 presents open RDRs used by respondents to store their research data.

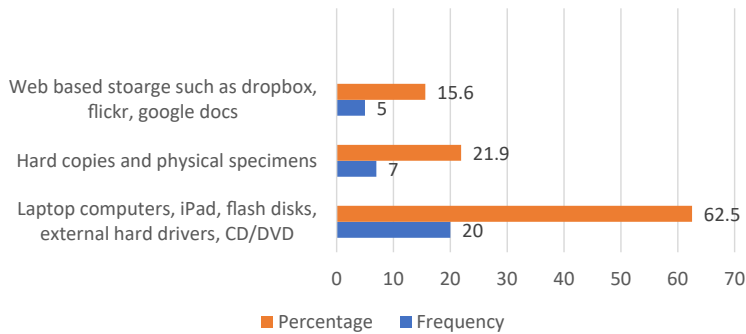
5.3 Willingness to share research data

A total of 10 (17%) out of 58 respondents were willing to share their research data. Figure 3 presents the willingness to share research data amongst respondents. On the other side, 48 (83%) respondents were not willing to share their research data.

Table 1.
Demographic
information of
respondents

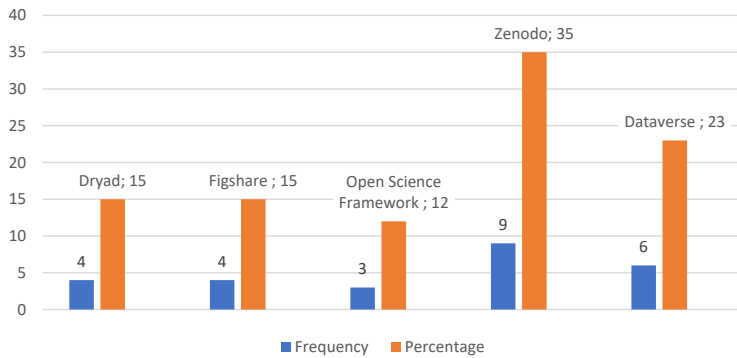
Item(s)	Categories	Frequency	Percentage
Gender	Male	38	66
	Female	20	34
Age (in years)	21–30	17	30
	31–40	31	53
	41–50	10	17
	51–60	0	0
School	CoCSE	21	36
	MEWES	13	22
	LiSBE	24	42

Source(s): Authors



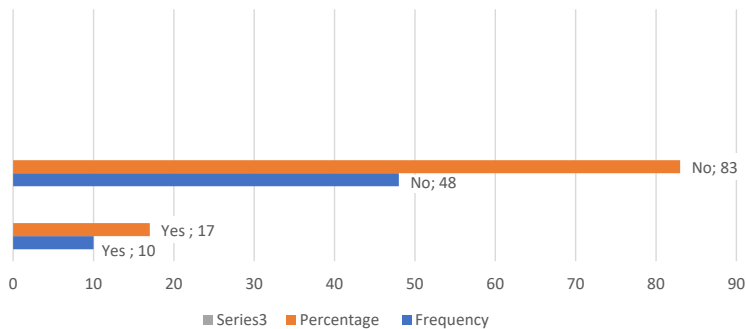
Source(s): Authors

Figure 1.
Personal research data
storage and sharing
facilities ($N = 32$)



Source(s): Authors

Figure 2.
Open RDRs for
research data
storage ($N = 26$)



Source(s): Authors

Figure 3.
Willingness to share
research data amongst
respondents ($N = 58$)

A total of 18 (37.5%) respondents said they did not share research data due to a lack of ownership after storing research data in most of the open RDRs, whereas 10 (20.8%) cited a lack of long-term preservation of research data as their reason for not sharing. Table 2 depicts reasons for not sharing research data amongst respondents.

5.4 Role of higher learning institutions in supporting data storage and sharing

Roles that need to be played by HLIs to ensure the storage and sharing of research data using open RDRs amongst their members were indicated whereas respondents 16 (27.6%) training on storing and sharing research data in open RDRs. [Table 3](#) provides different HLIs' roles to support research data storage and sharing using open RDRs.

5.5 Requirements for trusted open RDRs

Based on the requirements needed for trusted open RDRs, 16 (27.5%) respondents preferred open RDRs which allow the storage and sharing of different types and formats of research data. [Table 4](#) illustrates the requirements for trusted open RDRs expected by the respondents.

5.6 Research data traceability information

Research data stored and shared using open RDRs need to be traceable and should be accompanied by storage details. In this scenario, 18(31.1%) respondents indicated that research data storage details that provide information about "when" the research data were

Table 2.
Reasons for not
sharing research data

Reasons	Frequency	Percentage
Lack of long-term preservation of research data	10	20.8
Lack of ownership after storing research data in open RDRs	18	37.5
Lack of evidence offered by open data sharing	8	16.7
Research data being scooped	7	14.6
Lack of proper credit or attribution for sharing data	5	10.4
Total	58	100

Source(s): Authors

Table 3.
HLIs' roles to support
research data storage
and sharing

HLIs' roles	Frequency	Percentage
Develop open RDRs for storing and sharing research data	11	19.0
Provide and guide researchers on the available and reliable open RDRs	12	20.7
Provide training on storing and sharing research data using Open RDRs	16	27.6
Provide facilities to enhance storing and sharing of research data	9	15.5
Remove restrictions on research data produced within HLIs	10	17.2
Total	58	100

Source(s): Authors

Table 4.
The requirements for
trusted open RDRs

Requirements	Frequency	Percentage
Ensure online registration and enhance open access	12	20.7
Ensure long-term preservation of research data stored	7	12.1
Allow the storage of different types and formats of data	16	27.5
Ensure data traceability (FAIR Principle)	8	13.8
Establish usage policies and operational procedures	15	25.9
Total	58	100

Source(s): Authors

created, submitted and published should be easily traced in open RDRs. [Table 5](#) depicts research data storage information in open RDRs.

Open research
data
repositories

5.7 Benefits of storing and sharing research data in open RDRs

Respondents mentioned various benefits of storing and sharing research data. A total of 16 (27.6%) respondents mentioned that open RDRs increases research output and reduces research data waste. [Table 6](#) mentions various the benefits of storing and sharing research data using open RDRs.

573

5.8 Challenges that hinder the storage and sharing of research data in open RDRs

Challenges that hinder the storage and sharing of research data in open RDRs were presented to the respondents. A total of 16 (27.6%) respondents indicated a lack of research data storage and sharing policies and guidelines posed a lot of difficulties for researchers. [Table 7](#) presents challenges that hinder the storage and sharing of research data.

Research data storage details	Frequency	Percentage
Metadata standards and schemes	9	15.5
Information about “when” the data was created, submitted and stored	18	31.1
Research goals, type of research and funding sources	10	17.2
Methodology, sources, instruments and software used	8	13.8
Author(s), keywords, codes, tags and subject	13	22.4
<i>Total</i>	<i>58</i>	<i>100</i>

Source(s): Authors

Table 5.
Research data storage
details

Benefits	Frequency	Percentage
Increases citations, transparency and accountability	10	17.2
Enhances job, funding and collaboration opportunities	13	22.4
Saves time and money	14	24.2
Encourages standards and codification	5	8.6
Increases research output and reduces research data waste	16	27.6
<i>Total</i>	<i>58</i>	<i>100</i>

Source(s): Authors

Table 6.
Benefits of storing and
sharing research data
in open RDRs

Challenges	Frequency	Percentage
Lack of legal and confidentiality information	12	20.7
Misuse or misinterpretation of stored research data	11	19
Lack of research data storage and sharing policies and guidelines	16	27.5
Lack of privacy and security of research data	10	17.2
Lack of curation and long-term preservation of data	9	15.5
<i>Total</i>	<i>58</i>	<i>100</i>

Source(s): Authors

Table 7.
Challenges hinder the
storage and sharing of
research data using
open RDRs

6. Discussion

The discussion in this study is presented based on the specific objectives as follows:

6.1 Awareness and usage of open RDRs amongst researchers

In this study, more than half of the respondents were not aware of the open RDRs for storing and sharing their research data. Most of them were using their personal devices to store and share their research data. [Goben and Sandusky \(2020\)](#) warned that using personal devices has drawbacks because of previous reliance on decentralised servers, laptops, printouts, or hard drives stashed under a desk, which has resulted in ongoing and frequent research data loss. Few respondents were using open RDRs for their research data storage and sharing. Most of them were using Zenodo Repository to store their research data other than open RDRs such as Dryad, Figshare and Mendeley. Similar findings were reported by [Sicilia et al. \(2017\)](#), who found that 3,828 research data were kept in Zenodo at the time of data extraction. Despite storing research data, Zenodo allows the publication of any research output, including papers, posters and presentations ([Assante et al., 2016](#)).

6.2 Willingness to share research data

Many respondents expressed a lack of willingness to share their research data. Only a total of 10 (17%) out of 58 respondents were willing to share their research data. The present study revealed different reasons provided by respondents for not sharing their research data using open RDRs. Majority of them presented a lack of ownership after storing their data in most of the open RDRs, whereas others presented a lack of long-term preservation of research data. The absence of ownership after storing research data in the majority of open RDRs was also supported by other studies ([Austin et al., 2015](#); [Bangani and Moyo, 2019](#); [Bezuidenhout, 2019](#); [Michener, 2015](#); [Thoegersen and Borlund, 2021](#)). [Austin et al. \(2015\)](#) added that a lack of ownership rights prohibits researchers to deposit their data in repositories for open access. Other studies presented various reasons for not sharing research data. [Longo and Drazen \(2016\)](#) underlined inadequate attribution, [Science et al. \(2019\)](#) discovered a lack of incentives for researchers to share research data as well as an inadequate infrastructure for storage and sharing of research data. Inappropriate use of the shared data, costs associated with data storage and sharing, or the trialist's ability to have a fair opportunity to publish research using the data set first were amongst the hurdles to sharing clinical trial data discovered by [Rathi et al. \(2012\)](#) and [Ross and Krumholz \(2013\)](#). [Abele-Brehm et al. \(2019\)](#) found that researchers possess some fears regarding sharing research data, however, such fears are highest amongst early career researchers and lowest amongst professors. It will be interesting to find out from other studies if such reasons inhibit data sharing in other academic environments.

6.3 Role of higher learning institutions in supporting data storage and sharing

The present study findings revealed the need for HLIs to conduct training to equip knowledge and skills on storing and sharing open RDRs. The same observation was made by [Austin et al. \(2015\)](#) and [Goben and Sandusky \(2020\)](#) that training and advice on how to utilise open RDRs for storing and sharing research data, as well as help find reputable open RDRs, are significant roles that HLIs need to play in facilitating the adoption of open RDRs amongst their users. In this case, HLIs must set up trusted open RDRs per requirements such as the capacity to store and facilitate the sharing of various types and formats of research data. Various training programmes were indicated including metadata standards, research data reuse, free access and version control ([Austin et al., 2015](#)) and uploading and updating data files (i.e. versioning) and formats and usability of the stored research data ([Donaldson and Koepke, 2022](#)).

6.4 Requirements for trusted open RDRs

The findings indicated that respondents preferred open RDRs which allow the storage and sharing of different types and formats of research data. [Lin *et al.* \(2020\)](#) supported the findings presented by respondents that research data types and formats are amongst the criteria for a trusted repository since they can allow that data to be accessible. [Faundeen \(2017\)](#) also recommended that data and file formats that ensure long-term storage are amongst the requirements for trusted open RDRs. Other requirements presented by [Lin *et al.* \(2020\)](#) were metadata schema, data file formats, controlled vocabularies and ontologies.

6.5 Research data traceability information

Research data stored and shared using open RDRs need to be traceable and should be accompanied by storage details. For the research data to be traceable, they should provide information about “when” the research data was created, submitted and published. The findings from this study seem to suggest that the respondents did not pay enough attention to this issue. Traceability information as presented by [Assante *et al.* \(2016\)](#) including creation, submission and publication dates of the research data, as well as a unique or persistent identifier (such as a URL or PID) was not prevalent in the research data of the respondents. [Assante *et al.* \(2016\)](#) and [Austin *et al.* \(2015\)](#), concluded that research data needs to be accompanied by Digital Object Identifier (DOI) standards for persistent identification and traceability.

6.6 Benefits of storing and sharing research data in open RDRs

Respondents mentioned various benefits of storing and sharing research data whereas majority mentioned to increase research outputs and to reduce research data wastes. Studies by [Jeng and He \(2022\)](#), [Piwowar *et al.* \(2007\)](#) and [Ross and Krumholz \(2013\)](#) presented various benefits of open RDRs including to improve academic citation, visibility, scholarly influence and future chances. Other benefits include increasing the visibility of research data for citation and reuse by other researchers, reinforcing scientific inquiry by ensuring that enthusiasts and sceptics can test, validate and replicate research results, promoting new research and different ways of testing and analysing research data, saving financial and other resources that are wasted when similar data sets are created by different researchers and enabling new discoveries from old research data sets ([Engineering and Physical Sciences Research Council \(EPSRC\), \(2018\)](#)). Increased datafication of research procedures and infrastructure, encouragement for researchers to publish raw data, data sets and metadata and increased visibility of HLIs were further benefits that were indicated in line with what is reported in the literature ([Kidwell *et al.*, 2016](#); [Nosek *et al.*, 2015](#); [Roche *et al.*, 2014](#)).

6.7 Challenges that hinder the storage and sharing of research data in open RDRs

The findings revealed a lack of research data storage and sharing policies and guidelines posed a lot of difficulties for researchers. A lack of strategic planning for the long-term preservation of research data was also noted by [Jeng and He \(2022\)](#). Moreover, [Austin *et al.* \(2015\)](#), identified inadequate platform support for file-level metadata descriptions for research data or files and high expenses related to data storage, dissemination, curation, or preservation as some of the difficulties encountered in open RDRs. According to [Goben and Sandusky \(2020\)](#), research data loss is the most frequent threat, followed by improper data handling, loss of data privacy and security, problems with reproducibility and loss of trust from research participants and the public. The existence of policies and guidelines can play a great role in reducing some of the identified difficulties.

7. Areas for further studies

- (1) The use of structured questions limits participants to provide more information and ideas concerning the use of open RDRs for storing and sharing research data. Further studies should include unstructured questions and will engage more respondents from different universities.
- (2) Examine the factors associated with the application of open RDRs in HLIs in Tanzania and Sub-Saharan African (SSA).
- (3) Engage library members especially students, staff and researchers to assess not only the usage but also the implementation strategies for more open RDRs.
- (4) Examine the factors that affect the application of open RDRs for storing and sharing research data HLIs in Tanzania and Africa.

8. Conclusion and recommendations

Many respondents were unaware of open RDRs and their functionality. They used their own devices including laptops, iPads, flash disks and the like to store their research data. There is a need for HLIs to raise awareness about open RDRs amongst the members of their community, especially as they relate to storing and sharing research data. HLIs should ensure the installation of trusted open RDRs considering all the requirements for trusted open RDRs. Researchers must make sure their research data are accompanied by traceable information to be traced and utilised for subsequent research projects. To raise researchers' understanding of why and how they must use open RDRs in HLIs, it is necessary to communicate and promote the benefits of storing and sharing research data for the effective utilisation of open RDRs. The challenges that hinder storing and sharing of research data need to be minimised. Reducing the difficulties associated with sharing research data is good for science and the progress of humanity.

It is recommended that:

- (1) HLI should increase the awareness on storing and sharing research data using open RDRs.
- (2) HLIs should provide training to their researcher to increase the usability of open RDRs for storing and sharing research data.
- (3) HLIs should minimise the challenges presented to increase the usage of open RDRs for storing and sharing research data.

This study is a case study. It provides an overview of how researchers at a Tanzanian higher education institution use open RDRs. The results cannot be generalised to other HLIs, theories about the utilisation of open RDRs in HLIs can be developed if more case studies are conducted on the subject.

References

- Abele-Brehm, A.E., Gollwitzer, M., Steinberg, U. and Schönbrodt, F.D. (2019), "Attitudes towards open science and public data sharing: a survey among members of the German psychological society", *Social Psychology*, Vol. 50 No. 4, pp. 252-260, doi: [10.1027/1864-9335/a000384](https://doi.org/10.1027/1864-9335/a000384).
- Assante, M., Candela, L., Castelli, D. and Tani, A. (2016), "Are scientific data repositories coping with research data publishing?", *Data Science Journal*, Vol. 15 No. 6, pp. 1-24, doi: [10.5334/dsj-2016-006](https://doi.org/10.5334/dsj-2016-006).

- Austin, C., Brown, S., Fong, N., Humphrey, C., Leahey, A. and Webster, P. (2015), "Research data repositories: review of current features, gap analysis, and recommendations for minimum requirements", *IASSIST Quarterly*, Vol. 39 No. 4, 24, doi: [10.29173/iq904](https://doi.org/10.29173/iq904).
- Bangani, S. and Moyo, M. (2019), "Data sharing practices among researchers in South African Universities", *Data Science Journal*, Vol. 18 No. 28, pp. 1-14, doi: [10.5334/dsj-2019-028](https://doi.org/10.5334/dsj-2019-028).
- Besançon, L., Peiffer-Smadja, N., Segalas, C., Jiang, H., Masuzzo, P., Smout, C., Billy, E., Deforet, M. and Leyrat, C. (2021), "Open science saves lives: lessons from the COVID-19 pandemic", *BMC Medical Research Methodology*, Vol. 21 No. 117, pp. 1-12, doi: [10.1186/s12874-021-01304-y](https://doi.org/10.1186/s12874-021-01304-y).
- Bezuidenhout, L. (2019), "To share or not to share: incentivizing data sharing in life science communities", *Developing World Bioethics*, Vol. 19 No. 1, pp. 18-24.
- Boyd, C. (2021), "Understanding research data repositories as infrastructures", *Paper Presented at the 84th Annual Meeting of the Association for Information Science and Technology (ASIS & T) from 29th October to 3rd November 2021*, Salt Lake City, UT, pp. 25-35, doi: [10.7910/DVN/OWISNH](https://doi.org/10.7910/DVN/OWISNH).
- Chawinga, W.D. and Zinn, S. (2019), "Global perspectives of research data sharing: a systematic literature review", *Library and Information Science Research*, Vol. 41 No. 2, pp. 109-122, doi: [10.1016/j.lisr.2019.04.004](https://doi.org/10.1016/j.lisr.2019.04.004).
- Creswell, J.W. and Creswell, J.D. (2018), *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, Sage, Los Angeles.
- Donaldson, D.R. and Koepke, J.W.A. (2022), "A focus groups study on data sharing and research data management", *Scientific Data*, Vol. 9 No. 345, pp. 1-7, doi: [10.1038/s41597-022-01428-w](https://doi.org/10.1038/s41597-022-01428-w).
- Engineering and Physical Sciences Research Council (EPSRC) (2018), "Scope and benefits", available at: <https://epsrc.ukri.org/about/standards/researchdata/scope/>
- Faunden, J. (2017), "Developing criteria to establish trusted digital repositories", *Data Science Journal*, Vol. 16 No. 22, pp. 1-13, doi: [10.5334/dsj-2017-022](https://doi.org/10.5334/dsj-2017-022).
- Fetters, M.D., Curry, L.A. and Creswell, J.W. (2013), "Achieving integration in mixed methods designs: principles and practices", *Health Services Research*, Vol. 48 No. 6, pp. 2134-2156, doi: [10.1111/1475-6773.12117](https://doi.org/10.1111/1475-6773.12117).
- Fottrell, E. and Byass, P. (2008), "Population survey sampling methods in a rural African setting: measuring mortality", *Population Health Metrics*, Vol. 6 No. 2, doi: [10.1186/1478-7954-6-2](https://doi.org/10.1186/1478-7954-6-2).
- Garijo, D., Ménager, H., Hwang, L., Trisovic, A., Hucka, M., Morrell, T. and Allen, A. (2022), "Nine best practices for research software registries and repositories", *Peer Journal Computer Science*, Vol. 8 No. 8, doi: [10.7717/peerj-cs.1023](https://doi.org/10.7717/peerj-cs.1023).
- Giaretta, D. (2007), "Trusted data repositories", *Paper Presented at the 17th Annual Meeting of the Astronomical Data Analysis, Software and Systems (ADASS) from 24th to 28th 2007*, London.
- Gilmore, J.K., Bonciani, M. and Vainieri, M. (2022), "A comparison of census and cohort sampling models for the longitudinal collection of user-reported data in the maternity care pathway: mixed methods study", *JMIR Medical Information*, Vol. 10 No. 3, e25477, doi: [10.2196/25477](https://doi.org/10.2196/25477).
- Goben, A. and Sandusky, R.J. (2020), "Open data repositories: current risks and opportunities", *Scholarly Communication*, Vol. 81 No. 2, pp. 1-12.
- Gonzales, S., Carson, M.B. and Holmes, K. (2022), "Ten simple rules for maximizing the recommendations of the NIH data management and sharing plan", *PLoS Computer Biology*, Vol. 18 No. 8, e1010397, doi: [10.1371/journal.pcbi.1010397](https://doi.org/10.1371/journal.pcbi.1010397).
- Gordon, A.S., Millman, D.S., Steiger, L., Adolph, K.E. and Gilmore, R.O. (2015), "Researcher-library collaborations: data repositories as a service for researchers", *J Library School Communication*, Vol. 3 No. 2, pp. 1-15, doi: [10.7710/2162-3309.1238](https://doi.org/10.7710/2162-3309.1238).
- Haixia, J., Xingyu, Z. and Wei, T. (2017), "Research and implementation of mobile storage devices monitor and control system", *Procedia Computer Science*, Vol. 107, pp. 710-714.

- Holdren, J.P. (2013), *Increasing Access to the Results of Federally Funded Scientific Research*, Office of Science and Technology Policy, pp. 1-20, available at: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf
- Hrynaskiewicz, I., Harney, J. and Cadwallader, L. (2021), "A survey of researchers' needs and priorities for data sharing", *Data Science Journal*, Vol. 20 No. 31, 1, doi: [10.5334/dsj-2021-031](https://doi.org/10.5334/dsj-2021-031).
- Jeng, W. and He, D. (2022), "Surveying research data-sharing practices in US social sciences: a knowledge infrastructure-inspired conceptual framework", *Online Information Review*, Vol. 46 No. 7, pp. 1275-1292, doi: [10.1108/OIR-03-2020-0079](https://doi.org/10.1108/OIR-03-2020-0079).
- Kidwell, M.C., Lazarevic, L.B., Baranski, E., Hardwicke, T.E., Piechowski, S., Falkenberg, L.S. and Nosek, B.A. (2016), "Badges to acknowledge open practices: a simple, low-cost, effective method for increasing", *PLoS Biology*, Vol. 14 No. 5, e1002456, doi: [10.1371/journal.pbio.1002456](https://doi.org/10.1371/journal.pbio.1002456).
- Lin, D., Crabtree, J., Dillo, I., Downs, R.R., Edmunds, R., Giaretta, D., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M.E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D.V., Stockhause, M. and Westbrook, J. (2020), "The TRUST Principles for digital repositories", *Scientific Data*, Vol. 7 No. 144, pp. 1-5, doi: [10.1038/s41597-020-0486-7](https://doi.org/10.1038/s41597-020-0486-7).
- Longo, D.L. and Drazen, J.M. (2016), "Data sharing", *New England Journal of Medicine*, Vol. 374 No. 3, pp. 276-277, doi: [10.1056/NEJMe1516564](https://doi.org/10.1056/NEJMe1516564).
- Mauthner, N.S. and Parry, O. (2013), "Open access digital data sharing: principles, policies and practices", *Social Epistemology*, Vol. 27 No. 1, pp. 47-67, doi: [10.1080/02691728.2012.760663](https://doi.org/10.1080/02691728.2012.760663).
- Michener, W.K. (2015), "Ecological data sharing", *Ecological Informatics*, Vol. 29 No. 1, pp. 33-44, doi: [10.1016/j.ecoinf.2015.06.010](https://doi.org/10.1016/j.ecoinf.2015.06.010).
- Mosha, N.F. and Ngulube, P. (2023), "Teaching research data management courses in higher learning institutions in Tanzania", *Library Management*, Vol. 44 Nos 1-2, pp. 166-179, doi: [10.1108/LM-04-2022-0033](https://doi.org/10.1108/LM-04-2022-0033).
- Nielsen, M. (2011), "Definitions of open science? The open-science archives", available at: <https://lists.okfn.org/pipermail/open-science/2011-July/005607.html>
- Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S.D., Breckler, S.J. and Yarkoni, T. (2015), "Scientific standards: promoting an open research culture", *Science*, Vol. 348 No. 6242, pp. 1422-1425, doi: [10.1126/science.aab2374](https://doi.org/10.1126/science.aab2374).
- Park, H.J. (2022), "How to share data through harvard dataverse, a repository site: a case of the world journal of men's health", *Science Education*, Vol. 9 No. 1, pp. 85-90, doi: [10.6087/kcse.270](https://doi.org/10.6087/kcse.270).
- Piwovar, H.A., Day, R.S. and Fridsma, D.B. (2007), "Sharing detailed data is associated with increased citation rate", *PLoS One*, Vol. 2 No. 3, e308.
- Ponto, J. (2015), "Understanding and evaluating survey research", *Journal of the Advanced Practitioner in Oncology*, Vol. 6 No. 2, pp. 168-171.
- Rathi, V., Dzara, K., Gross, C.P., Hrynaskiewicz, I., Joffe, S., Krumholz, H.M. and Strait, K.M. (2012), "Sharing of clinical trial data among trialists: a cross sectional survey", *British Medical Journal*, Vol. 345, e7570.
- Roche, D.G., Lanfear, R., Binning, S.A., Haff, T.M., Schwanz, L.E., Cain, K.E. and Kruuk, L.E.B. (2014), "Trouble-shooting public data archiving: suggestions to increase participation", *PLoS Biology*, Vol. 12 No. 1, e1001779, doi: [10.1371/journal.pbio.1001779](https://doi.org/10.1371/journal.pbio.1001779).
- Ross, S.J. and Krumholz, H.M. (2013), "Ushering in a new era of open science through data sharing", *JAMA*, Vol. 309 No. 13, pp. 1355-1356.
- Scherer, D. and Valen, D. (2019), "Balancing multiple roles of repositories: developing a comprehensive repository at Carnegie Mellon University", *Publication*, Vol. 7 No. 30, doi: [10.3390/publications7020030](https://doi.org/10.3390/publications7020030).
- Science, D., Fane, B., Ayris, P., Hahnel, M., Hrynaskiewicz, I. and Baynes, G. (2019), "The state of open data report 2019", *Digital Science and Figshare (Open Repository)*. doi: [10.6084/m9.figshare.9980783.v1](https://doi.org/10.6084/m9.figshare.9980783.v1).

- Sicilia, M., Garcia-Barriocanal, E. and Sanchez-Alonso, S. (2017), "Community curation in open dataset repositories: insights from Zenodo in 13th international conference on current research information systems (CRIS2016)", *Procedia Computer Science*, Vol. 106, pp. 54-60.
- Tenopir, C., Dalton, E.D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D. and Dorsett, K. (2015), "Changes in data sharing and data reuse practices and perceptions among scientists worldwide", *PLoS One*, Vol. 10 No. 8, e0134826, doi: [10.1371/journal.pone.0134826](https://doi.org/10.1371/journal.pone.0134826).
- Thoegersen, J. and Borlund, P. (2021), "Researcher attitudes toward data sharing in public data repositories: a meta-evaluation of studies on researcher data sharing", *Journal of Documentation*, Vol. 78 No. 7, pp. 1-17, doi: [10.1108/JD-01-2021-0015](https://doi.org/10.1108/JD-01-2021-0015).
- Uzwysyn, R. (2016), "Research data repositories: the what, when, why, and how", *Computers in Libraries*, Vol. 36 No. 3, pp. 18-21.
- Van Wyk, J. and Van der Walt, I. (2020), *Criteria and Evaluation of Research Data Repository Platforms @ the University of Pretoria*, University of Pretoria, Pretoria, RSA.
- Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R. and Mons, B. (2016), "The FAIR guiding principles for scientific data management and stewardship", *Scientific Data*, Vol. 3, 160018.
- Xafis, V. and Labude, M.K. (2019), "Openness in big data and data repositories: the application of an ethics framework for big data in health and research", *Asian Bioethics Review*, Vol. 11, pp. 255-273, doi: [10.1007/s41649-019-00097-z](https://doi.org/10.1007/s41649-019-00097-z).
- Xu, Z., Zhou, X., Kogut, A. and Clough, M. (2022), "Effect of online research data management instruction on social science graduate students' RDM skills", *Library and Information Science Research*, Vol. 44 No. 4, pp. 101-190.

Further reading

- Corti, L., Van den Eynden, V., Bishop, L. and Woollard, M. (2014), *Managing and Sharing Research Data: A Guide to Good Practice*, Sage, London.
- Enwald, H., Grigas, V., Rudzionienė, J. and Kortelainen, T. (2022), "Data sharing practices in open access model: a study of the willingness to share data in different disciplines", *Information Research*, Vol. 27 No. 2, pp. 1-15, doi: [10.47989/irpaper932](https://doi.org/10.47989/irpaper932).
- Hansson, K. and Dahlgren, A. (2021), "Open research data repositories: practices, norms, and metadata for sharing images", *Journal of the Association for Information Science and Technology*, Vol. 73 No. 6, pp. 1-14, doi: [10.1002/asi.24571](https://doi.org/10.1002/asi.24571).
- Kurata, K., Matsubayashi, M. and Mine, S. (2017), "Identifying the complex position of research data and data sharing among researchers in natural science", *SAGE Open*, pp. 1-13, doi: [10.1177/2158244017717301](https://doi.org/10.1177/2158244017717301).
- Lee-Hwa, T., Abrizah, A. and Noorhidawati, A. (2013), "Availability and visibility of open access digital repositories in ASEAN countries", *Information Development*, Vol. 29 No. 3, pp. 274-285, doi: [10.1177/0266666912466754](https://doi.org/10.1177/0266666912466754).
- Ohmann, C., Moher, D., Siebert, M., Motschall, E. and Naudet, F. (2021), "Status, use and impact of sharing individual participant data from clinical data: a scoping review", *British Medical Journal Open*, Vol. 11 No. 8, pp. 1-20, doi: [10.1136/bmjopen-2021-049228](https://doi.org/10.1136/bmjopen-2021-049228).
- Popkin, G. (2019), "Data sharing and how it can benefit your scientific career: open science can lead to greater collaboration, increased confidence in findings and goodwill between researchers", *Nature*, Vol. 569, pp. 445-447, doi: [10.1038/d41586-019-01506-x](https://doi.org/10.1038/d41586-019-01506-x).

About the authors

Neema Florence Mosha is Postdoctoral Researcher at the University of South Africa (UNISA). She is also Director of Library Services at the Nelson Mandela African Institution of Science and Technology (NM-

AIST) and member of The International Federation of Library Associations and Institutions (IFLA)-HQ, Science and Technology Libraries Section Standing Committee 2023–2027. She holds a PhD in Information Studies from the UNISA, a Master of Information Technology (M.IT) from the University of Pretoria (UP), South Africa and a Master of Arts (Information Studies) from the University of Dar-es-Salaam (UDSM), Dar-es-Salaam, Tanzania. Her research focusses on the development of Research Data Management (RDM) services in higher learning institutions (HLIs). She is interested in research data management, digital and open repositories, artificial intelligence, knowledge management, open science, digital scholarships, and innovation in libraries, eResearch and Web 2.0 tools application in academic libraries. She is also involved in a project to train universities' staff and researchers in the fundamentals and essentials of open science and RDM. Neema Florence Mosha is the corresponding author and can be contacted at: moshanf@unisa.ac.za

Patrick Ngulube is Professor in the Department of Interdisciplinary Research and Postgraduate Studies at the UNISA. He is Visiting Professor at the National University of Science and Technology, Zimbabwe. He holds a PhD in information studies from the University of Natal in South Africa. He has received research awards from the University of South Africa and grants from the National Research Foundation. His research interests include research design and methodology, indigenous knowledge systems, knowledge management, records management, application of and the preservation of access to information. He has published over 50 articles in journals such as *African Journal of Library, Archives and Information Science*, *Government Information Quarterly and Library and Information Science Research*. In addition, he has edited 2 books and contributed to 10 book chapters. He has served in editorial boards of the *African Journal of Library, Archives and Information Science* and the *Journal of the Eastern and Southern Africa Regional Branch of the International Council on Archives*.