Future models and architecture of data repositories in African universities

Josiline Chigwada[1] and Elisha Chiware[2]

Abstract

Research data repositories as part of research infrastructures are being developed and are important tools and components that help to store, preserve, and allow for the re-use of data. As the technologies, networks, and systems that the data repositories are built upon are advancing, this study explores the future models and architectures that African universities can follow to have reliable and sustainable systems for the preservation of research data. A scoping review was done to focus on the future shape of data repositories based on past experiences of the last 10 years of research institutions in establishing data repositories. This study was done to gauge the communities' responses to the architecture of existing platforms to prepare other institutions planning to establish digital research data repositories. Articles were retrieved from Scopus, Web of Science, and Dimensions databases using relevant keywords. The content analysis approach was used to establish the requirements for establishing digital research data repositories to develop a framework that can be utilised by other research institutions to develop their repositories. The framework would be handy in providing a roadmap for research institutions that want to establish research data services in Africa enhancing the future of research infrastructure in African universities.

Keywords

Research Data Management, Research Data Services, Research Data Repositories, Data Repository Models, Data Repository Architecture.

Introduction

The pace of research data repositories' development in African universities has been slow compared to other continents, especially those in the global north (Patterton *et al*., 2018). Various factors can contribute to this slow progress including lack of financial resources; inadequate research infrastructures, lack of open science and research data management policies and frameworks, unstable electricity grids, and poor internet connectivity as well as limited skills (Chiware and Mathe, 2015; Chiware and Becker, 2018). In the last two decades, African universities and other knowledge production centers have developed and implemented digital repositories to showcase their research outputs including collections of Electronic Theses and Dissertations (ETDs), research articles, conference proceedings, technical reports, and book chapters. However, most of these repositories, built on platforms like Dspace, cannot host research datasets.

Over the last decade, several studies have emerged on how African institutions can develop research data management services. The majority of the writings are based on reviewing existing research data management practices and thereafter proposing frameworks on how new services can be developed and implemented. Several authors however recognise the existing challenges and have often questioned these common narratives that seem to ignore the reality of resource constraints in African research institutions. For instance, Abebe et al. (2021) argue that these narratives often overlook power imbalances and there is a need for solutions that are grounded in the African context.

Data Repositories Infrastructure in Africa

The Registry of Research Data Repositories (Re3data.org) currently lists disciplinary-specific and general research data repositories in several African countries. The majority of the listed repositories are based in South African institutions, followed by six in Kenya, three in Burkina Faso, two in Ghana and Benin, and single repositories in Cameroon, Egypt, Ethiopia, Ivory Coast, Malawi, Namibia, Niger, Senegal, Sudan and Tunisia. There are no other recorded research data repositories in any of the other African countries. The types of data repositories found in these African countries are at two levels: disciplinary (subject) specific and general repositories which are mostly found in South African university libraries. Disciplinary-specific data repositories include DataFirst, a research data service dedicated to providing open-access data from South Africa and other African countries. It also promotes high-quality research through the provision of essential Open Research Data Infrastructure for discovering and accessing data and skills development. Another existing project is the H3AbioNet (https://www.h3abionet.org/) which was established to develop bioinformatics capacity in Africa and specifically to enable genomics data analysis by H3Africa researchers across the continent. H3ABioNet is developing human capacity through training and support for data analysis, and facilitating access to informatics infrastructure by developing or providing access to pipelines and tools for human, microbiome, and pathogen genomic data analysis. The other discipline-specific continental data repositories include; Africa Rice Dataverse (https://www.re3data.org/repository/r3d100011251), AfDB Statistical Data portal (https://www.ruforum.org/directory/afdb-statistical-data-portal), ROCEEH Out of Africa Database (https://www.re3data.org/repository/r3d100013419), Square Kilometre Array (SKA) Telescope (https://www.sarao.ac.za/about/the-project/), Africa Health Research Institute Data repository (https://data.ahri.org/index.php/home), and the West African Vegetation (http://westafricanvegetation.senckenberg.de/menu/home.aspx). These repositories seek to enhance accessibility and research data utilisation across disciplines.

There are eighteen South African institutions, mainly university libraries and research councils' facilities that currently host data repositories of a general nature. The majority of these platforms run on proprietary platforms, especially on Figshare which was acquired through a national consortium arrangement to enable more uptake among interested institutions. As these data repositories grow there are increasing calls for their integration with existing platforms that host other research outputs (especially those on Dspace) (Mehnert *et al.,* 2019). There are also calls for institutions to go through the processes of certification to ensure that the repositories are internationally recognised as holding trusted data deposits. Trust in repositories will ensure maximisation of research outputs as well as facilitate collaboration and sharing of existing research outputs (Mehnert *et al.,* 2019). Trusted repositories usually have demonstrated high levels of trustworthiness, through certifications like the Core Trust Seal and membership of the International Science Council's World Data System (WDS) (Lin et al., 2020). They will have implemented robust policies, procedures, and technology infrastructures to ensure data quality, security, and preservation. Another important aspect relates to the culture of data sharing, which is the widespread adoption of open data practices, where researchers and organizations share data freely and willingly. The culture of sharing also encourages collaboration, reuse, and building upon existing research as well as fostering community norms and values around data sharing.

As research data management continues to grow and with support from funders, journal publishers, and the need to respond to international and national calls for coordinated data sharing mechanisms among and beyond research teams and enable more transparency as well as provide impetus to more accelerated scientific discoveries in developing countries and in Africa in particular, it is important to find models and solutions for data infrastructure architectures that can be adopted at minimal cost. The goal of this study was to explore how new models can be developed to encourage the uptake of data repositories in African institutions through minimal costs, as well as, ensuring sustainability.

Scope

The pace of research data repositories development on the African continent has been limited due to several challenges, including, limited funding, lack of policy framework, skills shortage and limited development within research infrastructures. Some of the well-resourced African countries have started using proprietary platforms to manage research data. Some of the existing institutional repositories like Dspace have no or very limited data management functionalities, thereby limiting the ability of most African universities to use open-source platforms to store and manage research data. For African universities to fully participate in the global open science agenda their scholarly outputs, including data, must be properly managed through data repositories that can be easily accessed. The aim of this paper was therefore to review past experiences and frame future models and define the architecture of data repositories that are more suitable for African university universities.

Objectives

The objectives of this study were:

1. To identify the requirements for establishing advanced data repositories.
2. To define successes and challenges in establishing and managing data repositories.
3. To develop a framework to be utilised when developing research data repository infrastructures.

Methodology

A scoping review was selected due to the complex nature of the topic and the wide range of information sources that might be available for the study since the issues of establishing research data repositories are topical. The review was guided by Levac's (Levac *et al*., 2010) scoping review methodology, which is an improvement of Arksey and O'Malley's (2005). The 5 stage methodological framework guided this study through identifying the research question, searching for relevant studies, selecting studies, charting the data, and collating, summarising, and reporting the results. The sixth, optional, stage of consulting with stakeholders to inform or validate study findings will be done as a way of developing the research paper.

Stage 1: involved the development of a research question. The following questions guided this review: 1) What are the requirements for establishing advanced research data repositories?; 2) What are the successes and challenges faced in establishing and managing research data repositories?

Stage 2: involved identifying relevant studies. Peer-reviewed articles, book chapters, and conference proceedings were retrieved from Scopus, Web of Science, and Dimensions. The following search terms were used: "Establish research data repository", "Research data repository requirements", "Research data repository AND Academic library", "Research data repository AND research institutions", and "Research data librarian experiences".

Stage 3: involved article selection using the inclusion and exclusion criteria. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) flowchart of article identification, screening, and extraction was used as shown in Figure 1. The selection was initially based on the titles, keywords, and abstracts.

Stage 4: involved data charting and extraction, where the selected articles were subjected to further screening and were retrieved from their databases, and each full-text article was examined. The data was documented on an Excel spreadsheet, and two reviewers reviewed the full-text articles to come up with relevant articles for the study. Eligible studies met the following criteria: 1) Published between

2013 and 2023; 2) Written in English; and 3) included a discussion of the establishment of research data repositories in research institutions.

Stage 5: involved collating, summarising, and reporting the results. Following a thorough reading of the articles, the authors used the research questions as a guide to identifying themes. All authors drafted and approved the report.

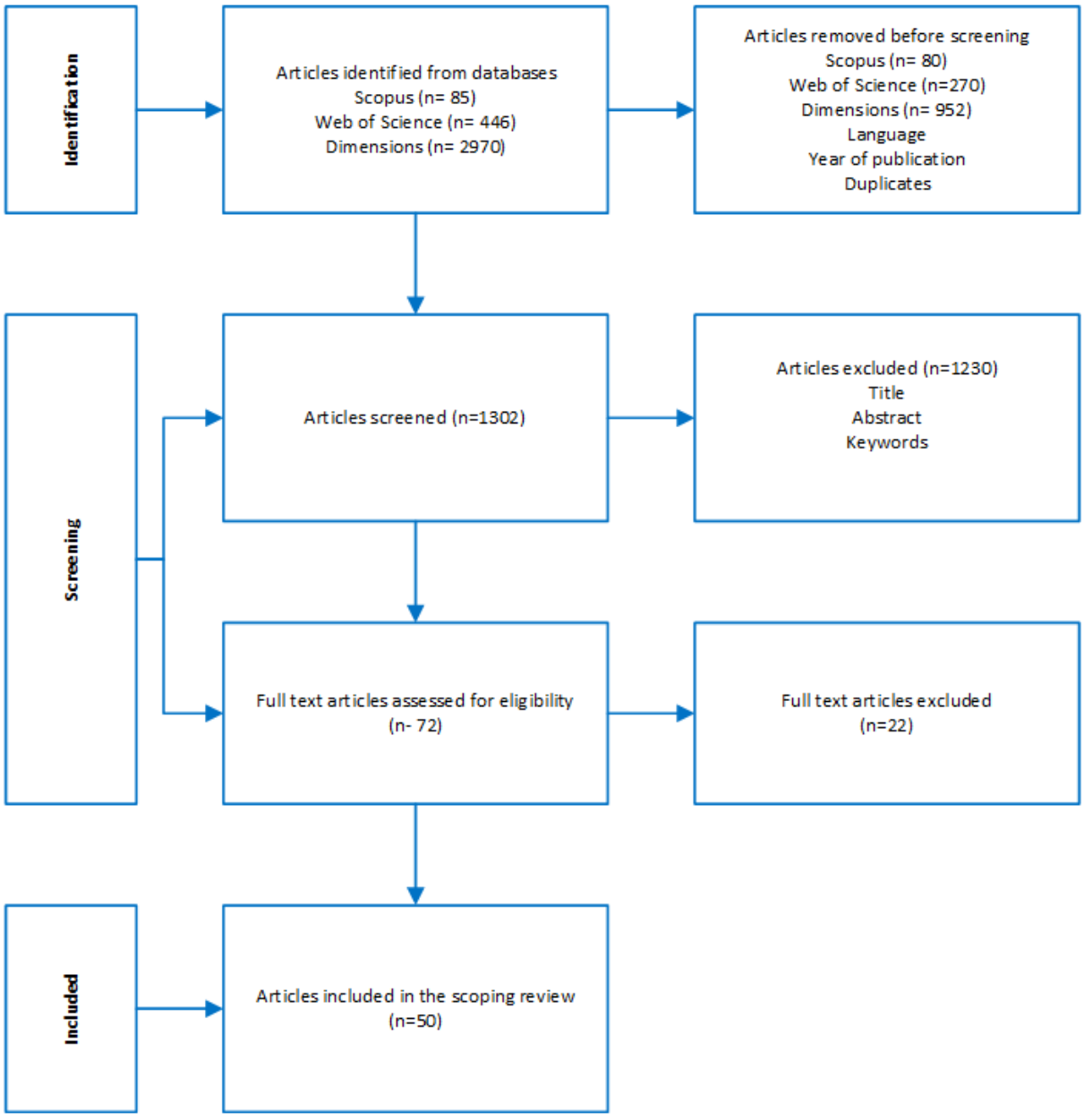

Figure 1: PRISMA flowchart

Findings and discussion

During the study, 3,501 articles were retrieved, and 50 articles were considered for this study after screening. Details for the screening, exclusion, and inclusion process can be found in Figure 1.

Requirements for establishing research data repositories

Repositories can be categorised into three groups based on the scope of content they collect and manage. The three groups are: domain repositories, discipline repositories, and institutional repositories and these differ according to the scope of content they collect and manage (Lee and Stvilia, 2017). This study was focused on institutional research data repositories. Patel (2016) developed a three-tier conceptual framework that is aimed at providing guidelines to address research data management issues at an institutional level. The first tier deals with data management, which involves developing institutional policy for data sharing, changing the mindset of researchers, data collection from researchers, copyright and data licensing, cross refer data to methodologies, data classification, data anonymization, data description and identification, data organisation, and an interoperability framework for data. The second tier is about data storage and hosting and covers the selection of file formats, data generated by private-public collaborations, data hosting services, independent data contributions, liability for hosted data, data security, data hosting software, and data backup. The third tier is about data usage and looks at access to data, copyright, data licensing, and rights in derivative works. A useful working guide for higher education institutions planning to start research data management services was developed by Jones et al. (2013). Nie et al. (2021) indicated that the implementation of research data management services included project kickoff, needs assessment, partnerships establishment, software investigation and selection, software customisation, and data curation services and training.

Cox *et al.* (2017) and Cox *et al*. (2019) developed a research data management landscape maturity model with four levels spanning from none, basic, developing, and extensive. Level zero deals with audits and surveys to solicit information concerning service and support, level one is the compliance stage with research data management governance boards and research data management policy, and level two deals with capacity-building and reengineering looking at skills, roles, and structures. The activities at levels one and two overlap and they deal with research data management training, data literacy, and advisory services (awareness of data archives, publication, citation storage, data management planning tools, and rights or intellectual property). Level three deals with stewardship where there are cultural acceptance and embedded practices looking at data repositories, technical support (selection, catalogue, curation, preservation, metadata), data analysis or visualisation, and research data management shared services. The model was revised and the new model retained the concept of four levels where level one was changed to compliance, level two stewardship, and level three transformation. The major change was in skills where there is the transition of existing skills on level one, reskilling of existing staff on level two, and new skills acquisition on level three (Cox *et al*., 2019). Chigwada *et al*. (2019) proposed a framework for establishing research data management services in Zimbabwe which consists of strategies, policies, guidelines, processes, technologies, and services.

Mushi *et al*. (2020) developed a planned implementation strategy that can be used by a university to establish research data management services. It includes four phases which include strategy, policy, procedures and infrastructure (phase 1), awareness creation, skills development and repository content development (phase 2), management of active data (phase 3), and data selection and preservation (phase 4) as shown in Figure 2. Knight (2015) also stated that it is important to determine the research data management requirements within the institutions so that the service would support the evolving needs of researchers. Issues such as funding, institutional data management infrastructure, research data management policies and procedures, a research data management website, and expertise should be considered when establishing research data management services at an institution. Dora and Kumar (2015) pointed out the factors that should be considered in designing and developing research data management services, they include understanding the needs of the various stakeholders, adopting standard recommendations, choosing the software (developing

one or adopting an existing commercial or open source software), reviewing the IT infrastructure and then developing institutional guidelines. They suggested that institutions should choose from Databank, CKAN, Dataverse, Figshare, Dryad, or Harvard Datacerse Network (Nie *et al*., 2021; Dora and Kumar, 2015).

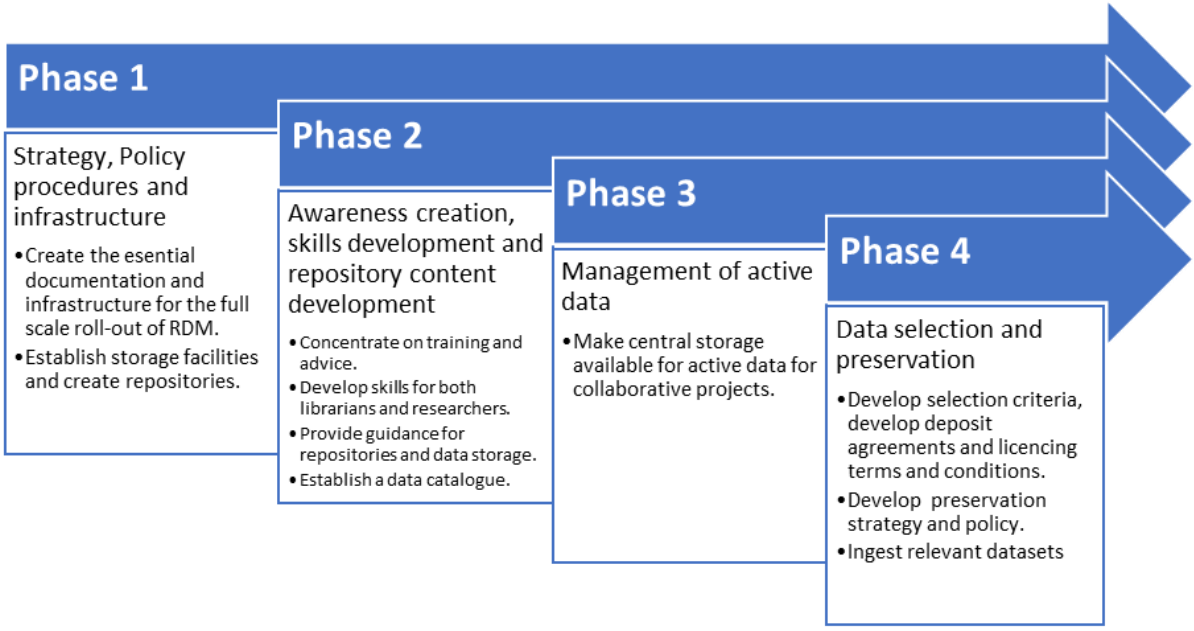

Figure 2: RDM implementation phases (Source: adapted from Mushi et al., 2020).

Experiences of universities in establishing research data repositories

The articles reviewed demonstrated the variety of experiences at universities from around the world. Knight (2015) noted that they identified various stakeholders that affect research data management and strengthened the institution's policy framework to address the needs of researchers. As a result, they came up with best practices that can be followed when offering research data management services. These include creating, managing, and sharing research data by contractual, legislative, regulatory, ethical, and other relevant requirements; creating a data management plan for all research projects that capture data; and registering all research data created, no matter where it is hosted. It was noted that some researchers were not willing to share their research data, although they wanted to use research data produced by others (Bangani and Moyo, 2019).

Chiware and Becker (2018) found out that research institutions in Southern Africa were offering various services such as support with data management plans, reskilling librarians, reference to high-performance computing centres, dedicated web pages, and advice on data preservation. At the Cape Peninsula University of Technology, research data management services were developed as an e-research information and communication infrastructure which included several components such as infrastructure development, information flow and management, communication with researchers, development of tools related to the full research life cycle and the means to store, curate, and retrieve data as well as the training of researchers (Chiware and Mathe, 2015). They emphasised the need for a national e-research infrastructure that would enable the preservation of research data. The University of Hong Kong utilised the research data stewardship framework which covered policy and

procedure settings for research data planning, the establishment of research data infrastructure, data curation services, and online resources and guidelines. John Hopkins University developed a new model of data management services involving storage, archiving, preservation, and curation layers (Shen and Varvel, 2013).

Cox *et al*. (2014), Naume (2014), Searle *et al*., (2015), Perrier and Barnes (2018) and Martin-Melon *et al*. (2023) stated that librarians had been playing a leading role in the establishment of research data management services and libraries had been leading in policy development, as supported by Cox and Pinfield (2014). It was noted that libraries were offering advisory, support, and training services rather than technical services, although there were some indications that librarians were upskilling to be able to remain relevant in the new research data management landscape. They added that librarians were not doing the research data management services alone but involved other key stakeholders from the IT services department and research support offices, legal office, including the researchers themselves (Akers *et al.*, 2014, Cox and Pinfield, 2014, Cox *et al*., 2017). Davidson *et al*. (2014) pointed out the activities that were done in supporting research data services by the Digital Curation Centre which include understanding funding bodies' policies, working with individual UK universities to scope RDM and data sharing challenges and opportunities, fostering RDM skills development, supporting data management planning, facilitating data discovery, and assessing research data management costs and benefits.

Challenges faced when establishing research data repositories

Institutions encounter several challenges in managing research data, as stated by Patel (2016), Chigwada (2022), Chiware (2020), Masenya (2021), Chigwada *et al*. (2017), Chiparausha and Chigwada (2019), Patterton *et al*. (2018), Tang and Hu (2019), Al-Jaradat, (2021), Ashiq *et al*. (2021), Huang *et al*. (2021), Ran *et al*. (2021), M'kulama *et al*. (2022), Chiware and Becker (2018), Koopman and De Jager (2016), Chiware and Mathe (2015), Raju (2014), Mohammed and Ibrahim (2019), and Nhendodzashe and Pasipamire (2017). It was noted that the challenges go beyond the institutional level but also include national challenges, as stated by Schopfel and Rebouillat (2022) and Knight (2015). The challenges that were pointed out include lack of storage space on institutional networks, limited computing power and cloud computing accessibility, poor state of research infrastructure, lack of government commitment to fund research data services, lack of clear policy guidelines, uncertainty of software tools to use, uncertainty on documentation standards to apply, security issues, interoperability issues, lack of skills and absence of research data management in some library schools, persistent brain drain, lack of awareness of RDM, and poorly resourced academic and research libraries as shown in Table 1.

Akers *et al*. (2014) indicated that the challenges that were faced by the eight US universities they studied included difficulties reaching out to researchers for assistance with research data management services and seeking funding for the human resources needed and infrastructure. Cox *et al*. (2017) indicated that libraries play a leading role in offering research data management services but are facing challenges such as low levels of engagement by key stakeholders, uncertainty on the technical infrastructure required, and funding issues. The findings from the study noted a lack of recognition of the need for tackling research data management at the institutional level, difficulties in getting institutional buy-in from the senior management, convincing some academics of the importance and worth of research data management services and some did not get support from the library management within the department (Cox *et al.,* 2014).

The future of research data repositories

The National Institute of Health (2023) pointed out the desirable characteristics for all data repositories which include unique persistent identifiers, long-term sustainability, metadata, curation and quality assurance, free and easy access, broad and measured reuse, clear use guidelines, security

and integrity, confidentiality, common format, provenance, and retention policy. In addition, the Confederation of Open Access Repositories (COAR) (2022) stated the framework of good practice in repositories, which includes discoverability through the use of metadata standards; harvesting of metadata using OAI-PMH; assigning of persistent identifiers, and registration on the registry of repositories; access including limiting access to sensitive research data; reuse through licencing information in the metadata record; integrity and authenticity to prevent unathorised manipulation of resources; quality assurance in line with the policies and procedures; preservation with a digital preservation and business continuity plan; sustainability and governance in terms of managing and funding the repository; and other documentation that provide the scope of the materials that are accepted in the repository. To be able to assess the usage of the research data repository, it should be able to show the usage metrics and citation count (Downs et al., 2023). Schopfel and Rebouillat (2022) stated that the international best practice which includes registration in the re3data directory, certification through the CoreTrustSeal certificate, the World Data System, or DTA Seal of Approval should be followed. A national approach to research data management was also suggested by Keller (2015) and Patterton *et al*. (2019). The issues of good practices and good standards were emphasised by Trippel and Zinn (2021). To develop institutional research data repositories that meet international standards, universities in Africa should work on the proposed framework in Figure 3.

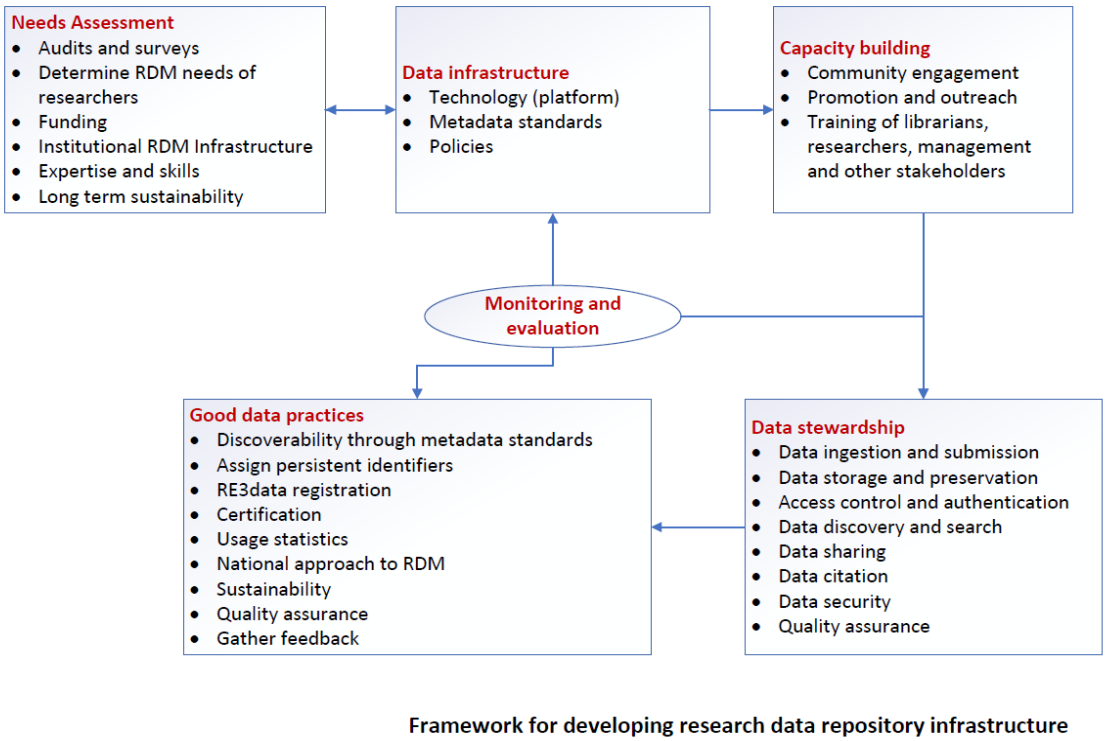

Figure 3: Proposed framework for developing data repository infrastructure.

Limitations of the study

The study only considered published literature, which was easier to retrieve. There is a need to consider grey literature as well, in terms of unpublished reports that document the experiences of data librarians in establishing and maintaining research data repositories in research institutions. This would be done by incorporating the sixth (optional) stage of the methodological framework to consult data librarians as a follow-up to this proposed framework as a way of getting consumer and stakeholder involvement to get additional references and insights beyond those in the literature.

Conclusion

Research data infrastructures across research domains, institutions, national boundaries, and beyond continue to grow as the need for good data management practices and sharing is now internationally recognised. The future of African data repositories depends on the development of sustainable platforms that have all the features of internationally trusted repositories, which are secure and driven by clear use guidelines and ensure integrity and confidentiality. The issue of costs is important and collaborative approaches in open source-based development are the only sustainable route to ensure the long-term curation and preservation of African-generated research data outputs. Continuous skills development especially among university librarians, research offices, and central Information Technology services is important as the technological landscape is always in a state of constant change. International pressure, especially from donors, funders, and publishers is likely to drive speedy development and uptake of data repositories across institutions in Africa.

Table 1: Challenges faced when establishing RDM services

Challenge	**Authors**
Lack of storage space on institutional networks	Chiware and Becker (2018), Knight (2015), Masenya (2021), Patterton *et al*. (2018), Tang and Hu (2019)
Limited computing power and cloud computing accessibility	Knight (2015), Patterton *et al*. (2018),
Poor state of research infrastructure	Chigwada et al, (2017), Chigwada (2022), Chiparausha and Chigwada (2019), Chiware (2020), Cox *et al*. (2019), Huang *et al*. (2021), Mohammed and Ibrahim (2019), Patterton *et al*. (2018), Tang and Hu (2019)
Lack of government commitment to fund research data services	Chiware (2020)
Lack of clear policy guidelines	Al-Jaradat, (2021), Ashiq *et al*. (2021), Chigwada et al, (2017), Chigwada (2022), Chiparausha and Chigwada (2019), Chiware and Becker (2018), Huang *et al*. (2021), Masenya (2021), M'kulama *et al*. (2022), Mohammed and Ibrahim (2019), Nhendodzashe and Pasipamire (2017), Ran *et al*. (2021),
Lack of mandate/ rewards	Chiware and Becker (2018), Cox *et al*. (2019), Huang *et al*. (2021), Masenya (2021),
Lack of institutional buy-in from senior management	Ashiq *et al*. (2021), Chigwada et al, (2017), Chiware and Becker (2018), Cox *et al.,* 2014; Cox *et al*. (2019), Mohammed and Ibrahim (2019), Tang and Hu (2019)
Uncertainty on documentation standards to apply	Knight (2015), Tang and Hu (2019)
Uncertainty of software tools to use and technical infrastructure required	Cox *et al*. (2017), Knight (2015), Mohammed and Ibrahim (2019),
Security issues	Al-Jaradat, (2021), Chigwada et al, (2017), Chiparausha and Chigwada (2019), Chigwada (2022), Cox *et al*. (2019), Huang *et al*. (2021), Knight (2015), Koopman and De Jager (2016), Patel (2016), Patterton *et al*. (2018),
Interoperability issues	Knight (2015), Mohammed and Ibrahim (2019),

Lack of skills	Al-Jaradat, (2021), Ashiq *et al*. (2021), Chigwada et al, (2017), Chiparausha and Chigwada (2019), Chigwada (2022), Chiware and Mathe (2015), Cox *et al*. (2019, Huang *et al*. (2021), Masenya (2021), M'kulama *et al*. (2022), Mohammed and Ibrahim (2019), Nhendodzashe and Pasipamire (2017), Patterton *et al*. (2018), Raju (2014), Ran *et al*. (2021), Tang and Hu (2019)
Absence of research data management in library schools	Raju (2014)
Persistent brain drain	Chiware (2020)
Poorly resourced academic and research libraries	Chigwada et al, (2017), Cox *et al*. (2019), Huang *et al*. (2021),
Funding	Ashiq *et al*. (2021), Akers *et al*. (2014), Chigwada et al, (2017), Chiparausha and Chigwada (2019), Chiware and Mathe (2015), Chiware and Becker (2018), Cox *et al*. (2017), Cox *et al*. (2019), Huang *et al*. (2021), Masenya (2021), Mohammed and Ibrahim (2019), Patterton *et al*. (2018), Tang and Hu (2019)
Researchers not willing to partner	Ashiq *et al*. (2021), Akers *et al*. (2014), Chigwada et al, (2017), Chiware and Becker (2018), Cox *et al.,* 2014, Huang *et al*. (2021), Patel (2016), Patterton *et al*. (2018), Tang and Hu (2019)
Low level of engagement by stakeholders	Ashiq *et al*. (2021), Chigwada et al, (2017), Cox *et al*. (2017), Cox *et al*. (2019), Huang *et al*. (2021), Patel (2016), Patterton *et al*. (2018), Ran *et al*. (2021), Tang and Hu (2019)
Lack of recognition for tackling RDM at the institutional level	Ashiq *et al*. (2021), Chigwada et al, (2017), Cox *et al.,* 2014, Cox *et al*. (2019), Huang *et al*. (2021), Patel (2016), Tang andHu (2019)
Lack of awareness	Ashiq *et al*. (2021), Chigwada (2022), Huang *et al*. (2021), Patel (2016), Tang and Hu (2019)

References

Abebe, R., Aruleba, K., Birhane, A., Kingsley, S., Obaido, G., Remy, S. L., and Sadagopan, S. (2021) "Narratives and Counternarratives on Data Sharing in Africa" in *Conference on Fairness, Accountability, and Transparency (FAccT '21)*, March 3–10, 2021, Virtual Event, Canada. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3442188.3445897.

Akers, K. G., Sferdean, F. C., Nicholls, N. H., and Green, J. A. (2014) "Building Support for Research Data Management: Biographies of Eight Research Universities", *International Journal of Digital Curation,* 9(2), pp. 171–191. doi:10.2218/ ijdc.v9i2.327.

Al-Jaradat, OM. (2021) "Research data management (RDM) in Jordanian public university libraries: Present status, challenges and future perspectives", *The Journal of Academic Librarianship*, 47 (5) 102378. https://doi.org/10.1016/j.acalib.2021.102378.

Arksey, H and O'Malley, L. (2005) ”Scoping studies: towards a methodological framework”, *International Journal of Social Research Methodology*, 8 (1), pp. 19-32, DOI: 10.1080/1364557032000119616.

Ashiq, M., Saleem, QUA and Asim, M. (2021) ”The Perception of Library and Information Science (LIS) Professionals about Research Data Management Services in University Libraries of Pakistan”. *Libri*, 71 (3), pp. 239-249. https://doi.org/10.1515/libri-2020-0098.

Bangani, S. and Moyo, M. (2019) “Data sharing practices among researchers at South African universities”, *Data Science Journal*, 18 (28), pp. 1–14. DOI: https://doi.org/10.5334/dsj-2019-028.

Chigwada, JP. (2022) "Management and maintenance of research data by researchers in Zimbabwe", *Global Knowledge, Memory and Communication*, 71 (4/5), pp. 193-207. https://doi.org/10.1108/GKMC-06-2020-0079.

Chigwada, J. P., Hwalima, T., and Kwangwa, N. (2019) ”A Proposed Framework for Research Data Management Services in Research Institutions in Zimbabwe” in R. Bhardwaj, and P. Banks (Eds.), *Research Data Access and Management in Modern Libraries.* Hershey, PA: IGI Global. doi:10.4018/978-1-5225-8437-7.ch002, (pp. 29-53).

Chigwada, J., Chiparausha, B. and Kasiroori, J. (2017) “Research data management in research institutions in Zimbabwe”, *Data Science Journal,* 16 (31), pp. 1–9, DOI: https://doi.org/10.5334/dsj-2017-031.

Chiparausha, B., and Chigwada, J. P. (2019) ”Accessibility of Research Data at Academic Institutions in Zimbabwe” in R. Bhardwaj, and P. Banks (Eds.), *Research Data Access and Management in Modern Libraries.* Hershey, PA: IGI Global. doi:10.4018/978-1-5225-8437-7.ch004. (pp. 81-89).

Chiware, E. (2020) ”Open research data in African academic and research libraries: a literature analysis”, Library Management, 41 (6/7,) pp. 383-399. DOI 10.1108/LM-02-2020-0027.

Chiware, E and Becker, DA. (2018) ”Research Data Management Services in Southern Africa: A Readiness Survey of Academic and Research Libraries”, *Afr. J. Lib. Arch. & Inf. Sc.,* 28, (1), pp. 1-16.

Chiware, E and Mathe, Z. (2015) ”Academic libraries’ role in Research Data Management Services: a South African perspective*”, SA Jnl Libs & Info Sci,* 81(2), pp. 1-10. doi:10.7553/81-2-1563.

COAR. (2022) *COAR community framework for good practices in repositories*, *Version 2 - July 19. Available at:* http://www.coar-repositories.org/ (Accessed 10 September 2023).

Cox, AM., Kennan, MA., Lyon, L., Pinfield, S and Sbaffi, L. (2019) ”Maturing research data services and the transformation of academic libraries”, *Journal of Documentation,* 75 (6), pp. 1432-1462. 0022-0418 DOI 10.1108/JD-12-2018-0211.

Cox, AM., Kennan, MA., Lyon, L., and Pinfield, S. (2017) ”Developments in Research Data Management in Academic Libraries: Towards an Understanding of Research Data Service Maturity”, *Journal of the Association for Information Science and Technology*, 68(9), pp. 2182–2200.

Cox, A.M., and Pinfield, S. (2014) ”Research data management and libraries: Current activities and future priorities”, *Journal of Librarianship and Information Science*, 46, pp. 299–316. http://doi.org/10.1177/ 096100061349254.

Davidson, J., Jones, S., Molloy, L and Kejser, UB. (2014) ”Emerging good practice in managing research data and research information within UK Universities”, *Procedia Computer Science,* 33, pp. 215 – 222.

Dora, M and Kumar, HA. (2015) ”Managing research data in academic institutions: role of libraries” in 10 th International CALIBER-2015 HP University and IIAS, Shimla, Himachal Pradesh, India March 12-14, INFLIBNET Centre, Gandhinagar, Gujarat, India, pp 484-495.

Downs, RR, Urquidi Díaz, A, Xu, Q, Wang, J, Chambodut, A, Liu, C, Flower, S and Payne, K. (2023) ”Harvestable Metadata Services Development: Analysis of Use Cases from the World Data System”, *Data Science Journal*, 22 (20), pp. 1–20. DOI: https://doi.org/10.5334/dsj 2023-020.

Huang, Y., Cox, A. and Sbaffi, L. (2021) ”Research data management policy and practice in Chinese university libraries”, *Journal of the Association for Information Science and Technology*, 72 (4), pp. 493-506. https://doi.org/10.1002/asi.24413.

Jones, S., Pryor, G. and Whyte,A. (2013) *How to Develop Research Data Management Services - a guide for HEIs*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available at: http://www.dcc.ac.uk/resources/how-guides/how-develop-rdmservices#sthash.kfiImTyy.dpuf. (Accessed 10 September 2023).

Keller, A. (2015) ”Research Support in Australian University Libraries: An Outsider View”, *Australian Academic & Research Libraries*, 46 (2), pp. 73-85, DOI: 10.1080/00048623.2015.1009528.

Knight, G. (2015) ”Building a research data management service for the London school of hygiene & tropical medicine”, *Program: electronic library and information systems,* 49 (4), pp. 424-439. DOI 10.1108/PROG-01-2015-0011.

Koopman, MM. and De Jager, K. (2016) ”Archiving South African digital research data: How ready are we?”, *S Afr J Sci*., 112 (7/8), pp. 1-7. http://dx.doi.org/10.17159/sajs.2016/20150316.

Lee, DJ. and Stvilia, B. (2017) ”Practices of research data curation in institutional repositories: A qualitative view from repository staff”, *PloS ONE,* 12 (3). https://doi.org/10.1371/journal.pone.0173987.

Levac, D., Colquhoun, H. and O’Brien KK. (2010) ”Scoping studies: advancing the methodology”, *Implement Sci.* 5 (69). doi:10.1186/1748-5908-5-69.

Lin, D., Crabtree, J., Dillo, I. et al. (2020). ”The TRUST Principles for digital repositories”. *Scintific Data.* 7, 144 (2020). https://doi.org/10.1038/s41597-020-0486-7

Martin-Melon, R., Hernandez-Perez, T. and Martinez-Gardama. (2023) ”Research data services (RDS) in Spanish academic libraries”, *The Journal of Academic Librarianship*. 49 (2023) 102732. https://doi.org/10.1016/j.acalib.2023.102732

Masenya, TM. (2021) "Research data management practices and services in South African academic libraries", *Library Philosophy and Practice (e-journal)*, 6311. https://digitalcommons.unl.edu/libphilprac/6311.

M’kulama, ACM., Zulu, Z, Chewe, P and Mwiinga, TM. (2022) "Preparedness for Open Science through Research Data Management at the University of Zambia in COVID-19 and Post-COVID Eras*", Library Philosophy and Practice (e-journal). 7268*. https://digitalcommons.unl.edu/libphilprac/7268.

Mohammed, MS. and Ibrahim, R. (2019) ”Challenges and Practices of Research Data Management in Selected Iraq Universities*”, DESIDOC Journal of Library & Information Technology*, 39 (6), pp. 308-314. DOI : 10.14429/djlit.39.6.14443.

Mushi, GE, Pienaar, H. and van Deventer, M. (2020) ”Identifying and Implementing Relevant Research Data Management Services for the Library at the University of Dodoma, Tanzania”, *Data Science Journal*, 19 (1), pp. 1–9. DOI: https://doi.org/10.5334/dsj-2020-001.

National Institute of Health. (2023*) Selecting a data repository.* Available at: https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/selecting-a-data-repository. (Accessed 10 September 2023).

Naum, A. (2014) ”Research data storage and management: library staff participation in showcasing research data at the University of Adelaide”, *The Australian Library Journal,* 63 (1), pp. 35-44. DOI: 10.1080/00049670.2014.890019.

Nhendodzashe, N. and Pasipamire, N. (2017) “Research data management services: are academic libraries in Zimbabwe ready? The case of University of Zimbabwe library”, *IFLA Satellite Meeting*, Wroclaw. Available at: https://library.ifla.org/id/eprint/1728/1/S06-nhendodzashe-en.pdf. (Accessed 10 September 2023).

Nie, H., Luo, P.C. and Fu, P. (2021) ”Research data management implementation at Peking University Library: Foster and promote Open Science and Open Data”, *Data Intelligence* 3(1), pp. 189-204. doi: 10.1162/dint_a_00088.

Patel, D. (2016) “Research data management: a conceptual framework”, *Library Review*, 65 (4-5), pp. 226-241. http://dx.doi.org/10.1108/LR-01-2016-0001.

Patterton, L., Bothma, T.J.D. and van Deventer, M.J. (2018) “From planning to practice: an action plan for the implementation of research data management services in resource-constrained institutions”, *South African Journal of Library and Information Sciences,* 84 (2), pp. 14-26. doi:10.7553/84-2-1761.

Perrier, L and Barnes, L. (2018) ”Developing research data management services and support for researchers: A mixed methods study”, *Partnership: The Canadian Journal of Library and Information Practice and Research*, 13, (1). http://dx.doi.org/10.21083/partnership.v12i2.4115.

Raju, J. (2014) ”Knowledge and skills for the digital era academic library”, *The Journal of Academic Librarianship*, 40 (2), pp. 163-170. https://doi.org/10.1016/j.acalib.2014.02.007.

Ran, C., Yang, L and Hu, L. (2021) ”Revisit the implementation status of research data management in Chinese academia”, *The Journal of Academic Librarianship*, 47 (2021) 102350. https://doi.org/10.1016/j.acalib.2021.102350.

Schopfel, J. and Rebouillat, V. (2022) ”The landscape of research data repositories in France”, in Schopfel, J. and Rebouillat, V. (eds) *Research Data Sharing and Valorization: Developments, Tendencies, Models*, pp31-48. https://doi.org/10.1002/9781394163410.ch2.

Searle, S., Wolski, M., Simons, N. and Richardson, J. (2015) "Librarians as partners in research data service development at Griffith University", *Program: electronic library and information systems*, 49 (4), pp. 440 – 460. http://dx.doi.org/10.1108/PROG-02-2015-0013.

Shen, Y. and Varvel, VE. (2013) ”Developing Data Management Services at the Johns Hopkins University” *The Journal of Academic Librarianship*, 39 (6), pp. 552-557. https://doi.org/10.1016/j.acalib.2013.06.002.

Tang, R. and Hu, Z. (2019) “Providing research data management (RDM) services in libraries: preparedness, roles, challenges, and training for RDM practice”, *Data and Information Management*, 3 (2), pp. 84-101. https://doi.org/10.2478/dim-2019-0009.

Trippel, T. and Zinn, C. (2021) ”Lessons learned: on the challenges of migrating a research data repository from a research institution to a university library”, *Lang Resources & Evaluation,* 55, pp. 91–207. https://doi.org/10.1007/s10579-019-09474-4.

Xiao, S., Ng, T.Y. and Yang, T.T. (2022) "Research data stewardship at the University of Hong Kong", *Library Management*, 43 (1/2) ,pp. 128-147. https://0-doi-org.oasis.unisa.ac.za/10.1108/LM-09-2021-0079.

Endnotes

[1] Josiline Chigwada is a Postdoctoral Fellow at the University of South Africa, School of Interdisciplinary Research and Graduate Studies. She can be reached by email: chigwaj@unisa.ac.za.

[2] Elisha Chiware is the Director at the Cape Peninsula University of Technology Libraries. He can be reached by email: chiwaree@cput.ac.za