Working towards securing and building a trusted institutional research data repository through the CoreTrustSeal process: case of Cape Peninsula University of Technology data repository

Janine Lockhart[i], Xabiso Xesi[ii] and Elisha R T Chiware[iii]

Abstract

In support of the Open Science movement and as a signatory of the Berlin Declaration, the Cape Peninsula University of Technology has since 2013 developed various systems, infrastructures and workflows to support open access and good research data management practices at the institution, providing a highly functional environment. Institutional policies that include a Research Data Management Policy and an Open Access Policy, data deposit guidelines and data deposit platforms are currently in place and utilized by affiliated postgraduate students and researchers from faculties, research units and entities as well as researchers from academic support units in alignments with FAIR principles. The requirement of postgraduate students to submit their research data with their theses for graduation purposes has increased the advocacy and publishing of datasets. The purpose of this paper is therefore to highlight the initial developmental trajectory of a research data repository and what was achieved to date. This includes the selection of the platform through the Ilifu project in the Western Cape, the implementation and strengthening of the repository review workflows to include a number of stakeholder players to ensure the quality and integrity of the data as well as ethics approval checks, the development of the data management planning tool and a more recent upgrade to include a section for the South Africa's Protection Of Personal Information Act (Nr 4 of 2013) compliancy, advocacy, training and processes that the institution has embarked on to secure the research data platform through proper preservation methodologies and approaches. Some challenges are discussed and how these were addressed. The paper also outlines the process of how the institution embarked on applying to have the research data repository certified as trustworthy through the CoreTrustSeal.

Keywords

Research data management, research data repositories, academic libraries, Cape Peninsula University of Technology, CoreTrustSeal, trusted digital repository (TDR), data preservation

Introduction

Research data repositories have been on the rise in the last few decades as various international and national organizations, funding agencies, publishers and research communities demand effective and efficient online access to digitally stored research data. Government agencies are mandating funding recipients to make research data publicly available in approved repositories (Hutchison et al., 2021). Research data repositories and archives are important components of the research infrastructure, providing resources and services to research communities. The growth and development of research data repositories is also seen as a direct response to several international and national declarations on open science that aim to advance discoverability and use of scientific research publications and data in an open and transparent manner. The provision of research data repositories is at several levels, including: institutional, regional, national, international and discipline specific. Institutional level research data repositories are often maintained to store internally generated research data whose curation, preservation and dissemination is governed by clear policies and guidelines. At regional, national and international levels data repositories can come in different forms, including discipline specific to generalist types of repositories.

Research data repositories have had the positive effect on the research processes by allowing researchers from different institutions and disciplines to share research workflows, experimental methods and data. To ensure that there is a global standard on the discoverability and usability of research data, key stakeholders from industry, academic, funding agencies and publishers have designed and endorsed a concise set of principles known as FAIR (Findable, Accessible, Interoperable, Reusable) Data Principles (Wilkinson et al., 2016).

South Africa stands out in Sub-Saharan Africa with its advances and sustained funding for the development of research infrastructure across institutions and at national level. The level of funding and development is enabling the development and integration of data repositories into existing research infrastructures (Chiware and Becker, 2018). According to the Registry of Research Data Repositories (re3data)[iv], South Africa has the most registered repositories than any other country in Sub-Saharan Africa with more than 17 institutional and discipline specific platforms. Most of the research data repositories are found in academic institutions which are the main knowledge production centres in the country. The growth of South African repositories has also been bolstered by the main government research funding agency, the National Research Foundation, which in 2015 mandated that all its grant recipients must deposit their research outputs, including research data, into trusted institutional repositories (NRF, 2015). A significant increase in the number of data repositories and data deposits in the past ten years is supported largely by several universities' Research Data Management (RDM) and Open Science policies.

One of the key issues in the development of research data repositories at the global level is a roadmap to achieving trustworthy digital repositories (TDR) status. Johnston (2012) outlines that a trusted digital repository is "a set of metrics that are used to certify that a given repository is an appropriate custodian of a collection of digital assets". Furthermore, Johnston (2012) emphasizes that a trustworthy digital repository must be a stable and sustainable platform, following a clear set of policies and procedures for the sound management of digital assets, housed within secure technical environments. Faundeen (2017) suggests that in order to secure and gain certification for digital

repositories, it is important to follow the guidelines set by national and international organizations and establish national policies and data governance guidelines. Bak (2016), was of the view that "the notion of trust within trustworthy digital repositories standards culture is itself evolving in the positive direction that emphasizes user perceptions of trust rather than seeking to establish objective evidence of trust". Yoon (2014) also emphasized that much attention has been paid to establishment of ISO standards towards trustworthy digital repositories, with very little attention on the users who are equally important. It is important for data repositories to maximize research outcomes and facilitate collaboration and sharing, as well as, ensuring the quality of the data and accompanying services (Mehnert et al., 2019).

Johnston (2012) clearly states that to be certified as a trustworthy digital repository, organizations must undergo an audit which will ensure that their repository meets all criteria of certifying body. Johnston outlines the need for overall information management processes, access, data security systems, and risk management parameters.

In this paper we describe the technical and non-technical process around the historical development of eSango, the research data repository at the Cape Peninsula University of Technology (CPUT) powered by Figshare, and the roadmap towards striving to achieve the status of a trustworthy research data repository through the CoreTrustSeal[v] process.

Cape Peninsula University of Technology Research Data Management Services

The Cape Peninsula University of Technology's Research, Technology and Innovation (RTI) 10-year blueprint (2012) outlined the key role that the university library was to play in supporting research, technology and innovation at the institution, which included: "Curation, dissemination and promotion of the traditional outputs of research in terms of articles and theses, and curation of research data and innovation output, including enhanced research data management systems". This recognition of the library's role in the provision of research data management services provided the basis on which RDM services were developed at the institution (Chiware and Mathe, 2015).

Since 2013, CPUT has put in place policies and developed systems and workflows to support good RDM practices supporting a strong open access environment at the institution. These guideline and infrastructure were used by students and researchers at all levels (Chiware and Mathe, 2015).

Research data environment

At the beginning, Research Data Management (RDM) at CPUT was placed in the library in a division called Knowledge, Information and Technology Services (KITS). This division was instrumental in creating platforms, systems, and processes for research data management. To advance the adoption of RDM practices, CPUT libraries established collaboration with several institutional stakeholders to develop policies, build infrastructure, train library staff, and conduct awareness and advocacy campaigns with academic staff and researchers. Tripathi et al. (2017) highlights the importance of Data Management Plans (DMP) and the role of libraries in supporting researchers in storing and accessing their data. According to Wilkinson et al. (2016), good RDM is important for knowledge discovery and innovation, and for subsequent data and knowledge integration and reuse by the community after the data publication process. To encourage data discoverability and reuse, in 2020 CPUT's Higher Degrees Committee (HDC) mandated that as part of the graduation process, Master

and Doctorate students must share the datasets used in their research on eSango, the CPUT research data repository. In addition to ensuring that essential research data is kept accessible, available for future reference, and verification and strengthens the transparency of the research this requirement also increased advocacy for RDM practices at the institution.

CPUT has taken a stance in applying the four foundational principles of good RDM that is Findability, Accessibility, Interoperability and Reusability, known as FAIR principles, in managing research data (Ntja, 2022). The driving force in promoting good RDM stewardship is due to the impact it has to high-quality digital publications that facilitate and simplify the ongoing process of discovery, evaluation, and reuse in teaching, learning and research.

Policy framework

Understanding the importance of data for scientific advancement and human development, CPUT's administration was interested in guaranteeing that data produced by researchers at the institution are of high quality and are openly accessible for reuse by other researchers. CPUT also recognized that good research data management procedures are essential for a productive and efficient research process. For example, it is important to follow the ethical and legal guidelines when working with sensitive data. Good RDM responds to the Open Science and Open Data movements, which urge for more transparency and efficiency in research to accelerate the scientific enterprise.

In 2019, the South African Government published a White Paper on Science, Technology, and Innovation designed to strengthen the National System of Innovation (NSI) through research output sharing. In response to this White paper and to establish an RDM policy landscape at CPUT, in 2020, the university formed the Policy Working Group (PWG) comprised of faculty and staff from across the different disciplines. The PWG reviewed the University's policies in place from 2013. The policies were revised and published as Open Access (OA) Policy (2021a) and Research Data Management (RDM) Policy (2021b).

The objective of the CPUT RDM policy is to govern research data management, promote reproducibility and ensure compliance by all university staff and students. The policy aims to establish guidelines and procedures for the management, ownership, sharing, access, storage, preservation, and ethical handling of research data within the university community. It requires all individuals affiliated with the University to adhere to the defined principles and practices of responsible research data management. Having a policy in place is only the first step towards good RDM practices. Often researchers find these policies to be cumbersome and are therefore reluctant to adopt them. Researchers may be more likely to adopt RDM policies when they are required by external funding agencies or publishers. Offering training, consultation, and collaboration to university staff and researchers can also help with achieving full implementation of the policies. Building trust in the value of repositories to the researchers' work may also help with their adherence to RDM policies. This trust is often built based on the value of the repositories to researchers' work, the visibility of reused data and how institutions, funders and publishers respond to the compliance mandates (Curty, 2016; Swauger and Vision, 2015).

Data repository

Ilifu[vi] is a regional node in the Western Cape, South Africa, known as a Tier II node, in the national infrastructure. It supports research mostly in the astronomy and bioinformatic fields. Ilifu is funded partly by the Department of Science and Technology (DST) through their Data-Intensive Research Initiative of South Africa (DIRISA). The regional project involved four universities, Cape Peninsula University of Technology, University of Cape Town, University of the Western Cape and Stellenbosch University. The Research Data Management and Open Science component of the project involved the development of policies and guidelines for research data management, sharing, reuse, governance, and quality. As part of this process, participating institutions worked collaboratively to negotiate for a platform that would be suitable for storage of research data and that a country wide licence was negotiated to make it more affordable for all universities in the country, and to also enable archiving of all South African universities' research data through one platform, see Figshare South Africa[vii]. Each of the four institutions listed above have its own instance and contract with Figshare as a proprietary software. The CPUT research data repository, called eSango, went live in early 2018. With the launch of eSango, CPUT worked on strengthening and implementation the repository review workflows and policies.

Data review workflow

According to Mayernik et al. (2015) peer review is critical to the scientific communication system. He sees reviewing as both a community responsibility and an opportunity to polish and expand one's understanding of cutting-edge research. Adding research data to the publication and peer review queues will put additional strain on the scientific publishing system, however, it will also increase the trustworthiness and value of individual datasets, strengthen findings based on cited datasets, and improve transparency and traceability of data and publications. The data review process at CPUT involves several stakeholders and is based on CoreTrustSeal (CTS) requirements. The review process, see Figure 1, includes the Digital Scholarship Librarian, Metadata Librarian, Ethics Manager and in cases of postgraduate students, their Research Supervisor. The process is as follows:

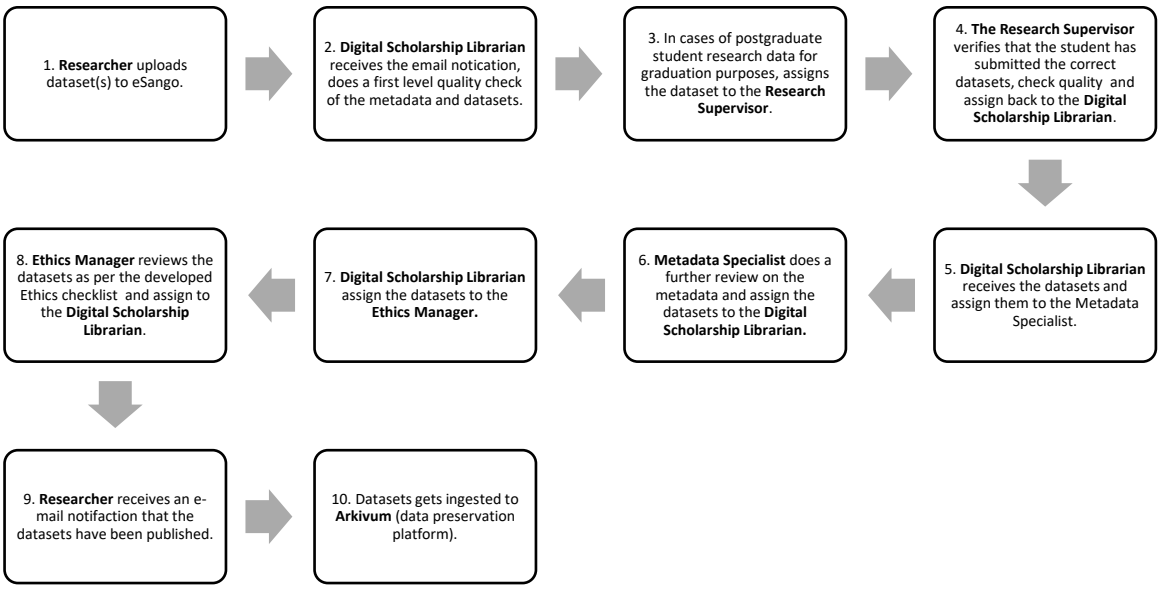

Figure 1: eSango Review Process Workflow

The standard submission form had to be adapted by Figshare so that the above workflow could be implemented at CPUT. It was important to know whether the submission was for graduation purposes, i.e., is it postgraduate students who submits research data related to their thesis (which is submitted through a different process). Therefore, the following question has been added to the submission: "Is this dataset for graduation purposes?" If yes, then the submitter needs to add the research supervisor's e-mail address. Adding this to the submission form provided the library team the needed information to identify postgraduate student research data submissions and to get the research supervisor's contact information so that they can be included in the review process. The library team does a basic level curation, but the quality and information in the actual dataset should be reviewed by an expert in the field.

Preservation

A preservation strategy is another important aspect to consider in securing and building trust in a research data repository, and one of the CTS requirements. To meet these requirements, CPUT acquired a cloud-based digital preservation and data management solution. The system called Arkivum, a leading digital preservation solution, was selected as a fully managed Software as a Service (SaaS), which manages the ingest process, safeguarding of data, preservation, supports over 100 formats, and providing discovery and access. Arkivum is a supporter of the Digital Preservation Coalition (DPC) and member of the National Data Stewardship Alliance (NDSA). While Figshare is a research data repository, Arkivum is a digital archiving and preservation software that archives and preserves data, therefore guaranteeing the longevity of the research data. Figshare and Arkivum had to develop and adapt their products so that they work seamlessly together.

Promotion of RDM

To promote effective RDM practices, CPUT Libraries offered several training options. First, Faculty Support Teams were trained on the eSango data repository as well as on the Data Management Planning (DMP) tool. This was followed by a series of online presentations (during lunchtimes) and workshops for faculty, aimed at highlighting various RDM tools and services available to researchers. One of the key tools discussed was the DMP tool, which is an essential aspect of responsible research data management. RDM training sessions always include a discussion of eSango and the DMP tool. The presentations were designed to provide researchers and postgraduate students (who are required to submit a DMP with their research proposal) with practical guidance on how to manage research data effectively and efficiently. During 2023 alone close to 30 training sessions were offered to CPUT researchers and postgraduate students. The workshops were well received, and the uptake and growth can be seen in the statistics of the DMP tool and the submission of datasets on eSango. Statistics obtained from the DMP tool suggests that as of the writing of this paper, 1,365 users registered and 1,360 DMP plans were started. Through these initiatives, the library has continued to play a vital role in supporting research excellence at CPUT.

Securing a trustworthy data repository status

General overview

Trust requirements in research data repositories are growing and as stated by Crabtree (2020) "trust in research data repositories is critical as they provide the evidence for past discoveries as well as the input for future discoveries". The CoreTrustSeal (CTS) is an internationally recognized standard for trustworthiness in digital repositories, ensuring that data is being managed in a secure and reliable manner. The CTS offers a process for core level certification based on the CTS 16+ requirements that reflect certain characteristics of trustworthy data repositories. These CTS requirements are a good assessment tool and are helpful in identifying gaps that need to be addressed. L'Hours, Kleemola and De Leeuw (2019) outline the history and background of the CTS and the formation of the requirements. As the CTS was officially launched in 2017, and it can take several years to get certified, to date there aren't many publications outlining this certification process in practice. Corrado (2019) examines the issue of trust in digital repositories. The author indicates that it is not clear if designated communities are influenced by certificates, however, repositories who meet the requirements may have a better foundation for building trust. The CTS requirements are reviewed every few years and adapted as needed. The certification process consists of several submissions and may take a few years to complete. Each submission of documentation to CTS is reviewed by two referees and the organization applying for certification has to respond. The time from submission to receiving reviewers' comments is about three months and it takes another three months to resubmit. The process can continue for up to five submissions. So far, the CTS has certified over 160 repositories around the world, one of which is on the African continent in South Africa.

CoreTrustSeal process

Data sharing is becoming an essential component of scientific research and scholarly publication. This necessitates informed and intentional planning, from early study planning to data and metadata collection, interoperability, deposit in data repositories, and curation (Austin et al., 2016). Magnuson and Thomas (2023) discuss their CTS application and list the five valuable lessons they learned:

institutional history, building the CTS application, business process model, leveraging documentation and preservation.

By applying for the CTS certification, CPUT demonstrated its commitment to promoting responsible RDM. To achieve the certification, CPUT has implemented a rigorous review process ensuring that its RDM policies and procedures meet the CTS requirements. By obtaining the CTS certification, CPUT will be positioned to better serve its researchers and enhance the visibility and impact of their research. The work on the CTS application was part of the larger Ilifu project undertaken during 2020. The four universities in the region worked on this together, holding regular meetings and supporting each other through the process. However, each university had their own Figshare data repository and applied separately for the CTS. CPUT's goal was to get the research data repository, eSango, which is powered by Figshare, to meet the core requirement of CTS and achieve a secured and trustworthy data repository. The 17 CTS Trustworthy Data Repositories Requirements (2020-2022) are detailed in Table 1.

Number	Requirement	Description
R0	Repository type, brief description, designated community, level of curation performed, insource/outsource partners.	Provide context of the repository
R1	Mission/Scope	The repository has an explicit mission to provide access to and preserve data in its domain.
R2	Licenses	The repository maintains all applicable licenses covering data access and use and monitors compliance.
R3	Continuity of access	The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.
R4	Confidentiality/Ethics	The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.
R5	Organizational infrastructure	The repository has adequate funding and enough qualified staff managed through a clear system of governance to effectively carry out the mission.
R6	Expert guidance	The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either in-

		house, or external, including scientific guidance, if relevant).
R7	Data integrity and authenticity	The repository guarantees the integrity and authenticity of the data.
R8	Appraisal	The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.
R9	Documented storage procedures	The repository applies documented processes and procedures in managing archival storage of the data.
R10	Preservation plan	The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.
R11	Data quality	The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.
R12	Workflows	Archiving takes place according to defined workflows from ingest to dissemination.
R13	Data discovery and identification	The repository enables users to discover the data and refer to them in a persistent way through proper citation.
R14	Data reuse	The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.
R15	Technical infrastructure	The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.
R16	Security	The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.

Table 1: CoreTrustSeal Trustworthy Data Repositories Requirements 2020-2022

Going through the submission and revision process streamlined and strengthened the internal workflows and led to better understanding of what it may take to build a trustworthy repository for CPUT. Most of the requirements for CTS focussed on the CPUT's internal environment regarding

policy, expertise, workflows, and preservation strategies that was in place. However, there were instances that documentation from Figshare was required. CPUT's application process is currently still in progress.

The scoring was done between 0 to 4 and described as seen in Table 2.

Score	Description
0	Not applicable
1	Not considered this yet
2	The repository has a theoretical concept
3	The repository is in the implementation phase
4	The guideline has been fully implemented in the repository

Table 2: CoreTrustSeal Scoring categories

Challenges

As much as libraries can implement new services, platforms and tools, the key challenge is often the advocacy and uptake of the services by the researchers and the university community at large. One of the important drivers in university environments is the policy landscape. Policies are key, but also take longer to develop, and review periods are often every three years and done consultatively as per the Policy on Policy Development. Once a policy is in place, advocacy is a key driver and this involves several different strategies which could include, being on the agenda of key university committee structures, regular communication via university communication channels, developing training by the library, sometimes as part other key departments training, e.g., the Directorate Research Development and Centre for Postgraduate Studies.

Since it was a new aspect to deal with the CTS process, it has been challenging. It was difficult to incorporate some of the discipline-specific repositories' requirements into CPUT's eSango research data repository since it is a generalist repository. However, working through the submission and revision process, helped the team improve and update the university's research data repository as well as the data deposit and preservation processes. While it was a steep learning curve, much was learned about good RDM practices.

A review process is essential to ensure the quality of research data. Since library staff may not have subject expertise, libraries are not able to provide more than a basic level curation of the datasets deposited. This highlights the importance of including discipline-specific specialists as part of the review process and this may be challenging to do. To overcome this challenge, CPUT decided to include the postgraduate students' research supervisors as part of the data review process. There are still hurdles in bringing students' research data and procedures into the review process and the library

has put in place webinars that equip students with the necessary skills on how to manage their research data throughout the research lifecycle.

Lessons learned

Embarking on this process may look daunting. However, establishing a small working group that includes two or three people, makes achieving this goal doable with limited additional workload for each person in the team. A systematic approach worked well for our team. We scheduled one-hour sessions over several days to go through each requirement and to write how we plan to meet each requirement. This approach was followed when addressing the reviewers' comments and making the necessary changes and/or notes to update workflows and guidelines. Additionally, it was helpful to look at other CTS approved repository documentation, to get a sense of what is required. If the repository uses proprietary software, it is helpful to ask them for support and to provide documentation and policies for aspect they are responsible for.

Future trends

Future trends could include more adoption by university libraries and research entities of certified research data repositories as a standard and thus ensuring improved research data management practices. This will lead to higher quality research data as per the FAIR principles. This will also lead to the requirement of increased preservation demands within university libraries; therefore, development of additional skills and preservation experts may be needed. An increase in the requirements of research dataset submissions from publishers and funders are expected, especially as governments put in place structures, white papers and statements regarding research funding from public funds.

Conclusion

RDM practices have evolved over the last decade and more publishers and funders require datasets as part of the publishing process and many governments have put in place strategies to ensure a good science, technology, and innovation landscape. To ensure best practices for RDM services and tools, it will be of excellent value to start with a self-assessment to measure a research data repository against the CTS requirements. This will assist with identifying gaps within the RDM environment at the university or research entity and lead to enhancements and standardization. A further step would be to submit the application to certify the research data repository for CTS approval.

References

Austin, C. C., Brown, S., Fong, N., Humphrey, C., Leahey, A. & Webster, P. 2016. Research data repositories: review of current features, gap analysis, and recommendations for minimum requirements. *IASSIST Quarterly*, 39, 24-24. DOI: https://doi.org/10.29173/iq904

Bak, G. 2016. Trusted by whom? TDRs, standards culture and the nature of trust. *Archival Science*, 16, 373-402. DOI https://doi.org/10.1007/s10502-015-9257-1

Cape Peninsula University of Technology. 2012. Research, Technology and Innovation (RTI) 10-year blueprint [Online]. Available: https://www.cput.ac.za/storage/research/research_directorate/RTI_Blueprint.pdf [Accessed 10 January 2024].

Cape Peninsula University of Technology. 2021a. Open Access (OA) Policy [Online]. Available: https://www.cput.ac.za/storage/library/pdf/research-support/CPUT-Open-Access-Policy.pdf [Accessed 11 January 2024].

Cape Peninsula University of Technology. 2021b. Research Data Management (RDM) Policy [Online]. Available: https://www.cput.ac.za/storage/library/pdf/research-support/Research-Data-Management-Policy.pdf [Accessed 09 May 2023].

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

Chiware, E. and Becker, D.A. 2018. Research Data Management Services in Southern Africa: a readiness survey of academic and research libraries. *African Journal of Library, Archives, and Information Science*, 28(1), 1-16.

Corrado, E.M. 2019. Repositories, trust, and the CoreTrustSeal. *Technical Services Quarterly,* 36(1), 61-72. DOI: https://doi.org/10.1080/07317131.2018.1532055

Crabtree, J. D. 2020. *Evidence for trusted digital repository reviews: an analysis of perspectives*. PhD Thesis. University of North Carolina. Available at: https://cdr.lib.unc.edu/concern/dissertations/2n49t9931 (Accessed: 17 January 2024)

Curty, R.G. 2016. Factors influencing research data reuse in the social sciences: an exploratory study. *Internal Journal of Digital Curation*, 11(1), 96-118. DOI: https://doi.org/10.2218/ijdc.v11i1.401

Department of Science & Technology. 2019. White paper on science, technology and innovation: science, technology and innovation enabling inclusive and sustainable South African development in a changing world. [Online]. Available: https://www.dst.gov.za/images/2019/White_paper_web_copyv1.pdf [Accessed 11 January 2024].

Faundeen, J. 2017. Developing criteria to establish trusted digital repositories. *Data Science Journal*, 16, 22-22. DOI: https://doi.org/10.5334/dsj-2017-022

Hutchison, V. B., Norkin, T., Langseth, M. L., Ignizio, D. A., Zolly, L. S., Mcclees-Funinan, R. & Liford, A. 2021. Leveraging existing technology: developing a trusted digital repository for the US Geological Survey. *International Journal of Digital Curation*, 16, 23-23. DOI: https://doi.org/10.2218/ijdc.v16i1.741

Johnston, W. 2012. Digital preservation initiatives in Ontario: trusted digital repositories and research data repositories. *Partnership: The Canadian Journal of Library and Information Practice and Research*, 7(2). DOI: https://doi.org/10.21083/partnership.v7i2.2014

L'Hours, H., Kleemola, M., De Leeuw, L. 2019. CoreTrustSeal: from academic collaboration to sustainable services. *IASSIST Quarterly*, 43(1). DOI: https://doi.org/10.29173/iq936

Magnuson, D. L., Thomas, W. L. 2023. Expanding our perspective: building a sustainable metadata culture. *IASSIST Quarterly*, 47(2). DOI: https://doi.org/10.29173/iq1046

Mayernik, M. S., Callaghan, S., Leigh, R., Tedds, J. & Worley, S. 2015. Peer review of datasets: when, why, and how. *Bulletin of the American Meteorological Society*, 96, 191-201. DOI: https://doi.org/10.1175/BAMS-D-13-00083.1

Mehnert, A. J., Janke, A., Gruwel, M., Goscinski, W. J., Close, T., Taylor, D., Narayanan, A., Vidalis, G., Galloway, G. & Treloar, A. 2019. Putting the trust into trusted data repositories: a federated solution for the Australian National Imaging Facility. *International journal of digital curation*. 14(1) DOI: https://doi.org/10.2218/ijdc.v14i1.594

National Research Foundation. 2015. Statement on Open Access to research publications from the National Research Foundation (NRF) funded research. [Online]. Available: https://chelsa.ac.za/wp-content/uploads/2018/12/NRF_open_access_-statement_19Jan2015_V6.pdf [Accessed 11 Jan 2024].

Ntja, B. 2022. Enhancing the role of the libraries in South African Higher Education Institutions through research data management: a case study of Cape Peninsula University of Technology, Faculty of Humanities, Library and Information Studies Centre (LISC). http://hdl.handle.net/11427/37692

Swauger, S., Vision, T.J. 2015. What factors influence where researchers deposit their data? A survey of researcher submissions to data repositories. *International Journal of Digital Curation*, 10(1), 68-81. DOI https://doi.org/10.2218/ijdc.v10i1.289

Tripathi, M., Shukla, A. & Sonkar, S. K. 2017. Research data management practices in university libraries: a study. *DESIDOC Journal of Library & Information Technology*, 37(6), 417-424. DOI https://doi.org/10.14429/djlit.37.6.11336

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. et al. 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific data*, 3(160018) DOI https://doi.org/10.1038/sdata.2016.18

Yoon, A. 2014. End users' trust in data repositories: definition and influences on trust development. *Archival Science*, 14, 17-34. DOI https://doi.org/10.1007/s10502-013-9207-8

Endnotes

[i] Janine Lockhart is the Library Manager: Scholarly Communication and Digital Scholarship at CPUT Libraries lockhartj@cput.ac.za.

[ii] Xabiso Xesi is the Digital Scholarship Librarian at CPUT Libraries xesix@cput.ac.za.

[iii] Prof Elisha R Chiware is the Director at CPUT Libraries chiwaree@cput.ac.za.

[iv] https://www.re3data.org/

[v] https://www.coretrustseal.org/

[vi] https://www.ilifu.ac.za/

[vii] https://southafrica.figshare.com/