

RESEARCH PAPER

Research Data Management in Research Institutions in Zimbabwe

Josiline Chigwada, Blessing Chiparausha and Justice Kasiroori

Bindura University of Science Education, ZW

Corresponding author: Josiline Chigwada (josyphiri@gmail.com)

The research was aimed at evaluating how research data are being managed in research institutions in Zimbabwe. The study also sought to assess the challenges that are faced in research data management by research institutions in Zimbabwe. Twenty five institutions of higher learning and other organisations that deal with research were selected using purposive sampling to participate in the study. An online questionnaire on SurveyMonkey was sent to the selected participants and telephone interviews were done to follow up on participants who failed to respond on time. Data that were collected using interviews were entered manually into SurveyMonkey for easy analysis. It was found out that proper research data management is not being done. Researchers were managing their own research data. Most of the research data were in textual and spreadsheet format. Graphical, audio, video, database, structured text formats and software applications research data were also available. Lack of guidelines on good practice, inadequate human resources, technological obsolescence, insecure infrastructure, use of different vocabulary between librarians and researchers, inadequate financial resources, absence of research data management policies and lack of support by institutional authorities and researchers negatively impacted on research data management. Authors recommend the establishment of research data repositories and use of existing research data repositories that are registered with the Registry of Research Data Repositories to ensure that research data standards are adhered to when doing research.

Keywords: research data; research data management; research data services; data sharing; research institutions

Introduction

The world has seen university libraries positioning themselves to support and gain authority on data management issues (Buys & Shaw 2015; Tenopir, Sandusky, Allard & Birch 2014; Whitmire, Boock, & Sutton 2015). This follows upon realising the importance of research data and proper research data management. Governments and funding organisations are increasingly demanding researchers to properly store and share data (Buys & Shaw 2015; Corti, Van den Eynden, Bishop & Woollard 2011; Kennan & Markauskaite 2015). Good data management is important as facilitates verification of research results thereby making it easier for other researchers to build on the existing research (Corti, Van den Eynden, Bishop & Woollard 2011). Currently, there is no evidence to show how research institutions are managing research data in Zimbabwe although a lot of research activities are being done. The research was aimed at evaluating how research data are being managed in research institutions in Zimbabwe and assess the challenges faced in research data management by research institutions in Zimbabwe.

Research data

The term research data was defined by Rice (2009) as data 'collected, observed or created for the purposes of analysing to produce original research results'. However, Kennan & Markauskaite (2015) went further to suggest that the data may not necessarily be used for research alone since the data include administrative records, log files of learning management systems and web portals and other behavioural traces used in learning analytics and traces of individual lives available from social media.

According to Kennan & Markauskaite (2015), research data, just like data sources, are heterogeneous because of the many forms depending on origins, research problem addressed and the discipline of the researcher. The authors note that in the life and physical sciences, researchers gather and produce data mostly through observations, experiments and computer modelling whilst in the social sciences researchers gather and produce data from interviews, surveys and questionnaires, and observations. The University of Essex (2017) listed several research data formats that include HTML, XML, MP4, MP3, JPEG, TIFF, CSV, DOC, PDF and TXT.

Research data management

According to Whyte & Tedds (2011), 'Research data management concerns the organisation of data, from its entry to the research cycle through to the dissemination and archiving of valuable results. It aims to ensure reliable verification of results, and permits new and innovative research built on existing information'.

Research data management is important because data are a valuable resource whose production requires time and money (Corti, Van den Eynden, Bishop & Woollard 2011). As a result, Corti, et al. (2011) put great emphasis on research data sharing because this enables scientific enquiry and debate, promotes innovation, transparency and accountability, examination of research findings, validation of research methods, avoiding duplicating data collection, research visibility, collaborations between and among data users and data creators. This would enable other researchers to discover, interpret and reuse the data as well as to sustain the value of the data by enabling others to verify and build upon the published results.

Corti, et al. (2011) suggested that best practice research data management should address issues relating to which data will be generated during research, metadata, standards and quality assurance measures, modalities for sharing and securing data, ethical and legal issues relating to data sharing that include copyright and intellectual property rights of data, data storage and backup, resources and costs associated with data management and, data management roles and responsibilities. Therefore, a data management plan addressing all these issues must be in place.

Fary & Owen (2013) stressed the importance of understanding the data lifecycle. The lifecycle informs data management for any institution. Fary & Owen (2013) summarised various data lifecycle models with their model presented in **Figure 1** below.

Tenopir, Sandusky, Allard & Birch (2014) acknowledged that there has been increasing need for libraries and librarians to play a leading role in research data management. The authors went further to give examples of new roles of librarians as a result of this new development. The new roles would see librarians as managers of data, datasets and data curation managers. Surkis & Read (2015) concurred and with Tenopir, et al. (2014) and further noted that librarians now provide a range of services in research data management that include teaching data management to researchers, assisting researchers to improve their data management practices, creating data management subject guides, and assisting in supporting funding agency and publisher data requirements.

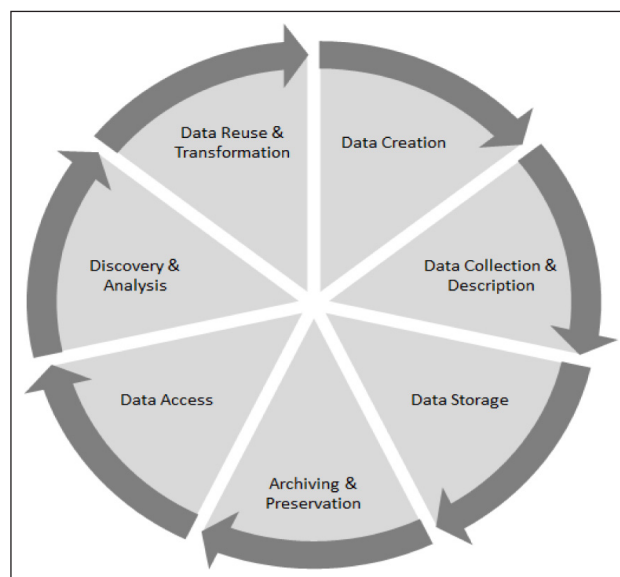


Figure 1: Data lifecycle model (Fary & Owen 2013).

According to Tenopir, Sandusky, Allard & Birch (2012) majority of librarians strongly felt that they have a responsibility of providing research data services to patrons and to increase institutional visibility and research impact. Tenopir, et al. (2015) found out that of those libraries offering research data services, librarians, committees and departments were responsible for research data services planning. The results also stressed that out of the 128 directors, 83.7% stated that librarians should be stewards of all types of scholarships including data sets; 68.6% indicated that losing data sets jeopardise future of scholarship while 76.7% pointed out that library needs to offer research data services to remain relevant to the institution. Surkis & Read (2015) stressed the importance of librarians in data management because there has been a paradigm shift from primarily focussing on publications as the only important research output as per past tradition towards recognising that research data are an important output of the research process. As such, involvement of librarians in data management has become so important as it facilitates data discoverability, accessibility, and understandability.

Challenges faced in research data management

Corti, Van den Eynden, Bishop & Woollard (2011) observed that research data management is not an easy task and data centres may not accept all data submitted to them; institutional repositories may not afford long-term maintenance of data, more complex research data may be difficult to store and manage and some websites are ephemeral with little sustainability.

Harvey (2010) as cited by Kennan & Markauskaite (2015) identified the following challenges associated with digital data management:

- Technology obsolescence;
- Technology fragility;
- Lack of guidelines on good practice;
- Inadequate financial and human resources to manage data well; and,
- Lack of evidence about best infrastructures.

Use of different vocabulary between librarians and researchers also hinders collaboration between these two players (Surkis & Read 2015). The cultures of the former and the latter are different. As Surkis & Read (2015) noted, researchers speak the language of research not the language of libraries.

In order to solve some of the challenges mentioned above, there is need to include all the data management stakeholders from the initial stages of the research to ensure that there is order throughout the research lifecycle. These include primary researcher or investigator, institution, data repository, user, funder, and publisher. If a researcher has a data management plan, it would help to describe the data produced during the research project as it outlines the strategies that would be implemented during the active phase of the research and after the project is complete. The data management plan may be one of the requirements to be submitted to funding agencies during the proposal stage.

Although Kennan & Markauskaite (2015) reported that there were inadequate human resources to manage data, a study carried out by Tenopir, et al. in 2012 revealed that 78% of the respondents for whom research data service was regarded as core of their duties indicated that they had the necessary skills, knowledge and training on research data management. Tenopir et al. (2016) concurred with Tenopir et al. (2012) that libraries offer opportunities for staff for research data services skills development by way of conferences, workshops, research data services related courses, professional development working groups and in house workshops and presentations.

Policies and guidelines on research data management

Section 2 of the Australian Code for the Responsible Conduct of Research states that 'policies are required that address the ownership of research materials and data, their storage, their retention beyond the end of the project, and appropriate access to them by the research community' (Australian National Data Service, 2015). The National University of Singapore (2016) stated that the design and adoption of policies for research data management help to safeguard valuable data. The policy helps to answer allegations of research misconduct and assists in the protection of intellectual property.

Libraries play a critical role in the implementation of research data policies, for example, the University of Leeds (2016) and the University of Manchester libraries (2016) host the Research Data Management policies of the universities. These libraries provide a research data management service to support researchers. Monash University (2013) stated that the purpose of a research data management policy is "to ensure

that research data is stored, retained, made accessible for use and reuse, and/or disposed of, according to legal, statutory, ethical and funding bodies' requirements." The Research Data Management Policy at Monash University is administered by the library. The Australian National Data Service (2015) suggested an outline for a research data management policy for Australian Universities/Institutions. They stated that the document 'is intended as a basic starting point for institutions that are intending to write, or update, their research data management policy. It is intended to be informative, not prescriptive.' As a result, the document can also be used by other research institutions in creating their own policies.

Tenopir et al. (2016) in a survey of directors of Association of European Research Libraries found out that almost all libraries collaborate with organisations within and outside the institutions in order to offer or develop policy related to research data services. They reported that Librarians collaborate with researchers, information technology centres, research offices, university archives, and legal offices.

Development of research data repository

Grace, Whyte & Rans (2015) stated that there is need to know where the repository will be housed within a research institution. At the University of East London, for example, the research committee provides general oversight for research data management and it includes representation from senior academics and services departments including the library, information technology (IT), and research development support. The authors added that there is need to choose a platform to use and the hosting service to run the repository. The University of Oxford used a joint approach to develop its research data management repository whereby the library focused on preservation and open access; IT services cared about infrastructure; and research services supporting funding body compliance. (Wilson, Fraser, Martinez-Urbe, Patrick, Akram & Mansoori: 2010). Throughout the process, Wilson et al. (2010) emphasised that there is need for data documentation, support and training, secure storage and linking data to the publication. Tenopir et al. (2016) stated that Librarians provide data storage facilities, tools for data analysis and virtual community support.

There is need to register the research data repository with the Registry of Research Data Repositories (re3data). This covers the research data repositories from different academic disciplines and it represents repositories for the permanent storage and access of datasets to researchers, funding bodies, publishers and scholarly institutions (Registry of Research Data Repositories 2016). As at 13 April 2016, the registry had listed 1500 repositories. It offers an avenue to select appropriate repositories for the storage and search of research data (Registry of Research Data Repositories 2016).

Research objectives

The study sought to:

1. Evaluate how research data are being managed in research institutions in Zimbabwe.
2. Assess the challenges that are faced in research data managed by research institutions in Zimbabwe.

Materials and methods

The study focussed on research institutions in Zimbabwe. A population of 25 research institutions were purposively selected that is, 16 institutions of higher learning and nine organisations that deal with research. The study was approved by the Research and Postgraduate Centre at the Bindura University of Science Education who provided the support letter to carry out the research. Informed consent was sought from the participants. This was indicated in the introductory part of the questionnaire and the telephone interview. Respondents consented to participate in the study. Personal identifiable data was anonymised for confidentiality purposes. Responses were received from 23 institutions which are 16 institutions of higher learning and seven research institutions. The population was composed of librarians, researchers, information officers and records managers and the institutions participating in the study were given the discretion to appoint the respondent. The breakdown of the respondents by profession were as follows: one researcher, two information officers, and 20 librarians. The study was conducted in mid-2016 using an online survey focusing on prevailing research data management practices. A link to the online questionnaire on SurveyMonkey was sent to all the participants. Fourteen responded to the online questionnaire on time and telephone interviews were done to follow up on nine participants who had failed to respond on time. The data that was collected using telephone interviews was entered manually into SurveyMonkey for easy analysis. The three respondents who indicated that they had research data repositories were phoned to seek further clarification on how they are managing research

data. One of the three respondents consented to a site visit which was done. The SurveyMonkey database was used to analyse the data which was then presented in tables and figures.

Results

The findings showed that researchers are responsible for managing their research data within their institutions as shown in **Figure 2**. One institution indicated that research data management is the responsibility of the research ethics committee and the other one the responsibility of the records managers. In terms of the availability of an institutional policy on research data management, 19 institutions indicated that they do not have a policy. Out of the five institutions that indicated that they have policies, only three have a research data repository in place; the oldest data repository was established in 2013 while the latest one was set up in 2016. For the three institutions that indicated that they have research data repositories, various professionals were involved in setting up the repository. The results showed that researchers and development partners were involved in setting up the research data repository in two research institutions. In the other institutions, librarians and information technology personnel were involved. However, records managers and research officers were not involved in setting up the repositories. Those institutions with research data repositories indicated that more than one department manage the repository that is the research office, information technology department, library and records centre. All the repositories are neither available on the Internet nor registered with the Registry of Research Data Repositories.

The respondents indicated that the reasons for keeping research data solely depend with the researchers in most institutions as shown by fourteen respondents in the study. Only one participant indicated that it is a requirement by publishers. The results showed that participants store the research data as text documents, spreadsheets, graphics, audio, databases, video, software applications and structured text. No research institution use software application source code or configuration data in storing research data. Seventeen institutions indicated that they do not archive research data which means that research data are destroyed soon after the data are analysed. Only one institution stated that they archive research data for a year or less as indicated in **Figure 3**.

Eleven respondents indicated that the choice to dispose the research data lies with the researcher, while six respondents pointed out that they erase the data from storage devices, four stated that they transfer research data to the records offices or archives, three shred the research data, three do permanent preservation and finally one indicated that there is no defined research data management at their institution. This is shown on **Figure 4**.

The findings showed that accessibility of the research data in eight institutions is provided for researchers and other partners. In five institutions, researchers decide who can access their research data, two participants stated that research data can be accessed by anyone, while one respondent indicated that access to research data depends on the data privacy classification. The challenges that are faced in research institutions in the management of research data are lack of guidelines on good practice, inadequate human resources, technological obsolescence, security, use of different vocabulary between librarians and researchers, inadequate financial resources, lack of institutional support and unavailability of software for research data

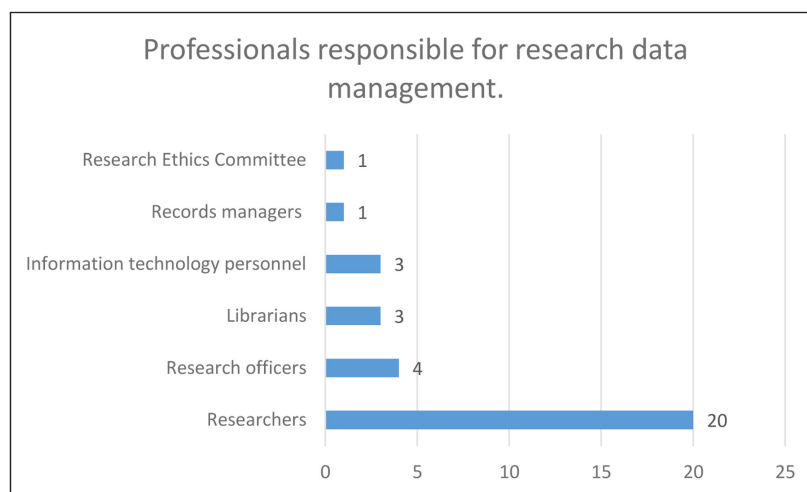


Figure 2: Professionals responsible for research data management.

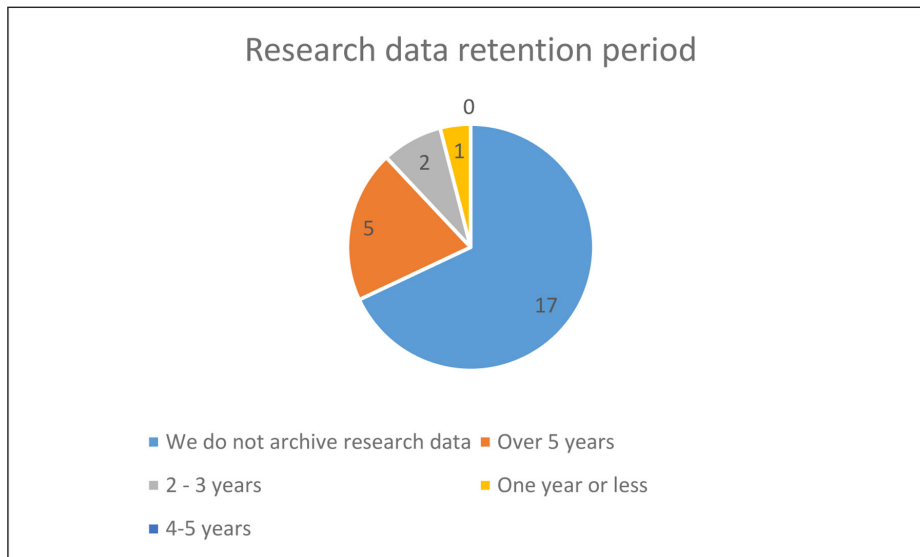


Figure 3: Research data retention period.

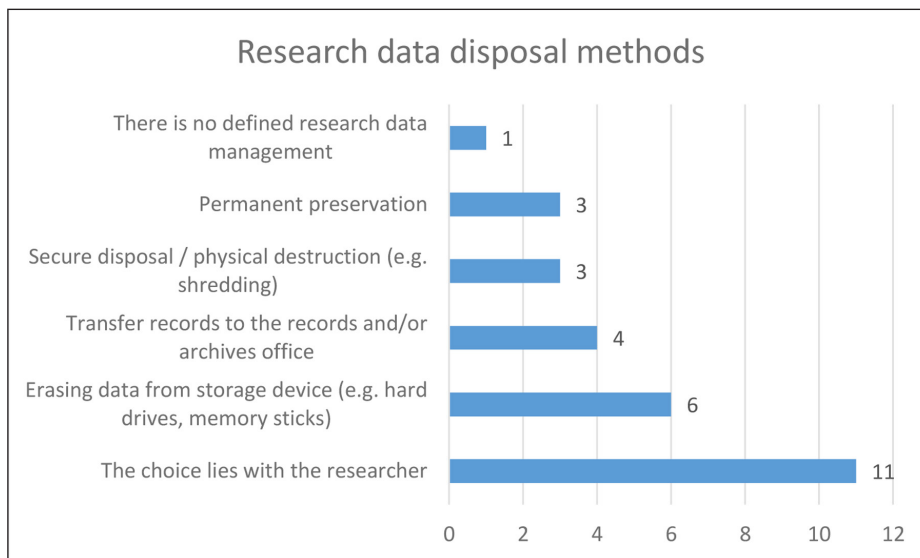


Figure 4: Research data disposal methods.

management at the institution. To mitigate these challenges, respondents indicated that they limit access to research data, conduct training, establish research policy, office and ethics committee, purchase online data management systems, and utilise the available equipment.

Discussion

The findings revealed that research data management is still a new concept to research institutions in Zimbabwe. Researchers are currently managing their own research data. This is in contrast with the findings from the studies by Tenopir et al. (2012, 2014, and 2016), Surkis and Read (2015) who reported that libraries in America and Europe have taken research data management as a key component of their duties and responsibilities. This implies that libraries in Zimbabwe have to take up research data management as part of their services.

Most of the research data were in textual format, spreadsheet format and graphical format that included images. Some of the research data were also reported to be in audio, video, database, structured text formats and software applications. The UK Data Archive (University of Essex 2017) acknowledges the availability of various formats of research data and the ones identified in this research were listed by the UK Data Archive.

It was found out that access to available research data is still a challenge considering that most of the research data is under the custody of the researchers. These findings differ from what Tenopir, et al. (2016) observed in European and American academic and research libraries whereby librarians were actively involved in research data management making access to research data easier. They also indicated that librarians at these institutions collaborated with researchers, information technology personnel, legal offices and university archives in research data management. This promoted seamless access to research data. However, only a few research institutions in Zimbabwe have started to manage research data and the data are only available on local area networks for access by a limited population.

There are a number of challenges being faced by research institutions relating to research data management key among them the lack of guidelines on good practice, inadequate human resources, technological obsolescence and insecure infrastructure. These challenges relating to inconsistencies and complexities in available data are in tandem with what is currently prevailing elsewhere (Corti, et al. 2011). Similarly, Kennan & Markauskaite (2015) pointed out that digital data management was affected by technological obsolescence, security, inadequate human resources, lack of guidelines on good practice, and lack of evidence about best infrastructures. Just as Surkis and Read (2015) noted, use of different vocabulary between librarians and researchers remains a challenge in research data management in Zimbabwe. This study confirms an assertion by Kennan & Markauskaite (2015) that librarians involved in research data management faced challenges to do with inadequate financial resources, absence of research data management policies and lack of support by institutional authorities and researchers have also negatively impacted on research data. The findings confirm that the challenges facing librarians involved in research data management across the world are the same despite the experience one would have with research data management.

In the absence of proper research data management, the challenges cited above were faced. Absence of proper research data management vis-à-vis the availability of research data in various formats cited above exacerbates the challenges faced in managing research data in Zimbabwe. However, a number of solutions were employed to solve these challenges. Training on research data management was introduced so as to improve librarians, researchers, research officers and records managers' competencies in handling research data. Literature consulted (Tenopir, et al. 2012, 2016) buttresses the importance of training librarians on research data management to ensure effective research data service delivery. In fact, Tenopir et al. (2016) show that even libraries with well-established research data services offered their staff opportunities for skills development for improved service delivery. Access to research data was also restricted to selected users to avoid unauthorised access, data loss and corruption. Some institutions established research ethics committees and research offices to steer good research data management practices and spearhead the crafting of research data management policies.

Conclusion and recommendations

Research data management is still a relatively new concept in Zimbabwe's research institutions as compared to other institutions in the developed countries. However, the concept is very important and librarians, research officers, records managers, information technology professionals and researchers need to explore the concept so as to effectively participate in good research data management practice. Authors recommend the establishment of research data repositories or the use of already established research data repositories that are registered with the Registry of Research Data Repositories to ensure that research data management standards are adhered to when doing research. There is also need to partner with international organisations such as DataCite and Research Data Alliance as these forums would assist Zimbabwean research institutions in managing research data professionally.

Additional File

The additional file for this article can be found as follows:

- **Research Data for Research Data Management in research institutions in Zimbabwe paper.** Summary of responses from SurveyMonkey before data analysis. DOI: <https://doi.org/10.5334/dsj-2017-031.s1>

Acknowledgements

The study was made possible with the block allocation grant that was provided by the Bindura University of Science Education Research and Postgraduate Centre. The authors are grateful for the support that was provided by the institution.

Competing Interests

The authors have no competing interests to declare.

References

- Australian National Data Service** 2015 *Australian Code for Responsible Research*, 15 June 2015. Available at: <https://www.nhmrc.gov.au/guidelines-publications/r39> [Last accessed 04 January 2016].
- Buys, C M and Shaw, P L** 2015 Data Management Practices Across an Institution: Survey and Report. *Journal of Librarianship and Scholarly Communication*, 3(2): eP1225. DOI: <https://doi.org/10.7710/2162-3309.1225>
- Corti, L, Van den Eynden, V, Bishop, L and Woollard, M** 2011 *Managing and Sharing Research Data: A Guide to Good Practice*. UK Data Archive. University of Essex.
- Fary, M and Owen, K** 2013 *Developing an institutional research data management plan service*, 10 June 2013. Available at: EDUCAUSE website: <https://net.educause.edu/ir/library/pdf/ACTI1301.pdf> [Last accessed 20 November 2016].
- Grace, S, Whyte, A and Rans, J** 2015 Building UEL's Research Data Repository Capabilities with EPrint. *DCC RDM Services case studies*. Edinburgh: Digital Curation Centre. Available at: <http://www.dcc.ac.uk/resources/case-studies> [Last accessed 19 November 2016].
- Kennan, M A and Markauskaite, L** 2015 Research data management practices: a snapshot in time. *International Journal of Digital Curation*, 10(2): 69–95. DOI: <https://doi.org/10.2218/ijdc.v10i2.329>
- Monash University** 2013 *Research Data Management Policy*. Available at: <http://policy.monash.edu.au/policy-bank/academic/research/research-data-management-policy.html> [Last accessed 24 June 2016].
- National University of Singapore** 2016 *Research Data management policy*. Available at: <http://www.nus.edu.sg/dpr/InfoForResearchers/research-dmp.html> [Last accessed on 24 June 2016].
- Registry of Research Data Repositories** 2016 *re3data.org Reaches a Milestone and Begins Offering Badges*. Available at: <http://www.re3data.org/> [Last accessed 01 July 2016].
- Rice, R** 2009 *DISC-UK DataShare project: Final report*. Retrieved from Jisc repository: <http://ie-repository.jisc.ac.uk/336/1/DataSharefinalreport.pdf> [Last accessed 24 June 2016].
- Surkis, A and Read, K** 2015 Research data management. *Journal of the Medical Library Association: JMLA*, 103(3): 154–156. DOI: <https://doi.org/10.3163/1536-5050.103.3.011>
- Tenopir, C, Hughes, D, Allard, S, Fram, M, Birch, B, Baird, L, Sandusky, R, Langseth, M and Lundeen, A** 2015 Research Data Services in Academic Libraries: Data Intensive Roles for the Future? *Journal of eScience Librarianship*, 4(2): e1085 [Last accessed 04 April 2017]. DOI: <https://doi.org/10.7191/jeslib.2015.1085>
- Tenopir, C, Sandusky, R J, Allard, S and Birch, B** 2014 Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*, 36(2): 84–90. DOI: <https://doi.org/10.1016/j.lisr.2013.11.003>
- Tenopir, C, Talja, S, Horstmann, W, Late, E, Hughes, D, Pollock, D, Schmidt, B, Baird, L, Sandusky, R J and Allard, S** 2016 Research Data Services in European Academic Research Libraries. Submitted to *LIBER QUARTERLY*. Available at: http://libereurope.eu/wpcontent/uploads/2016/10/LIBER_RDM_SURVEY_FINAL_MANUSCRIPT.pdf [Last accessed 04 April 2017].
- University of Essex** 2017 *File formats table*. Available at: <http://www.data-archive.ac.uk/create-manage/format/formats-table> [Last accessed 7 April 2017].
- University of Leeds** 2016 *University of Leeds Research Data Management Policy*. Available at: <https://library.leeds.ac.uk/research-data-policies> [Last accessed 24 June 2016].
- University of Manchester** 2016 *The University of Manchester Research Data Management Policy*. Available at: <http://www.library.manchester.ac.uk/services-and-support/staff/research/services/research-data-management/> [Last accessed 24 June 2016].
- Whitmire, A L, Boock, M and Sutton, S C** 2015 Variability in academic research data management practices: implications for data services development from a faculty survey. *Program: electronic library and information systems*, 49(4): 382–407. URL: <http://hdl.handle.net/1957/57240>. DOI: <https://doi.org/10.1108/PROG-02-2015-0017>
- Whyte, A and Tedds, J** 2011 'Making the Case for Research Data Management'. *DCC Briefing Papers*. Edinburgh: Digital Curation Centre. Available at: <http://www.dcc.ac.uk/resources/briefing-papers> [Last accessed 23 June 2016].
- Wilson, F, Martinez-Uribe, P and Akram, M** 2010 *Developing Infrastructure for Research Data Management at the University of Oxford*. Available at: <http://www.ariadne.ac.uk/issue65/wilson-et-al> [Last accessed 01 July 2016].

How to cite this article: 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>

Submitted: 06 January 2017 **Accepted:** 09 June 2017 **Published:** 16 June 2017

Copyright: © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

]u[*Data Science Journal* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 