A Proposed Framework for Research Data Management Services in Research Institutions in Zimbabwe

Josiline Phiri Chigwada
Bindura University of Science Education, Zimbabwe

Thembelihle Hwalima
Lupane State University, Zimbabwe

Nancy Kwangwa
University of Zimbabwe, Zimbabwe

Citation

ABSTRACT
The chapter documents the proposed framework for the establishment of research data management services in research institutions in Zimbabwe. It has been indicated that there are no formal research data management services that are taking place in Zimbabwe since researchers are managing their own data. It is against such a background that a literature review was done to understand how research institutions in other countries are engaging in research data services. E-mails were sent to the pioneers of research data services. It was discovered that there are challenges that are faced when establishing research data management services and it is important to incorporate all the stakeholders from the planning stage. The framework consists of strategies, policies, guidelines, processes, technologies and services.

Keywords: Research Institutions – Zimbabwe, Research Data Management – Zimbabwe, Research Data Services – Zimbabwe, Research Data Repository – Zimbabwe.

INTRODUCTION
The chapter documents a proposed framework that can be used in Zimbabwe to develop research data management (RDM) services in research institutions. Published research (Chigwada, Chiparausha and Kasiroori 2017) showed that RDM is a new concept in Zimbabwe and research institutions are working towards the establishment of RDM services. This chapter would assist in unpacking the activities and processes that should be done in establishing RDM
services by proposing a framework that can be used. The objectives that are addressed by this chapter are:

1. To identify the strategies that can be used to establish RDM services in Zimbabwe.
2. To identify the stakeholders that can be involved in establishing RDM services in Zimbabwe.
3. To assess the challenges faced when establishing RDM services in Zimbabwe.
4. To propose a framework that can be used to establish RDM services in Zimbabwe.

A literature survey and document analysis was done to get the information from those with experience in the establishment of RDM services. Emails were sent to the champion persons to gather as much detail as possible on how research institutions in Zimbabwe can establish RDM services.

**WHAT IS RESEARCH DATA MANAGEMENT?**

Research data management (RDM) is regarded as the organisation and description of data from its entry to the research cycle to dissemination and archiving of the results (Whyte and Tedds 2011). Pinfield, Cox, and Smith (2014) point out that RDM addresses a number of information needs and it is driven by the need to provide immediate storage facilities; the requirements to ensure the security of data; the need for medium and long term preservation of data; compliance with the requirements and policies of other agencies such as funders; the quality of the research activity and research data itself; the need to share data and make it open access; and the jurisdiction to be involved in RDM as well and the involvement of other players in offering RDM services. In view of the above, there were various initiatives that were done in Africa to establish research data management services and some institutions are now archiving research data locally, nationally, regionally, and internationally.

**What is a research data repository?**

Research data repository is a database infrastructure that is set up to manage, share, access and archive researchers’ datasets (Uzwyshn 2016). It can be institutional, national, regional or international and can be specialised or general. The main objective of establishing a research data repository would be to share data which would allow the examination, proof, review, transparency, validation of researchers’ results by other experts at the same time allowing simultaneous access by many researchers. Uzwyshn (2016) indicates that libraries should seriously consider partnerships with national and international organisations working towards the development of research data repositories in order to provide the required information services and infrastructure. In order to manage research data effectively, there is need to comply with the funders and institutional policies. Jones, Pryor and White (2013) wrote a guide for those who work in higher education institutions.

In the event that the institution does not have a data repository but wants to promote RDM services amongst its researchers, there are other avenues that can be utilised to ensure that research data is managed. In some cases, the funder might state the data repository that should be utilised by researchers when they are applying for funds. If the funder does not have a preferred repository, researchers can use discipline specific repositories such as UK Data
Archive for social sciences and humanities, arXiv for mathematics and physical science, GEO for genomic datasets, and some that are provided by PLOS journals or Scientific Data. There are general purpose repositories such as Zenodo and Mendeley data. In the absence of research data services, research institutions in Zimbabwe can encourage researchers to visit the registry of research data repositories (re3data) to find the best repository to host their data.

**Initiatives to establish research data management services in Africa**

Literature indicates that efforts to establish research data management services are being done in South Africa, Botswana and Zimbabwe (Chiware and Becker 2018, Peters and Hahnel 2017). University of Cape Town, Cape Peninsula University of Technology and other institutions of higher learning in the Western Cape took a proactive approach in establishing RDM services taking advantage of the open access movement. In South Africa, there is the African Open Science Platform (AOSP) which is managed by the Academy of Science of South Africa (ASSAf), and funded through the South Africa National Research Foundation in the Department of Science and Technology. Direction for the provision of funding is given by the International Council for Science (ICSU) Committee on Data for Science and Technology (CODATA). The work of AOSP benefits the African content by looking into progress in terms of open science policy, ICT infrastructure, data sharing, collaboration, and capacity building. In South Africa, the South African National Research Foundation (NRF) policy states that research funded by public funds should be archived in trusted repositories. The Science Granting Councils Initiative (SGCI) in Sub-Saharan Africa also works with 15 science granting councils in Sub-Saharan Africa to strengthen research in the region. The Data Intensive Research in South Africa (DIRISA) was created to enable universities and research institutions to access facilities for data intensive projects. A research done by Chiware and Becker (2018) indicated that there were no responses from Namibia, Malawi, Zambia, Lesotho and Swaziland in terms of reported activities in relation to RDM in university and research libraries or at national level again pointing to the need for clear national guidelines on how institutions should plan and develop this important area for research support.

**Role of the library in RDM**

The role of the library and librarians has been evolving in the 21st century because of the rising need to keep the library aligned to current trends in the field. This has found the librarians having a more outgoing approach in their service delivery and having to support research, that is, research data management, bibliometrics, research support so as to remain relevant to their users. To fully understand the thrust of RDM, it is important to relate to the definition of research data management by Whyte and Tedds (2011). It is easy to understand the increasing prominence and importance of RDM but this has come with a diverse or rather transformed the role that librarians play within their research institutions. Cox and Pinfield (2014) observe that RDM consists of a number of different activities and processes associated with the data lifecycle, involving the design and creation of data, storage, security, preservation, retrieval, sharing, and reuse, all taking into account technical capabilities, ethical considerations, legal issues and governance frameworks.
The debate about the library’s involvement in data management has been conducted at both strategic and operational levels. Lewis (2010) argues that data from academic research projects represents an integral part of the global research knowledge base, and so managing it should be a natural extension of the library’s current role in providing access to the published part of that knowledge base, while also noting the scale of the challenge in terms of infrastructure, skills and culture. Therefore, librarians are encompassing new RDM responsibilities which involve being asked to do things which at present are beyond their usual expertise (Chiware & Mathe, 2015). This has been posited by Nitecki & Davis (2017) who points out that librarians given their long experience with information organisation and documentation are becoming more involved in the development of principles and best practices for managing digital data for long term use. This has been exacerbated by some research institutions who now have a new role of a data librarian which is emerging, but there can be ambiguity over the meaning of this designation: does it embody a librarian with a particular skill set, or is it more of a description of new duties? In the case of the latter, a data librarian may be someone who thrusts into a new role and must acquire a range of new skills on the job as posited by Pinfield (2014).

**Librarians’ response to RDM services**
Hey and Hey (2006) suggest that if librarians can respond effectively to the challenge by engaging the e-Science revolution which will put libraries and repositories center stage in the development of the next generation research infrastructure, then RDM will be very prevalent. Services and tools often managed by libraries are now becoming more widespread. It is clear that at certain institutions librarians are playing significant and growing roles in the research data management process. Recognising the potential that librarians can offer, and the need to develop skills, there are a growing number of online training courses being developed at national and institutional levels such as Data Scientist Training for Librarians, DIY RDM Training Kit for Librarians, Data Intelligence for Librarians (DCC 2017). It is evident also that much of the work is being done collaboratively between the library and the institutional Information Technology departments, the latter dealing with the technical aspects of handling petabytes of data, the former developing the organisational and service aspects of data management (Chiware & Mathe, 2015).

As different approaches to RDM are developing in research institutions different stakeholders have become involved, including support services staff as well as the faculty. University libraries have moved into this space and are increasingly seen as major contributors to RDM activity in general and in the design of research data services (RDS) in particular (Pinfield, Cox & Smith, 2014). Therefore, from this assertion, it is pivotal to train librarians so as to fully align them to this changing spectrum of RDM.

Swan and Brown (2008) see data management as a strategic issue for libraries and librarians and point out that the role of the library in data-intensive research is important and a strategic repositioning of the library with respect to research support is now appropriate. They add that three main potential roles for the library could be increasing data-awareness amongst researchers; providing archiving and preservation services for data within the institution
through institutional repositories; and developing a new professional strand of practice in the form of data librarianship. Figure 1 summarises the role of the librarian in RDM services.

![Figure 1: The role of the librarian in RDM services](image)

From the above diagram, RDM challenges librarians to become self-motivated, research-grounded, intellectual entrepreneurs, and specifically to become proactive designers of services that enable productive knowledge workers; partner in knowledge-generating activities bringing understanding of the information and data landscape and its tools for discovery and utilization; share project management roles to increase research team productivity; and be change agents that build evidence to monitor efficiencies and gauge impact (Nitecki & Davis, 2017).

**ESTABLISHING RESEARCH DATA MANAGEMENT SERVICES**

Shen and Varvel (2013) indicate that human, financial and technological resources should be available for the RDM services to be successful. When establishing a research data repository, institutions should develop a research data management infrastructure profile (Davidson 2015). The profile would help to provide a better understanding of the infrastructure that is available to avoid duplication of effort by providing an inventory of the existing services. Davidson (2015) provided the following components as the desirable scope of RDM infrastructure provision:

- Means of raising staff awareness of funders’ research data requirements
- Research data policy
- Strategy or implementation plan for research data services
- RDM advice and support services
- Active data storage
- Persistent identification for datasets
- Data register or catalogue
- Data access procedures
- Secure data access
· Institutional publications repository (if it includes research data or metadata)
· Data repository for longer term access and preservation

There is need to determine the software that would be used for archiving data. Several possibilities can be used whereby research data repository software specifically created for data or other digital library software can be used. According to Uzwyshn (2016) software specifically created for data are Dataverse, HUBzero, and Chronopolis while those for general digital libraries are DSpace, Fedora, and Hydra. The institution can install the software on its servers or it can be hosted by other organisations. The size of the data determines the information, communication technology (ICT) infrastructure that would be chosen for establishing a research data repository. The data sizes can be divided into three, that is, small/medium, large, and very large (Uzwyshn 2016). Small-to-medium datasets can be stored on the researcher’s computer and can be uploaded by a researcher, emailed, or transferred through university network drives to a server or the cloud. Data from medium-to-large projects require special back-end storage systems while very large projects can be preserved and archived in consortial or national research data repositories. In Zimbabwe, research institutions should choose the software that matches their needs together with the available resources. There should be IT experts that should be responsible for choosing and maintaining the software.

Human and financial resources should be available for the RDM services to flourish. The RDM experts should be able to offer guidance on how best to tackle such a project. They should also be available to train the stakeholders involved as a way of capacity building. Financial resources and costs include the price for setting up and managing a research data repository and payment of expertise needed to run the repository. Further costs would be incurred depending on the duration of keeping the data and the preservation requirements.

Reliable ICT infrastructure is fundamental to achieve the aim of archiving research data both at institutional and national level. In Africa, the research intensive institutions are connected through the National Research and Education Network (NREN) and some institutions in Kenya, South Africa, Uganda and Zambia are running data intensive applications and sharing of high end computing assets (Adam 2016). This shows that Zimbabwe is not yet in the picture and should work towards the development of good ICT infrastructure to enable the development of RDM services.

STAKEHOLDERS IN ESTABLISHING RESEARCH DATA MANAGEMENT SERVICES
The stakeholders that should be involved in establishing RDM services are librarians, records managers, research institutions, research administration, researchers, library schools, IT specialists, funders, government and other RDM providers (Ingram 2016). It has been stated that the roles and responsibilities of establishing RDM services is shared among university management, support and administrative services, and researchers (Jones, Pryor and Whyte 2013, Cox and Verbaan 2016, Latham 2017). It has been reported that librarians should work with IT staff and research administrators to offer RDM services (Latham 2017). As a result,
when establishing RDM services a minimum of the following offices should be involved: the research office, library, IT, researchers and the ethics committee. For this chapter, the stakeholders are divided into three i.e. management, support and administrative services, and researchers. Management would define expectations, support staff would deliver the services while the researchers would create and use research data.

**Management**

Management should ensure that the project for RDM services is feasible by providing support and making sure that the infrastructure is available and working. According to Jones, Pryor and Whyte (2013), the roles of senior management are to provide a champion person at the level of a Pro Vice-Chancellor Academic to chair the working group; ensure that all the stakeholders are represented in the working group including the library, research office, IT office and the researchers; approve proposals and endorse RDM budgets; and policy ratification and implementation. In addition to providing policies on how research data can be managed, research institutions must ensure that researchers have educational and support services in line with research data management as a way of encouraging data sharing (Tenopir, Birch, & Allard, 2012; Tenopir, Sandusky, Allard, & Birch 2013).

**Researchers**

Conrad, Shorish, Whitmire, and Hswe (2017) state that some researchers did not receive formal training on research data management and as a result are not able to personally manage their own data. Researchers are one of the major stakeholders in the development of research data services. These are crucial since they are the major provider and consumer of research data. Researchers can be from universities, think tanks, institutes, organisations or companies which have dedicated research and development departments. They collect, analyse, find and reuse data. They collect data from interviews or surveys which they deposit in the research data repository. Whyte and Wilson (2010) point out that researchers are data creators and they are responsible for providing enough information that would enable other researchers to assess the quality of that data and whether it complies with the ethics of the subject. The research would also provide information that would indicate the users who can access the data including any access requirements or constraints. The research data would be provided in the recommended formats and the metadata requested by the repository is also provided by the researcher.

**Support and administrative staff**

The support staff is composed of the library, information technology, records management, research administration, and other outside RDM service providers. The research office plays a key role as the link between researchers, management and the funders. Therefore, they should work together to establish the team that would be in the working group. The administrative people should do the groundwork so that they know the policy requirements and have a bigger role to play in the development and implementation of proposals, plans and budgets needed for RDM services. As a result, capacity building is key to these people and should advocate for the establishment of RDM services at their research institutions.
Data curators can be the library or the research office depending on what has been agreed during the planning stage. It should be clear on its mission of digital archiving to ensure that digital objects are selected according to the policy. The repository should be able to house data in all the formats be it audio, video, audio visual, tif, gif, tsv or scanned documents. The curator has a duty to ensure that the repository is in compliance with the legal regulations and the authenticity and integrity of the data. This is because the data curator would assume responsibility of ensuring that the data is accessible and available to the approved users. The major role of the data curator is to plan for long term preservation of the data.

Conrad, Shorish, Whitmire, and Hswe (2017) observe that libraries and records centres are given the responsibility to assist research in managing and curating research data despite the limited training opportunities. Librarians are now partners in the research data life cycle and are now actively involved in curating, advising, and preserving research data. Libraries therefore are assuming additional duties of creating awareness amongst researchers, archive and preserve data, and to train researchers on best practices. Peters and Dryden (2011) found out that the researchers are mostly interested in getting assistance on grant proposal support such as the process of data management planning, locating research data and other data related services, publication support, and dealing with the specific data during data collection. In order to successfully assist the researchers, librarians must understand the researchers’ needs so that the services provided would meet the research data requirements by these researchers. Librarians contribute to research data management from the proposal writing stage where they are supposed to assist in creating the data management plan. During the project start up stage, librarians help develop the data model and identify appropriate standards and recommended tools and resources for organising and sharing data complying with the requirements. After the end of the project, librarians support efforts to archive data by placing it in repositories for preservation. They also assist researchers to locate already existing public data that answers their research questions.

In Zimbabwe, there are various stakeholders that can be involved in the development of RDM services. These institutions play a major role in research and development activities within the country and would be in a better position to work towards the RDM services as pointed out in the literature review. Figure two is a diagrammatic indications of those who should be involved in RDM services.
Figure 2: RDM Stakeholders in Zimbabwe

The diagram indicates that all the institutions that are involved in research should take part in the development of RDM services, both those in the public and the private sector. Various ministries including the Ministry of Information Communication Technology (ICT) and Cyber security, Ministry of Higher and Tertiary Education, Science and Technology Development, and Ministry of Information, Media and Broadcasting Services should be actively involved since they are responsible for providing the infrastructure and the resources that are needed for the development of RDM services. Research Council of Zimbabwe is the overall board that is responsible for all the research activities and funding in the country. It would assist in ensuring that the research data is available for archiving.

There is a lot of research that is going on in institutions of higher learning as evidenced by the advent of institutional repositories to archive the research output from these institutions. These can develop RDM services at an institutional level if they have the resources. Another option is to develop national RDM services so that there is a large pool of resources. Within the institutions of higher learning, there are already established boards that can be used to develop
national RDM services. The Research Board deals with all the research activities of a college or university and the expertise can be used to develop policies and guidelines that would enable the archiving of research data in Zimbabwe. There are some research and development departments within the private sector and other research institutes such as Scientific and Industrial Research and Development Centre (SIRDC) which would play a very crucial role in the development and management of RDM services in Zimbabwe. Figure 2 indicated that the development of RDM services is a team effort which require the joint effort of both the private and public sector.

STRATEGIES FOR CREATING AWARENESS TO RESEARCHERS
As the creators and users of research data, researcher engagement is crucial in the development of RDM services (Jones, Pryor and Whyte, 2013; Wilson and Jeffrey 2013; Buchhorn and McNamara 2010). Furthermore, Sompel et al. (2004) aptly reminded “Like any technology, success will depend not only on technical soundness but on the willingness of the participants in the system that is publishers, scholars, academic institutions, funding institutions, and others to adopt new tools and develop new organisational models on top of them.” It is therefore imperative to have strategies in place to raise RDM awareness to researchers. The strategies include advocacy and training, use of social media tools and embedding RDM in existing research support activities.

Advocacy and training programmes have proven to be on top of the list in promoting new initiatives in the library and information science fraternity. In the recent past when institutional repositories and electronic resources were introduced in academic libraries in Zimbabwe, advocacy and training were central in raising awareness among researchers. A study by Chigwada et al. (2017) on RDM services in research institutions in Zimbabwe revealed that research data management is still relatively a new concept compared to other institutions in the developed countries. It is therefore, pertinent to have RDM advocacy and training programmes in place to bring researchers on board. The training programmes should not target researchers only but all the stakeholders in RDM since data management is a relatively new concept in the country.

The use of social media in promotional and advocacy initiatives cannot be overemphasised. Social media tools provide an opportunity to reach out to a wider audience at a very low cost. There are several social media tools that can be utilized in raising RDM awareness to researchers. These include Researchgate, Academia.edu, LinkedIn, Mendley, Facebook, YouTube and Twitter. Libraries in Zimbabwe have made commendable progress in the use of social media to promote their collections and they also act as an online noticeboard for announcements. Some researchers utilize platforms such as Researchgate and Academia.edu to disseminate their research findings. Librarians should maximize the opportunity to make use of these platforms to reach out to researchers.

Libraries need to provide support for the complete research cycle, and therefore, they need to analyse what researchers require to manage their data from creation or compilation to archive and preservation. In this context, many university libraries are seeking to enhance their support
for research (Digital Curation Centre, 2016). Hence they are rethinking the roles of their collection-based services; changing the roles of liaison librarians; and developing new services for researchers including advice on scholarly communications and open access, bibliometrics services, research data management and library-led publishing services. RDM services can be embedded in existing research support services which are currently being offered by libraries. There is commendable progress in research support services in academic libraries. It should be easier to incorporate RDM in research methods courses and information literacy courses given their popularity in the research landscape.

**CHALLENGES IN ESTABLISHING RESEARCH DATA MANAGEMENT SERVICES**

Establishing a research data repository is a complex issue involving multiple activities carried out by various actors addressing a range of drivers and influenced by a large set of factors (Pinfield et al., 2014). Given this complexity, stating a research data repository is bound to face some challenges that include, skills, costs, policies, infrastructure and sustainability. This section will provide a highlight of some of the challenges which are likely to be faced when setting up a research data repository.

The success of RDM programmes partly depends on the skills and knowledge of the people involved. Henderson and Knott (2014) observe that the introduction and success of RDM services in academic libraries calls for the need to hire new staff or re-skilling and up skilling of librarians to take up new roles and responsibilities. Many librarians especially those who trained in Zimbabwe did not have any courses or formal training in data management because the subject was not covered in the Library and Information Science curriculum. A study by Nhendodzashe (2017) on the feasibility of offering RDM services at the University of Zimbabwe revealed that only a handful of librarians had knowledge of RDM. This lack of knowledge and skills on RDM presents a hurdle in the establishment of research data repositories in Zimbabwe. Interestingly, there is now a course in graduate Library and Information Science programme being offered by the National University of Science and Technology. In addition to this there are other free online RDM courses such as MANTRA and free reference material provided by the Digital Curation Centre.

A research data repository highly depends on the content and willingness by researchers to submit their data in the data repository. Research has shown that there are mixed views from researchers regarding their willingness to submit research data to data repositories (Wilson et al., 2010; Buchhorn and McNamara 2010). Keil (2014) in a study on research data needs from academic libraries taking a perspective of the faculty researchers revealed that sharing data can be particularly unnerving to scientists who may perceive a loss of a competitive edge for their next follow-up manuscript or grant proposal. Similarly, Kennan and Markauskaite (2015) conducted a study on research data management and sharing practices of academics at ten universities in New South Wales, Australia. Researchers were asked whether, once they had collected their own data, they would be willing to share them outside of their research team or project. While more than half (54.7%) indicated that they would not be prepared to share any
of their data, 36.4% indicated they would be prepared to share some of their data and 8.9% indicated that they would be prepared to share most of their data. On the same note, a study by Buchhormand and McNamara (2010) reveals that researchers are often reluctant to share their data. Data is viewed as a personal good, created by researchers and to be exploited by them. While many researchers feel data should be available to the research community there is a very strong and unanimous view that researchers should be able to exclusively exploit ‘their’ data for a period of time before it becomes available to others. On the contrary, a study by Kennan and Markauskaite (2015) in South Africa showed that the majority 73% of the researchers were willing to submit their research data to data repositories. It is discouraging to note that in institutions that have established institutional repositories, some researchers still resist to submit their research articles for archiving, what more research data?

Another challenge that can be encountered in establishing a research data repository is on metadata interoperability. Repositories often include metadata coming from a range of disciplines, each of which have different citation traditions and different emphases on the type of information they share (Chapman et al., 2009). Given the diversity in which metadata originate from, it may be difficult to enforce unwavering use of metadata and entry of metadata values. Metadata can be sparse or lack important contextual information particularly when that context is held at a collection level. The breadth and depth of disciplines across an academic institution means that use of controlled subject terms is possible at only the highest levels.

The long term preservation of research data in a repository should be sustainable. A study by Chapman et al (2009) showed that responsibility for the long-term management of research data is ill-defined. This is a challenge especially for researchers who conduct research for different organisations because the responsibility for research data management will be scattered. Gold (2007) argues that it is fair to say there is still a substantial amount of uncertainty about the roles libraries can play in scientific data management, reflecting an environment of ongoing experimentation and negotiation. In addition to responsibilities, most institutions lack clear guidelines and policies that govern the operations of the data repository leading to a situation where obligations for data retention may not always be met and longer-term access may not be possible. In the Zimbabwean context, most institutions that have established institutional repositories are operating without policies which makes their operations difficult.

The establishment of a research data repository calls for a conducive legal environment. Understanding the legal obligations surrounding research data management is crucial as it guides the preservation and access of research data. Fitzgerald and Pappalardo (2007) assert that it is apparent that to achieve seamless access to data it is necessary not only to adopt appropriate technical standards, practices and architecture, but also to develop legal frameworks that facilitate access to and use of research data, whether on an inter-organisational basis or across national borders. The UK Data Archive (2015) emphasises that before embarking on a RDM project, it is imperative to know your legal, ethical and other obligations regarding research data, towards research participants, colleagues, research funders and institutions. RDM legal obligations broadly include intellectual property rights which
encompass trademarks, design rights, patents and copyright. Trade secrets protect confidential business information.

Incurring costs when setting up a research data repository is inevitable. Hole et al. (2010) assert that predicting the costs of long-term digital collection, storage, preservation of and access to research data is a crucial yet complex task for even the largest repositories and institutions. Costs are incurred in acquiring hardware and software for the repository, hiring staff and staff training costs and costs associated with the maintenance of the repository. Most Zimbabwean institutions are hit by economic challenges to the extent that they are failing to recruit staff with the requisite skills to run data repositories. Additionally, foreign currency shortages is another stumbling block as some software programmes require to be imported from other countries.

Davidson et al. (2014) assert that preserving research data for the long term has a cost; although the infrastructure itself is costly, more significant is the cost associated with human resources, such as personnel to manage and maintain the archive. Storage costs for digital data are decreasing, but costs related to storage, such as power, data curation and annotation, and personnel, are not decreasing (Berman 2008 as cited by Strasser, 2014). Increasing amounts of digital data, and the need to comply with regulations regarding backup and monitoring, emphasise that these costs should not be underestimated or overlooked. Strasser (2014) posits that short-term costs for data preservation are primarily those related to storing data rather than archiving it. This may include software or hardware for backing up data or personnel costs for managing and organising data for storage. Longer-term preservation costs are associated with archives. Many repositories and archives use annual pricing schemes for a set amount of data; this situation is changing, however, to better meet the needs of researchers whose costs are intertwined with grant cycles.

Another challenge that is likely to be encountered in establishing a data repository is poor technological infrastructure most importantly the ability to integrate the repository with existing systems. Most institutions in Zimbabwe rely on an outside vendor to update the programming of their information technology systems. The approach is costly and not sustainable given the financial challenges highlighted above. This is coupled by lack of institutional support especially if the project is not laid down well and advocated for. In some institutions, the relationship of libraries with other departments is not good. This can be as a result of the perceived role of the librarians in the research life cycle where librarians are not seen as partners. That negative attitude towards the librarians’ role in research data management activities would mean that researchers would not look to the library for research support.

**SOLUTIONS AND RECOMMENDATIONS**
To counter the challenges highlighted in the preceding section, there is need for training on data repositories. There should be training and mentorship programmes to upskill the people who will be responsible for research data management in research institutions. Zimbabwe can take advantage of the highlighted courses so that expects can be created. The use of free and open source software by organisations that plan to establish data repositories cannot be
overruled. This helps to reduce setting up and operational costs. Software programmes such as Fedora, DSpace, DataVerse can be used. There is need to develop clear guidelines for operational and administrative responsibilities for the data repository. Research institutions must establish strategic partnerships with other institutions that have well established data repositories. Universities in Western Cape in South Africa took advantage of experts in the developed areas to train the fieldwork. There is need to engage in fundraising activities for the setting up of data repositories. Research institutions should develop multidisciplinary metadata standards. Also to be addressed as part of policy is copyright which was stated by Terroir (2016) as undeveloped, inefficient, expensive and unenforced.

The authors recommend that research institutions in Zimbabwe should develop an RDM workgroup to help establish and implement RDM policies and services. In the process, the workgroup should be able to match best practices with practical realities. The workgroup should engage and advocate for research data management both at institutional and national levels with stakeholder communities. At an institutional level, if it is a university, the working group can take advantage of the already established committees such as the senate, library committee, research board, faculty boards and research units such as the research office and library to advocate for RDM services. At a national level, there are various government bodies that should be involved such as Universities, Research institutes, Scientific and Industrial Research and Development Centre (SIRDC), Research Council of Zimbabwe, Ministry of Higher and Tertiary Education Science and Technology Development, Ministry of ICT and Cyber Security, Ministry of Information, Media and Broadcasting Services, and Zimbabwe Universities’ Vice-Chancellors’ Association.

During implementation, the research institutions can start small by piloting one department and then continue collaborating with other departments through capacity building. When other researchers register strategic interest in managing their research data, they can be roped in thereby increasing the number of datasets in the repository. There is need for capacity building to skill the librarians and other stakeholders so that they are able to assist lecturers throughout the research data life cycle. This is done to ensure that the librarians are supported for them to acquire relevant expertise that is needed to deliver RDM services to the researchers. Shipman, Martin, Kaplan, and Albright (2018) emphasised the importance of training librarians.

Research institutions in Zimbabwe can build a community of practice for research data management. This would be composed of researchers, institutions, and a network of RDM experts. Researchers would be from academic institutions, government and other research institutes. Institutions would provide liaison librarians, IT specialists, University research offices, ethics boards, and other stakeholders. This would help to facilitate and provide leadership in the development of RDM infrastructure. Institutions should also develop their RDM policies and funding agencies should ensure that researchers who receive public funds should manage and share publicly their research data. Library schools should be actively involved in RDM services since they are responsible for the curriculum development of the programmes that are offered in colleges and universities. Involving them would ensure that students are taught how to manage research data. There should be international collaborations
both for training and infrastructure development. Research institutions in Zimbabwe should also register their presence on the registry of open data repositories when they establish research data repositories.

FRAMEWORK FOR ESTABLISHING RESEARCH DATA MANAGEMENT SERVICES

The digital revolution has made it easier to store, share and re-use data, and scientific research data is almost universally created and collected in digital form. Data sharing increases the potential return on the large investments made into research by reducing costly data duplication. But data sharing requires data to be stored efficiently, maintained and preserved for re-use, discovered by secondary users, and used with confidence in its authenticity and integrity therefore the Library & Librarians have to align themselves with this fundamental change. Repositories are combinations of software and hardware that together provide a set of services that manage and disseminate digital works (Lynch, 2003; Jantz and Wilson, 2008) and in which authors are encouraged to deposit copies of their own work. The term commonly used for depositing ones’ work is to self-archive (Xia and Sun, 2007). Repositories come in two main types, institutional and disciplinary. There are many disciplinary repositories, some of which are very successful in covering, preserving and making accessible the literature of their discipline.

In an effort to come up with a viable framework for data repositories, there is need to borrow ideas from other frameworks that have been used and then be able to relate to an institution’s current context or setting. Kennan (2011) in relation to an institutional repository program identifies some components that can make up a tangible framework and these are stated in figure 3.

![Figure 3: Proposed framework for research data management in Zimbabwe](image)
Strategies: Research institutions in Zimbabwe must decide whether they can develop institutional or national research data services. They have to define the vision for research data management within the institution and how it relates to the institutional mission and priorities so that it can assist in meeting the institutional goals. There is need to outline major developmental goals and principles which inform research data management activities both at institutional and national level. Research institutions must understand their current position and where they want to be so that they would be in a position to define their strategy.

Policies: There is need to specify how the strategies are to be operationalised through regular procedures and how RDM would be incorporated in other policies such as Open Access and Intellectual Property. These policies should complement each other and therefore cannot be considered in isolation. In Zimbabwe, the research institutions deal with copyright issues in relation to the use of institutional repositories that were established in institutions of higher learning.

Guidelines: The guidelines would provide detail on how the policies will be implemented. These would be directed to a particular user group and would help to pinpoint the roles, responsibilities and activities that would be carried out by different stakeholders. The guidelines for institutional RDM services would be different from those for the national one.

Processes: There is need to specify and regulate activities within the research data life-cycle including research data management planning for individual projects, data processing, ingesting data into central systems, selecting data for preservation and involving the use of standards and standardised procedures wherever possible. This can encompass the roles and responsibilities of the stakeholders involved in deciding on, setting up, maintaining and managing the research data repository as pointed out by Flores, Brodeur, Daniels, Nicholls and Turnator, (2007). The processes would point out the steps that should be taken when establishing RDM service in Zimbabwe.

Technologies: Research institutions in Zimbabwe through the working group should ensure that there are functional data repositories and networking infrastructures to allow the storage and transport of data. A major consideration for all of those offering, or planning, a repository service is selecting an appropriate software platform. They should choose whether to use open source or proprietary software considering the cost implications.

Services: Research institutions should ensure that there is end-user access to systems and support for research data life-cycle activities. The activities include supporting the creation of data management plans, providing skills training, and delivering helpdesk services to encourage researchers to archive their research data. There is need to demonstrate the value of the research data repository to the research community. Flores et al (2007) mentions that it is one thing to have the support of senior administrators who can see the practical value of providing access to research outputs and research resources, but it is another to convince people to go through the necessary steps to ensure their materials are deposited.
FURTHER RESEARCH DIRECTIONS
A study can be done on the willingness of Zimbabwean researchers to archive their research data in data repositories. It was indicated that researchers are not at ease when they are asked to contribute their research data since it can be prone to abuse by other researchers.

CONCLUSION
It can be concluded that various strategies can be used to establish research data management services in Zimbabwe. RDM services can be established at institutional level or at national level or researchers can be encouraged to use already established research data repositories. There are various research data repositories that can be utilised by researchers whereby some are research specific or general repositories. In Zimbabwe various stakeholders such as the Government, institutions of higher learning, research institutes, non-governmental organisations, funders and other boards should be involved in the establishment of RDM services. A working group should be formed to steer the process of establishing RDM services in Zimbabwe. It was noted that a number of challenges are encountered when working towards the development of RDM services. These include but are not limited to lack of skills, lack of resources, poor ICT infrastructure and unwillingness of researchers to submit their research data for archiving. Therefore, there is need to develop capacity and advocate for the development of RDM services at every level to ensure that it is supported by all the stakeholders. A proposed framework consisting of strategies, policies, guidelines, processes, technologies and services should be followed when establishing RDM services in Zimbabwe.

REFERENCES


KEY TERMS AND DEFINITION

**Research data:** Research data are recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings irrespective of the format in which it is created.

**Research data repository:** A database infrastructure that is set up to manage, share, access and archive researchers’ datasets.

**Research data management:** It covers the planning, collecting, organising, managing, storage, security, backing up, preserving, and sharing data. It ensures that research data are managed according to legal, statutory, ethical and funding body requirements.

**Research data services:** A network of services throughout the library to assist researchers during all phases of the research data lifecycle.

**Research institutes:** An establishment founded for doing research.

**Workgroup:** Two or more individuals who routinely function like a team, are interdependent in achievement of a common goal.

**Community of practice:** Informal, self-organized, network of peers with diverse skills and experience in an area of practice or profession. Such groups are held together by the members' desire to help others by sharing information and the need to advance their own knowledge by learning from others.

**Data Librarian:** Data librarians are professional library staff engaged in managing research data, using research data as a resource, or supporting researchers dealing with data. They equip participants with the necessary knowledge to develop and implement services for research data management.
**Framework:** a basic structure underlying a system, concept, or text.

**Stakeholders:** anybody who can affect or be affected by an organisation, strategy or project.