

Teaching research data management courses in higher learning institutions in Tanzania

Neema Florence Mosha

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

Patrick Ngulube

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

Abstract

Purpose – The study investigated teaching research data management (RDM) courses in higher learning institutions (HLIs) in Tanzania to enable postgraduate students to work with their research data.

Design/methodology/approach – The study triangulated research methods. Postgraduate students were investigated using survey questionnaires to learn about their needs and perceptions of the teaching RDM courses in HLIs. Key informants (academicians, information and communication technologists and library staff) were also investigated using in-depth interviews to explore their experiences and knowledge of teaching RDM courses. SPSS statistical software was used for analysing quantitative data; qualitative data were analysed thematically.

Findings – A total of 70 questionnaires were distributed to postgraduate students with a returning rate of 44 (69%). On the other hand, 12 key informants were interviewed. A low level of RDM literacy was revealed among 38 (86%) respondents. Most respondents 40 (91%) reported the need for HLIs to start teaching RDM courses. A lack of skills and knowledge in teaching RDM courses was revealed among key informants. The competency-based, adaptive and constructive teaching techniques were selected for teaching RDM courses, whereas intensive training and online tutorials were revealed as teaching formats.

Research limitations/implications – This study focused on teaching RDM courses in HLIs. The survey questionnaires were distributed to all 2nd year postgraduate students, however, the findings cannot be generalised to all postgraduate students due to the response rate obtained. The findings obtained from key informants can also not be used as a basis for generalization across HLIs.

Practical implications – This study concluded that postgraduate students need to be well equipped with skills and knowledge on RDM and its related concepts; teaching RDM courses should be regarded as a continuous programme for benefit of students, researchers and the community at large.

Social implications – Appropriate teaching of RDM courses among students not only ensures that students meet the funders' and publishers' requirements, but also encourages students to store and share their research among researchers worldwide; thus increasing collaboration and visibility of the datasets and data owners through data citations and acknowledgements.

Originality/value – This is a comprehensive study that provides findings for HLIs to teach RDM courses in HLIs, especially for postgraduate students. The findings revealed the need for teaching RDM courses in HLIs. The study provides the basis for further RDM research in HLIs and research institutions.

Keywords Teaching, Tanzania, Research data management, Postgraduate students, Higher learning institutions, Research data management courses

Paper type Research paper



Introduction

The development of research data as a key source of information is now becoming apparent; thus, to take advantage of this revolution, students and researchers in higher learning institutions (HLIs) need to be trained in research data management (RDM) and related fields. Research data are defined as the output from any systematic investigation that involves a process of investigation, experiment or testing of a hypothesis, as well as from artefacts of cultural heritage (Pryor, 2012; Nielsen and Hjørland, 2014). Lewis (2010) defined RDM as the processing of all types according to raw materials, which is created along every research process. Research data is increasingly recognised as a vital resource whose value needs to be preserved for future research (Perrier *et al.*, 2017); and thus, it places a huge responsibility on HLIs to teach RDM courses. Students in HLIs are always working with research data and data management in different ways (Radecki and Springer, 2020); however, formal education in RDM is very rare at the university level (Whitmire, 2015; Wiljes and Cimiano, 2019).

Studies indicate the need for teaching RDM concepts and practices to enable students to apply new methods of managing their research data very (Tammaro and Casarosa, 2014; Carlson and Bracke, 2015; Whitmire, 2015; Pasek and Mayer, 2019; Perrier *et al.*, 2017; Wiljes and Cimiano, 2019; Yang and Li, 2020). There is a larger gap in understanding graduate students' RDM needs (Pasek and Mayer, 2019). In this regard, HLIs need to teach RDM courses to ensure students meet their RDM needs in all stages as specified in the research data life cycle (Schneider, 2013). According to Wiljes and Cimiano (2019), the main aim of teaching RDM courses in HLIs is to enable students to apply the contents of the courses in their research work.

Teaching RDM courses is gaining importance worldwide as it ensures data quality, integrity, shareability, discoverability and reuse of data over time (Henderson and Corry, 2021; Perrier *et al.*, 2017; Wiljes and Cimiano, 2019). For example, the Oregon State University (OSU) has established a RDM course that is open to graduate students from all disciplines without any prerequisites and is taught by the libraries' data management specialists. Curricular materials are drawn from existing resources where possible, but much of the lecture and computer laboratory content was developed from scratch (Whitmire, 2015). Syracuse University developed a program "to train students with knowledge and skills in collecting, processing, managing, evaluating, and using data for scientific inquiry, the project known as Science Data Literacy" (Qin and D'ignazio, 2010). Schneider (2013) adds that forthcoming generations of students of almost all disciplines will have to work with their research data in systematic and scientific ways.

Studies identified gaps based on the lack of teaching RDM courses in HLIs. O'Reilly *et al.* (2012) explain that researchers do not undergo any training, which leads to unreliability, inaccessibility and short lifespan of their data and data management in general, and they do not receive any help on how to manage their data (Buys and Shaw, 2015). This means that they learn as they go, unsystematically, during their day-to-day research work (Fear, 2011; Peters and Dryden, 2011). On the other hand, research indicates a lack of knowledge and skills in RDM, from data generation to data storage and sharing (Perrier *et al.*, 2017; Wiljes and Cimiano, 2019; Henderson and Corry, 2021).

Lack of competencies, lack of practices and perceptions of data use, range of data sharing and data type, as well as incorrect researchers' perceptions toward RDM hinder HLIs from teaching RDM courses (Peters and Dryden, 2011; Jahnke *et al.*, 2012; Hickson *et al.*, 2016; Kurata *et al.*, 2017). For example, Whitmire (2015, p. 2) explain the RDM components of open data and data sharing as "... that many graduate students would probably not be prepared to deposit a dataset into the IR because they would not have received any formal training in data management during their programs, and as such, their data would not be an insufficient condition to be shared (i.e. be well organized and sufficiently documented)". The quote insisted that competencies in working with data are often not included as a part of a student's formal education (Carlson and Stowell-Bracke, 2013).

This study investigated the teaching of RDM courses in HLIs in Tanzania. Additionally, the present study focused on postgraduate students who are working on their research. Graduate students do not have a lot of knowledge of RDM; thus, HLIs shall be required to obtain RDM skills for their research progress and job requirements in educational fields (Pasek and Mayer, 2019).

Methods

Study area

This multimethod design study was conducted at the Nelson Mandela African Institution of Science and Technology (NM-AIST), one of the public universities in Tanzania. The institution has four schools, namely the School of Computational and Communication Science Engineering (CoCSE), the School of Materials, Energy, Water and Environmental Sciences (MEWES), the School of Life Sciences and Bio- Engineering (LiSBE), and the School of Business and Humanities (BuSH). However, the School of BuSH does not have any students, since it was established to enable the other three schools to integrate science and engineering subjects with business studies. Therefore, the School of BuSH was not part of this study. The NM-AIST was selected based on the initial preparation for the implementation of the RDM services such as the approved RDM policy and operational procedures, as well as, a well-equipped Information and Communication Technology (ICT) department with equipment like servers and High-Performance Computing (HPC) for storing research data and handling big data.

Study population

The study involved postgraduate students who were in their 2nd year, and who had already started their data collection. The study selected postgraduate students because they have already presented their research proposal, have collected their research data and they were at the stage of analysing data and report writing; thus they have the experience in RDM and its related concepts. Mannheimer and Banta (2017), RDM instruction programs often focus on faculty and graduate students who normally work with research data, thus, improving their data management strategies. Key informants involving academic, library and ICT staff were selected to add more information and experiences on teaching RDM courses and the advantages of RDM in HLIs.

Sample size and sampling procedures

The whole population (2nd postgraduate students) were involved in this study because the population size was very small (70 postgraduate students) then surveying all 2nd year students was ideal and more rigorous. Key informants were conveniently selected. The interview data were studied and collected until it was clear that perspectives were being repeated and data saturation was reached. In this study, the point of saturation was reached when the researcher conducted interviews with 12 key informants. The interview venue was selected based on the convenience and preferences of the respondents, taking into consideration their privacy and comfort. The library staff were assisted to set up appointments for the interviews and distributing the questionnaires to the respondents. All respondents were requested to fill out the informed consent before the interviews.

Data collection

In line with the multimethod design, qualitative data were collected from key informants to supplement quantitative data collected from postgraduate students. Survey questionnaires

were distributed to 70 postgraduate students, with a return rate of ($n = 44, 69\%$), whereas 12 key informants were interviewed using a developed interview guide.

Data analysis

In this study, both quantitative and qualitative data were analysed separately, and then qualitative data were merged with the quantitative data set to provide support to the results of the quantitative data set (Creswell and Plano-Clark, 2007). The study performed data cleaning and analysis using SPSS version 20 (Statistical Package for Social Science Software (SPSS), IBM SPSS Statistics). The thematic analysis approach was used to analyse qualitative data from key informants. The analysis was carried out in three stages: first, the line-by-line coding of field notes and transcripts (unpacking of text into discrete elements to expose underlying thoughts and meanings); second, the in-depth examination and interpretation of the resultant codes into descriptive themes and; third, interpretation of the descriptive themes into more abstract analytical themes. The study further employed between-method triangulation, whereby qualitative data were used to corroborate the findings obtained from quantitative data.

Results

A total of 44 postgraduate students aged 21–50 years were enrolled in this study of which 33 (75%) were males. The demographic information for postgraduate students was provided (Table 1). Further qualitative data was collected from 12 key informants (6 Academicians, 4 Librarians and 2 IRC staff).

RDM literacy level

Overall, postgraduate students were not aware of the RDM and its related concepts. Respondents 38 (86%) reported a low level of RDM literacy, with four respondents reporting a high level and two respondents saying they had a moderate level (Table 2).

Item(s)	Frequency (<i>n</i>)	Percentage (%)	
Gender	Male	33	75
	Female	11	25
Age	21–30	10	23
	31–40	25	57
	41–50	9	20
Education level	Bachelor Degree	27	61
	Master Degree	17	39
School	School of Computational and Communication Science Engineering (CoCSE)	15	34
	School of Materials, Energy, Water and Environmental Sciences (MEWES)	12	27
	School of Life Sciences and Bio- Engineering (LiSBE)	17	39

Table 1. Demographic information of postgraduate students ($N = 44$)

RDM literacy level	Frequency	Percentage
High	4	10
Moderate	2	4
Low	38	86
<i>Total</i>	<i>44</i>	<i>100</i>

Table 2. RDM literacy level

Key informants described RDM literacy level as the skills obtained by students to enable them to work with their research from data collection/generation to data curation. They saw RDM courses as being essential for students to work with their research. They believed that students gain elementary knowledge in RDM during research methodology courses offered at NM-AIST as follows: research methods and communication (Master and PhD), applied research methods (Master and PhD), research methods and statistics (Master), and biostatistics (Master). Indeed, they identified a need for RDM for students to learn how to work with their research. As one respondent reported:

RDM courses shall be taken as among the core courses to enhance RDM literacy level for students to work with their research" [Academician #1]. Capacity building among librarians, IRC and academicians was considered important in teaching RDM courses. Another respondent reported, "We think that there is a need for the institution to teach RDM courses among students; In fact, the institution should empower its staff especially librarians to be competent in RDM [Librarian #1].

The need for teaching RDM courses in HLIs

Respondents were further asked to indicate if there is a need for HLIs to start teaching RDM courses in HLIs. Respondents 40 (91%) reported the need for teaching RDM courses in HLIs. Key informants considered RDM as among the topics in HLIs and they supported the need for teaching RDM courses in HLIs. However, capacity building for facilitators and instructors for teaching RDM courses is highly recommended.

One respondent reported that:

Even though RDM skills to teach our students are needed, we cannot wait until we get training since most of us, we have research methodology skills to start with [Academician #2].

Teaching techniques for RDM courses

Generally, teaching techniques for teaching RDM courses in HLIs were required. In this case, competency-based teaching and adaptive teaching techniques were selected by most respondents 39 (23%) (Table 3).

Key informants selected the constructive teaching technique since it involves both lecturers and students and it encourages teamwork. One respondent added:

I think constructive teaching technique enables learners to contribute and share their knowledge [Librarian #2].

Another respondent insisted:

Since RDM is among the new courses in HLIs, constructive teaching technique could enable both students and lecturers to participate well in teaching and learning activities [Academician #5].

Teaching techniques	Frequency	Percentage
Distributed teaching techniques	20	13
Collaborative teaching techniques	32	19
Adaptive teaching techniques	39	23
Competency-based teaching techniques	39	23
Constructive teaching techniques	37	22
<i>Total</i>	<i>167</i>	<i>100</i>

Table 3.
RDM teaching
techniques

Note(s): NB: Multiple responses were allowed

In addition, constructive teaching technique provides both practical and theory to enable students to learn in deep and be competent in the subject area. One of the respondents contributed that:

I also think that any teaching technique that involves both hand-on and theory will enable students to be able to practice well, thus, HLIs should use techniques which are focusing more on practical than theory [IRC Staff #1].

Another respondent added that:

I encourage RDM modules which involve “hands-on activities” to enable researchers to learn about RDM software and tools, therefore, instead of just demonstrating the tools to allow students use them [Academician #6].

Teaching format for RDM courses

Respondents were also required to select a teaching format that will enable them to catch lessons very easy. Many respondents 34(27%) preferred intensive training (online) (Table 4).

The findings from key informants mentioned intensive training (online) and online tutorials (video clips, information etc.) as teaching formats for teaching RDM in HLIs. One respondent added that:

Online teaching and learning formats need to be considered in teaching RDM courses in HLIs. These formats are very important since they can enable students and other learners to study in their own time and to have learning resources to assist them in their studies [Librarian #3].

Another respondent insisted on the sharing of learning materials to students using online mechanisms, and he, therefore, reported that:

We like facilitators to produce and circulate teaching and learning materials and share them online to enable students to use them at their own time [Librarian #4].

Another respondent provided the advantages of online formats:

I think online teaching formats will enable tutors to reach a big number of students wherever they are and these are good formats especially, in this time when the world experiencing covid-19 pandemic in which social distancing is among the preventive measures [Academician # 8].

RDM topics to be covered

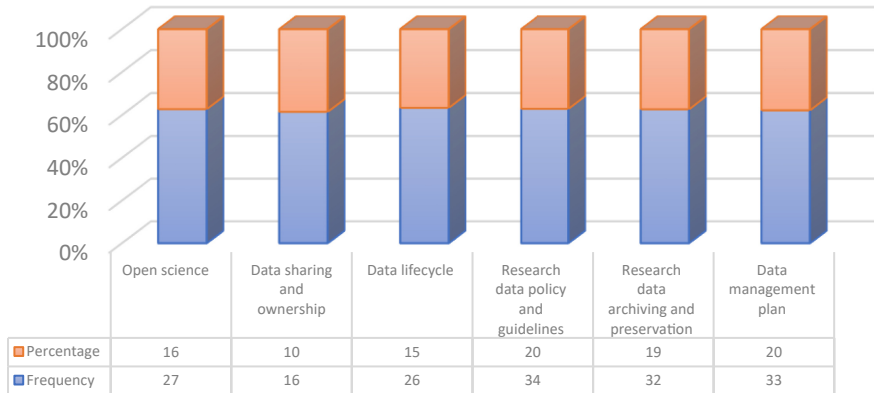
RDM topics were also identified, and the majority of respondents 34(20%) selected research data policy and guidelines (Figure 1).

Key informants' respondents selected research data life cycle and Data Management Plan (DMP). One respondent reported that:

RDM topics	Frequency	Percentage
Presentation or seminars (online)	17	14
Presentation or seminars (live)	30	24
Online tutorials (video clips, information etc.)	12	9
Interactive workshops	6	5
Intensive training (online)	34	27
Intensive training (live)	26	21
<i>Total</i>	<i>125</i>	<i>100</i>

Note(s): NB: Multiple responses were allowed

Table 4.
RDM teaching formats

Figure 1.
RDM topics

I suggest practical implementation and follow-up of the DMP. DMP is very important during early stages of data collection to data storage [Academician #7].

Benefits of teaching RDM courses in HLIs

There are several benefits for teaching RDM courses in HLIs. Respondents 41(24%) indicated meeting publishers' requirements, whereas key informants mentioned funders' requirements. One of the respondents felt that meeting funder requirements were of primary importance for students, HLIs and the nation in general:

Most of the funders' requirements need to see their data to be open, hence to enable other researchers to use the data for more research findings and to use the fund for the development of the country and its citizens [Academician #3].

Another respondent added:

Data sharing and open data are among the components in RDM that enable funders to invest more in research [Academician #4].

Other benefits mentioned included saving money and time for recollecting research data on the same area and strengthening scholarly communication.

Discussion

The findings revealed that RDM literacy was low among respondents. A similar finding was provided by [Mushi et al. \(2020\)](#), which revealed that the University of Dodoma (UDOM) lecturers and stakeholders (students) have a low level of RDM and related areas. This finding was inconsistent with the study of [Adika and Kwanya \(2020\)](#) which indicated RDM literacy levels among 50 students interviewed on research data literacy skills as 2(4%) poor, 38(76%) good, and 10(20%) very good. Also, the study conducted at the University of Oregon by [Carlson et al. \(2015\)](#) revealed that "... students may have picked up their skills on the job training because a lot of them had a former life in a professional field ... or it's something they got as an undergraduate". Again, the study by [Carlson and Stowell-Bracke \(2013\)](#) noted that graduate students at Purdue were involved in the collecting, processing and analysing of research data. It is thus concluded that postgraduate students should have a kind of RDM skills to work with their research as compared to undergraduate students.

Generally, teaching RDM courses is mandatory in HLIs. The study findings revealed that RDM courses will equip researchers and students in HLIs with more skills and knowledge.

This finding was consistent with preceding studies by [Carlson *et al.* \(2011\)](#), [Buys and Shaw \(2015\)](#), [Mandinach *et al.* \(2015\)](#) and [Cowie and Cooper \(2017\)](#). The study by [Wiljes and Cimiano \(2019\)](#) also added that The Cluster of Excellence Cognitive Interaction Technology (CITEC) [1] released the CITEC Open Science Manifesto [2] stated that “CITEC recognizes the need to extend the educational curriculum for young scientists towards topics of research data management and offers training and personal consulting for advanced researchers, thus contributing to awareness among young researchers of good practice in scientific research”. Teaching RDM courses in HLIs enables researchers to work with their research data as well as ensure risks regarding data security and privacy, which many researchers are unaware of ([Jahnke *et al.*, 2012](#)). [Buys and Shaw \(2015\)](#) found that researchers need help, guidance and training regarding long-term access to data, long-term preservation of data, becoming acquainted with services that support these processes, as well as providing backup copies of data during the project lifetime.

The study found competency-based and adaptive teaching techniques in teaching RDM courses in HLIs. A similar finding emerged from the study done by [Wiljes and Cimiano \(2019\)](#), namely that the competency teaching technique that relies on practical exercise is useful for teaching RDM courses in HLIs. The study is inconsistent with the study done by [Biernacka *et al.* \(2021\)](#), which proposed ways of teaching and learning data mining and learning analytical skills by using adaptive learning (adaptive teaching technique) to orchestrate the interaction with learners and deliver customised resources and learning activities to address the unique need of each learner. [Whitmire \(2015\)](#) added that the active learning exercises were successful and generally enjoyed by students. In addition, the active approach is driven by practical reality and thus enables students to engage with the content and techniques of RDM rather than lecture alone ([Whitmire, 2015](#)). The study also found that a constructive teaching technique should be used in teaching RDM in HLIs, which is similar to the study of [Mannheimer and Banta \(2017\)](#), which found that to make RDM relevant among students in HLIs, instructors need to use constructive learning for strategies that create meaningful learning experiences. This study found that any mode which involves “hands-on activities” enables students to work with their research data practically. [Wiljes and Cimiano \(2019\)](#) add that hands-on demonstrations, group discussions and individual presentations allow students to link the acquired knowledge to their research works. HLIs should use diverse teaching and activation methods to increase engagement and active participation among students ([Biernacka *et al.*, 2021](#)). In addition, application, knowledge transfer, and replication methods need to be integrated to facilitate the learning process ([Biernacka *et al.*, 2021](#)).

The study findings revealed that an intensive training (online) format should be considered in teaching RDM courses in HLIs. Online learning systems learning management systems, learning platforms, and learning software can capture streams of fine-grained learner behaviours, and the tools and techniques described above can operate on the data to provide a variety of stakeholders with feedback to improve teaching, learning, and educational decision making. A study by [Mushi *et al.* \(2020\)](#) recommended several online training modules available freely for teaching RDM courses.

The study finding also revealed that DMP should be considered among the topics to be covered in teaching RDM in HLIs. The finding is consistent with studies done by [EDINA and Data Library \(2011\)](#), and [Henkel *et al.* \(2012\)](#) on DMP, metadata and documentation, organisation, storage, backup and preservation practices, and data publication and sharing are among the topics to be considered in teaching RDM and their instructional programs. In addition, [Baker *et al.* \(2015\)](#) recommended topics that include data collection techniques, development of a data management plan, keeping data collection activity logs and providing training in summarising the results. [DataONE \(2016\)](#) proposed RDM topics to be covered including data management, data sharing, DMP, data manipulation, data quality control assurance, data protection, metadata, data citation, analysis and workflow, as well as legal

and policy issues. [NLM \(2012\)](#) developed seven modules for teaching RDM courses, namely Module 1: an overview of RDM; Module 2: data types, strategies and format; Module 3: metadata; Module 4: data storage, backup and security; Module 5: legal and ethical considerations; Module 6: data sharing and reuse policies; and Module 7: archiving and preservations.

Parallel to previous studies done by [Kennan and Markauskaite \(2015\)](#), [Patterton *et al.* \(2018\)](#), [Bangani and Moyo \(2019\)](#), [Chigwada *et al.* \(2019\)](#), [Adika and Kwanya \(2020\)](#), and [Mushi *et al.* \(2020\)](#), the present study found that publishers' and funders' requirements are among the benefits for teaching RDM courses in HLIs. [Kennan and Markauskaite \(2015\)](#) added that "There is increasing pressure from funders, publishers, the public, universities and other research organisations for researchers to improve their data management and sharing practices". Funders' pressure to add value to expensive research and to solve cross-disciplinary grand challenges as well as pressure from publishers to ensure transparency and reproducibility of the scientific record, made researchers acquire RDM knowledge ([Kennan and Markauskaite, 2015](#); [Patterton *et al.*, 2018](#); [Bangani and Moyo, 2019](#)). [López-Meneses *et al.* \(2020\)](#) explored the benefit of RDM, namely to enable students for their professional development. Additionally, [Yang and Li \(2020\)](#) added that teaching RDM courses enable students to conduct their research activities effectively and HLIs to improve the quality of their research services. [EPSRC \(2018\)](#) mentioned the following benefits of RDM in HLIs, especially on data storage and sharing:

- (1) Increases the visibility of research data for citation and reuse by other researchers.
- (2) Reinforces scientific enquiry by ensuring that enthusiasts and sceptics alike can test, validate and replicate research results.
- (3) Promotes new research and different ways of testing and analysing research data.
- (4) Discourages unnecessary duplication of effort by researchers.
- (5) Saves financial and other resources that are wasted when similar data sets are created by different researchers.
- (6) Enables discoveries from old research data sets.

The main strength of this study is its composition of respondents (students, academicians, librarians, ICT Personnel) involved in teaching RDM courses in HLIs. Awareness of these issues is essential for HLIs' management together with their units within the HLIs, to be able to design RDM courses and develop RDM services that suit the needs of the research community. The findings can also be used to develop RDM curriculum based on different RDM topics and shall be recognized as a standard guideline.

Conclusion

This study concludes that there is a need for HLIs to start teaching RDM courses to their students and thus, RDM courses shall be considered among the topics to be included in curriculums. Despite the benefits mentioned in this study, RDM shall be required to enable students in HLIs to share their data and other researchers worldwide and store their data in data repositories for future use. Research data has a long lifespan than the project, thus, students need to plan for how long they will manage their data after the project has finished, as well as before and during research. Research data literacy skills are very important for students even after completing their studies. Therefore, teaching RDM in HLIs shall be considered an essential part of academic education ranging from online guides, classroom presentations, and workshops. Teaching techniques that includes both theoretical and

practical sessions shall be accommodated for teaching RDM courses in HLIIs. Additionally, the blended learning format is appropriate for teaching RDM courses since it combines both traditional classroom (face-to-face) and online courses. This study will be the basis for other HLIIs to set up similar courses for their students, researchers and staff.

Recommendations

- (1) HLIIs' should think of teaching RDM courses to their students especially for those engaging in research activities.
- (2) HLIIs' worldwide should implement the teaching technique which will involve both theory and practical sessions to enable students to be well informed about the intended courses.
- (3) HLIIs should ensure that capacity building is provided for all HLIIs' staff who will be involved in teaching RDM and related courses. This can be through long and short course training, workshop programs, advocacy and seminars.
- (4) HLIIs shall ensure that the RDM infrastructure is well implemented to enable data analysis, storage and sharing.

Notes

1. <http://cit-ec.de>.
2. <https://www.cit-ec.de/en/open-science/manifesto>

References

- Adika, F.O. and Kwanya, T. (2020), "Research data management literacy amongst lecturers at Strathmore University", *Library Management*, Vol. 41 Nos 6-7, pp. 447-466, doi: [10.1108/LM-03-2020-0043](https://doi.org/10.1108/LM-03-2020-0043).
- Baker, K.S., Duerr, R.E. and Parsons, M.A. (2015), "Scientific knowledge mobilization: co-evolution of data products and designated communities", *International Journal of Digital Curation*, Vol. 10 No. 2, pp. 110-135, doi: [10.2218/ijdc.v10i2.346](https://doi.org/10.2218/ijdc.v10i2.346).
- Bangani, S. and Moyo, M. (2019), "Data sharing practices among researchers at South African Universities", *Data Science Journal*, Vol. 18 No. 1, pp. 1-14, doi: [10.5334/dsj-2019-028](https://doi.org/10.5334/dsj-2019-028).
- Biernacka, K., Helbig, K. and Buchholz, P. (2021), "Adaptable methods for training in research data management", *Data Science Journal*, Vol. 20 No. 1, pp. 1-9, doi: [10.5334/dsj-2021-014](https://doi.org/10.5334/dsj-2021-014).
- Buys, C.M. and Shaw, P.L. (2015), "Data management practices across an institution: survey and report", *Journal of Librarianship and Scholarly Communication*, Vol. 3 No. 2, pp. 0-24, doi: [10.7710/2162-3309.1225](https://doi.org/10.7710/2162-3309.1225).
- Carlson, J. and Bracke, M. (2015), "Planting the seeds for data literacy: lessons learned from a student centered education program", *Data Information Literacy Case Study Directory*, Vol. 1 No. 1, pp. 1-5, doi: [10.5703/1288284315518](https://doi.org/10.5703/1288284315518).
- Carlson, J. and Stowell-Bracke, M. (2013), "Data management and sharing from the perspective of graduate students: an examination of the culture and practice at the water quality field station", *Libraries Faculty and Staff Scholarship and Research*, Vol. 13 No. 4, pp. 343-361, doi: [10.1353/pla.2013.0034](https://doi.org/10.1353/pla.2013.0034).
- Carlson, J.R., Fosmire, M., Miller, C. and Nelson, M.R.S. (2011), "Determining data information literacy needs: a study of students and research faculty", *Libraries and the Academy*, Vol. 11 No. 2, pp. 629-657, doi: [10.1353/pla.2011.0022](https://doi.org/10.1353/pla.2011.0022).

- Carlson, J., Jeffryes, J., Johnston, L.R., Nichols, M., Westra, B. and Wright, S.J. (2015), "An exploration of the data information literacy competencies: findings from the project interviews", in Carlson, J. and Johnston, L.R. (Eds), *Data Information Literacy: Librarians, Data and Education of a New Generation of Researchers*, Prude University Press, London, pp. 52-69.
- Chigwada, J.P., Hwalima, T. and Kwangwa, N. (2019), "A proposed framework for research data management services in research institutions in Zimbabwe", in Bhardwaj, R.K. and Banks, P. (Eds), *Research Data Access in Modern Libraries*, IGI Global, Hershey, pp. 29-53, doi: [10.4018/978-1-5225-8437-7.ch002](https://doi.org/10.4018/978-1-5225-8437-7.ch002).
- Cowie, B. and Cooper, B. (2017), "Exploring the challenge of developing student teacher data literacy", *Assessment in Education: Principles, Policy and Practice*, Vol. 24 No. 2, pp. 147-163, doi: [10.1080/0969594X.2016.1225668](https://doi.org/10.1080/0969594X.2016.1225668).
- Creswell, J.W. and Plano-Clark, V. (2007), *Designing and Conducting Mixed Methods Research*, Sage Publishing, Thousands Oaks, CA.
- Data Observation Network for Earth (DataONE) (2016), "Data management education modules", available at: <https://old.dataone.org/education-modules> (accessed 15 January 2022).
- EDINA and Data Library (2011), "Research data MANTRA project at EDINA, information services, university of Edinburgh", available at: <http://datalib.edina.ac.uk/mantra> (accessed 15 January 2022).
- Engineering and Physical Sciences Research Council (EPSRC) (2018), "Data management plan (DMP) information, scope and benefits", available at: <https://epsrc.ukri.org/about/standards/researchdata/scope/> (accessed 2 December 2021).
- Fear, K. (2011), "You made it, you take care of it: data management as personal information management", *International Journal of Digital Curation*, Vol. 6 No. 2, pp. 53-77, doi: [10.2218/ijdc.v6i2.190](https://doi.org/10.2218/ijdc.v6i2.190).
- Henderson, J. and Corry, M. (2021), "Data literacy training and use for educational professionals", *Journal of Research in Innovative Teaching and Learning*, Vol. 14 No. 2, pp. 2397-7604, doi: [10.1108/jrit-11-2019-0074](https://doi.org/10.1108/jrit-11-2019-0074).
- Henkel, H., Viv Hutchison, C.S., Stacy, R.H., Kristin, V., Lynda, W., Stephanie, H., Amber, B., Yiwei, W., David, B., Amy, H., Gail, S., Stephanie, W. and Matt, M. (2012), "DataONE education modules", available at: <https://www.dataone.org/education-modules> (accessed 26 January 2022).
- Hickson, S., Poulton, K.A., Connor, Richardson, J. and Wolski, M. (2016), "Modifying researchers' data management practices: a behavioural framework for library practitioners", *IFLA Journal*, Vol. 42 No. 4, pp. 253-265, doi: [10.1177/0340035216673856](https://doi.org/10.1177/0340035216673856).
- Jahnke, L., Asher, A. and Keralis, S.D.C. (2012), *The Problem of Data*, Council on Library and Information Resources (CLIR, Alexandria, VA.
- Kenan, M.A. and Markauskaite, L. (2015), "Research data management practices: a snapshot in time", *International Journal of Digital Curation*, Vol. 10 No. 2, pp. 69-95, doi: [10.2218/ijdc.v10i2.329](https://doi.org/10.2218/ijdc.v10i2.329).
- Kurata, K., Matsubayashi, M. and Mine, S. (2017), "Identifying the complex position of research data and data sharing among researchers in natural science", *SAGE Open*, pp. 1-13, doi: [10.1177/2158244017717301](https://doi.org/10.1177/2158244017717301).
- Lewis, M.J. (2010), "Libraries and the management of research data in Envisioning future academic library services : initiatives, ideas and challenges", in McKnight, S. (Ed.), *Envisioning Future Academic Library Services*, Facet Publishing, London, pp. 145-168.
- López-Meneses, E., Sirignano, F.M., Vazquez-cano, E. and Rumirez-Hurtado, J.M. (2020), "University students' digital competence in three areas of the DigCom 2.1 model: a comparative study at three European Universities", *Australasian Journal of Educational Technology*, Vol. 36 No. 3, pp. 69-88, doi: [10.14742/AJET.5583](https://doi.org/10.14742/AJET.5583).
- Mandinach, B.E.B., Parton, B.M. and Gummer, E.S. (2015), "Ethical and appropriate data use requires data literacy", *Phi Delta Kappan*, Vol. 96 No. 5, pp. 25-28, doi:[10.1177/003172171556946589ik](https://doi.org/10.1177/003172171556946589ik).
- Mannheimer, S. and Banta, R. (2017), "Personal digital archiving as a bridge to research data management : theoretical and practical approaches to teaching research data management

- skills for undergraduates”, *The Complete Guide to Personal Digital Archiving for Librarians, Archivists, and Information Professionals*, p. 15.
- Mushi, G.E., Pienaar, H. and Deventer, M. (2020), “Identifying and implementing relevant research data management services for the library at the University of Dodoma, Tanzania”, *Data Science Journal*, Vol. 19 No. 1, pp. 1-9, doi: [10.5334/dsj-2020-001](https://doi.org/10.5334/dsj-2020-001).
- National Library of Congress (NLM) (2012), “Frameworks for a data management curriculum for science, health sciences, and engineering students”, *University of Massachusetts and New England Area Librarian e-Science Symposium*, available at: http://escholarship.umassmed.edu/escience_symposium/2012/posters/6/ (accessed 10 January 2022).
- Nielsen, H.J. and Hjørland, B. (2014), “Curating research data: the potential roles of libraries and information professionals”, *Journal of Documentation*, Vol. 70 No. 2, pp. 221-240, doi: [10.1108/JD-03-2013-0034](https://doi.org/10.1108/JD-03-2013-0034).
- O'Reilly, K., Johnson, J. and Sanborn, G. (2012), “Improving university research value : a case study”, *SAGE Open*, pp. 1-13, doi: [10.1177/2158244012452576](https://doi.org/10.1177/2158244012452576).
- Pasek, J.E. and Mayer, J. (2019), “Education needs in research data management for science-based disciplines: self-assessment surveys of graduate students and faculty at two public universities”, *Issues in Science and Technology Librarianship*, Vol. 92, pp. 1-28, doi: [10.29173/istl12](https://doi.org/10.29173/istl12).
- Patterton, L., Bothma, T. and Van Deventer, M. (2018), “From planning to practice: an action plan for the implementation of research data management services in resource-constrained institutions”, *South African Journal of Libraries and Information Science*, Vol. 84 No. 2, pp. 14-36, doi: [10.7553/84-2-1761](https://doi.org/10.7553/84-2-1761).
- Perrier, L., Blondal, E., Ayala, A.P., Dearborn, D., Kenny, T., Lightfoot, D., Reka, R., Thuna, M. and Trimble, L., M.H. (2017), “Research data management in academic institutions: a scoping review”, *PLoS One*, Vol. 12 No. 5, pp. 178-261, doi: [10.1371/journal.pone.0178261](https://doi.org/10.1371/journal.pone.0178261), eCollection 2017.
- Peters, C. and Dryden, A.R. (2011), “Assessing the academic library’s role in campus-wide research data management: a first step at the University of Houston”, *Science and Technology Libraries*, Vol. 30 No. 4, pp. 387-403, doi: [10.1080/0194262X.2011.626340](https://doi.org/10.1080/0194262X.2011.626340).
- Pryor, G. (2012), *Managing Research Data*, Facet Publishing, Cambridge. doi: [10.29085/9781856048910](https://doi.org/10.29085/9781856048910).
- Qin, J. and D’ignazio, J. (2010), “The central role of metadata in a science data literacy course”, *Journal of Library Metadata*, Vol. 10 No. 23, pp. 188-204, doi: [10.1080/19386389.2010.506379](https://doi.org/10.1080/19386389.2010.506379).
- Radecki, J. and Springer, R. (2020), “Research data services in US Higher Education, a web-based inventory, ITHAKA”, available at: <https://doi.org/10.5281/zenodo.4270332> (accessed 10 April 2022).
- Schneider, R. (2013), “Research data literacy”, *paper presented at the 1st European Conference on Information Literacy (ECIL)*, Istanbul, 22-25 October 2013, pp. 134-150, doi: [10.1007/978-3-319-03919-0](https://doi.org/10.1007/978-3-319-03919-0).
- Tamaro, A.M. and Casarosa, V. (2014), “Research data management in the curriculum: an interdisciplinary approach”, *Procedia Computer Science*, Vol. 38, pp. 138-142, doi: [10.1016/j.procs.2014.10.023](https://doi.org/10.1016/j.procs.2014.10.023).
- Whitmire, A.L. (2015), “Implementing a graduate-level research data management course: approach, outcomes, and lessons learned”, *Journal of Librarianship and Scholarly Communication*, Vol. 3 No. 2, doi: [10.7710/2162-3309.1246](https://doi.org/10.7710/2162-3309.1246).
- Wiljes, C. and Cimiano, P. (2019), “Teaching research data management for students”, *Data Science Journal*, Vol. 18 No. 1, pp. 1-9, doi: [10.5334/dsj-2019-038](https://doi.org/10.5334/dsj-2019-038).
- Yang, N. and Li, T. (2020), “How stakeholders’ data literacy contributes to student success in higher education: a goal-oriented analysis”, *International Journal of Educational Technology in Higher Education*, Vol. 17 No. 41, pp. 1-18, doi: [10.1186/s41239-020-00220-3](https://doi.org/10.1186/s41239-020-00220-3).

Further reading

- Akmon, D., Zimmerman, A., Daniels, M. and Hedstrom, M. (2011), “The application of archival concepts to a data-intensive environment: working with scientists to understand data

- management and preservation needs”, *Archival Science*, Vol. 11 Nos 3-4, pp. 329-348, doi: [10.1007/s10502-011-9151-4](https://doi.org/10.1007/s10502-011-9151-4).
- Boyle, J.A. (2004), “Bioinformatics in undergraduate education”, *Biocemistry and Molecular Biology Education*, Vol. 32 No. 4, pp. 236-238, doi: [10.1002/bmb.2004.494032040376/pdf](https://doi.org/10.1002/bmb.2004.494032040376/pdf).
- Briney, K., Coates, H. and Goben, A. (2020), “Foundational practices of research data management”, *Research Ideas and Outcomes (RIO)*, Vol. 6, pp. 1-17, doi: [10.3897/rio.6.e56508](https://doi.org/10.3897/rio.6.e56508).
- Carlson, J., Johnson, L., Westra, B. and Nichols, M. (2013), “Developing an approach for data management education: a report from the data information literacy project”, *International Journal of Digital Curation*, Vol. 8 No. 1, pp. 204-217, doi: [10.2218/ijdc.v8i1.254](https://doi.org/10.2218/ijdc.v8i1.254).
- Chiparousha, B. and Chigwada, J.P. (2019), “Accessibility of research data at academic institutions in Zimbabwe”, in Bhardwaj, R.K. and Banks, P. (Eds), *Research Data Access in Modern Libraries*, IGI Global, Hershey, pp. 81-89, doi: [10.4018/978-1-5225-8437-7.ch004](https://doi.org/10.4018/978-1-5225-8437-7.ch004).
- Cox, J. (2018), “Positioning the academic library within the institution: a literature review”, *New Review of Academic Librarianship*, Vol. 24 Nos 3-4, pp. 219-243, doi: [10.1080/13614533.2018.1466342](https://doi.org/10.1080/13614533.2018.1466342).
- Creswell, J.W. and Plano-Clark, V. (2018), *Designing and Conducting Mixed Methods Research*, 3rd ed., Sage Publishing, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., Gutmann, M. and Hanson, W. (2003), “Advanced mixed methods research designs”, in Tashakkori, A. and Teddlie, C. (Eds), *Handbook of Mixed Methods in Social and Behavioural Research*, Sage Publishing, Thousand Oaks, CA, pp. 209-240.
- Ditty, J.L., Kvaal, C.A., Goodner, B., Freyermuth, S.K., Bailey, C., Gordon, S.G., Heinhorst, S., Reed, K. and Xu, Z. (2010), “Incorporating genomics and bioinformatics across the life sciences curriculum”, *PLoS Biology*, Vol. 8 No. 8, pp. 1-4.
- Fearon, D.J., Gunia, B., Pralle, B.E. and Sallans, A.L. (2013), *Research Data Management Services*, Association of Research Libraries SPEC Kit, Washington, DC, Vol. 334, doi: [10.29242/spec.334](https://doi.org/10.29242/spec.334), available at: <https://publications.orl.org/Research-Data-Management-Services-SPEC-Kit-334/> (accessed 5 December 2021).
- Higgins, S. (2012), “The lifecycle of data management”, in Pryor, G. (Ed.), *Managing Research Data*, Facet Publishing, London, pp. 17-46.
- Koltay, T. (2015), “Data literacy: in search of a name and identity”, *Journal of Documentation*, Vol. 71 No. 2, pp. 401-415, doi: [10.1108/JD-02-2014-0026](https://doi.org/10.1108/JD-02-2014-0026).
- Koltay, T. (2016), “Data governance, data literacy and the management of data quality”, *IFLA Journal*, Vol. 42 No. 4, pp. 303-312, doi: [10.1177/0340035216672238](https://doi.org/10.1177/0340035216672238).
- Koltay, T. (2017), “Data literacy for researchers and data librarians”, *Journal of Librarianship and Information Science*, Vol. 49 No. 1, pp. 3-14, doi: [10.1177/096100615616450](https://doi.org/10.1177/096100615616450).
- Koltay, T. (2019), “Research data management and data literacy as we see them today”, *paper presented at the Library And Information Science XXVIII*, November, 2019, Comenius University, pp. 0-10.
- Mazon, J.N., Lloret, E., Gomez, E., Aguilar, A., Mingot, I., Perez, E. and Quereda, L. (2014), “Reusing open data for learning database design”, *paper presented at the International Symposium on Computers in Education*, pp. 59-64, SIIE 2014, doi: [10.1109/SIIE.2014.7017705](https://doi.org/10.1109/SIIE.2014.7017705).
- Piorun, M., Kafel, D., Leger-Hornby, T., Najafi, S., Martin, E.R., Colombo, P. and Nancy, R. (2012), “Teaching research data management: an undergraduate/graduate curriculum”, *Journal of eScience Librarianship*, Vol. 1 No. 1, 8 pp. 46-50, doi: [10.7191/jeslib.2012.1003](https://doi.org/10.7191/jeslib.2012.1003).
- Prado, J.C. and Marzal, M.Á. (2013), “Incorporating data literacy into information literacy programs: core competencies and contents”, *Libri*, Vol. 63 No. 2, pp. 123-134, doi: [10.1515/libri-2013-0010](https://doi.org/10.1515/libri-2013-0010).
- Sayogo, D.S. and Pardo, T.A. (2013), “Exploring the determinants of scientific data sharing: understanding the motivation to publish research data”, *Government Information Quarterly*, Vol. 30, pp. S19-S31, doi: [10.1016/j.giq.2012.06.011](https://doi.org/10.1016/j.giq.2012.06.011).

-
- Schumacher, J. and Vandecreek, D. (2015), "Intellectual capital at risk: data management practices and data loss by faculty members at five American Universities", *International Journal of Digital Curation*, Vol. 10 No. 2, pp. 96-109, doi: [10.2218/ijdc.v10i2.321](https://doi.org/10.2218/ijdc.v10i2.321).
- Teddle, C. (2009), *Foundation of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioural Sciences*, Sage Publishing, Thousands Oaks, CA.
- Vilar, P. and Zabukovec, V. (2019), "Research data management and research data literacy in Slovenian science", *Journal of Documentation*, Vol. 75 No. 1, pp. 24-43, doi: [10.1108/JD-03-2018-0042](https://doi.org/10.1108/JD-03-2018-0042).
- Voet, J.G., Bell, E., Boyer, R., Boyle, J., O'Leary, M. and Zimmerman, J.K. (2003), "Recommended curriculum for a program in biochemistry and molecular biology", *Biochemistry and Molecular Biology Education*, Vol. 31 No. 3, pp. 161-162, doi: [10.1002/bmb.2003.494031030223](https://doi.org/10.1002/bmb.2003.494031030223).

About the authors

Neema Florence Mosha is currently a Postdoctoral Research Fellow at the University of South Africa (UNISA). She is a Librarian at the Nelson Mandela African Institution of Science and Technology (NM-AIST), Arusha. She obtained her PhD in Information Studies from the University of South Africa (UNISA) and her Master of Information Technology (MIT) from the University of Pretoria (UP), South Africa. Her research focuses on the development of Research Data Management (RDM) services in Higher Learning Institutions (HLIs). She is interested in knowledge management, open science, digital scholarship and innovation in libraries, eResearch and Web 2.0 tools application in academic libraries. She is also involved in a project to train universities' staff and researchers in the fundamentals and essentials of RDM. Neema Florence Mosha is the corresponding author and can be contacted at: moshanf@unisa.ac.za

Patrick Ngulube is currently the Dean of the College of graduate studies at the University of South Africa (UNISA) and a professor in the department of interdisciplinary research and postgraduate studies. He has published scholarly works in indigenous knowledge systems, knowledge management, e-government, the preservation of access to information and research methodology. He is a rated national research foundation researcher.