libres volume 32, issue 1, pages 44-63 (2022). https://doi.org/10.32655/LIBRES.2022.1.4

Eliciting Researchers' Behaviour as the Foundation of Research Data Management Service Development

Mohd Ikhwan Ismail (corresponding author)
Department of Library & Information Science, Faculty of Arts and Social Sciences, Universiti Malaya, Malaysia
Hamzah Sendut Library, Universiti Sains Malaysia, Malaysia
ikhwanismail@usm.my

Cik Ramlah Che Jaafar
Hamzah Sendut Library, Universiti Sains Malaysia, Malaysia
cramlah@usm.my

Noor Adilah Azmi
Hamzah Sendut Library, Universiti Sains Malaysia, Malaysia
nooradilah@usm.my

Muaz Mohd Zaini Makhtar
School of Industrial Technology and Centre for Global Sustainability Studies, Universiti Sains Malaysia, Malaysia
muazzaini@usm.my

Samsul Farid Samsuddin
Department of Library & Information Science, Faculty of Arts and Social Sciences, Universiti Malaya, Malaysia
samsulfarid@um.edu.my

A. Abrizah (corresponding author)
Department of Library & Information Science, Faculty of Arts and Social Sciences, Universiti Malaya, Malaysia
abrizah@um.edu.my

ABSTRACT

Background. Research data management (RDM) has become an important activity in universities, for researchers to fulfil funding agencies' and journal publication requirements, and to promote open science practices. Academic libraries have been identified as the locations to base RDM services. However, to develop effective RDM services, an understanding of RDM from the researchers' perspectives is needed, including how researchers manage their research data.
Objectives. This study aims to discover researchers' behaviours and practices in RDM, and propose how the library can incorporate RDM into the research services offered.
Methods. This case study, carried out at a research university in Malaysia, involved both quantitative and qualitative data gathering, focusing on three

Published by WKW School of Communication & Information & NTU Libraries, Nanyang Technological University

aspects of RDM: data creation, data storage and preservation, and data sharing. Quantitative data were collected via a survey of 113 researchers, and qualitative data were gathered through semi-structured interviews with 12 researchers.
Results. It was found that the researchers had been generating research data irrespective of format and types. Most of the researchers managed their research data based on their own perspectives and practices, without following proper guidelines and standards. The researchers used personal solutions for research data storage and preservation, and utilized less than 10 gigabytes of storage for the short term. The researchers also did not share their research data due to privacy and confidentiality issues.
Contributions. Researchers need support for organizing, archiving and preserving research data for future use, and libraries can provide this important service. The study reflects the library's transformative role starting with conducting needs assessment of the academic research community, and establishing an RDM service.

INTRODUCTION

The importance of managing research data at universities worldwide has prompted university libraries to incorporate digital data and research data management (RDM) in the research support services offered. To be able to offer effective services, libraries should comprehend the researchers' needs in managing the data, and ensure that the services align with their needs (Syn & Kim, 2019). The library's role is to assist the researchers in managing the research data, and this involves curating, advising, and preserving research data (Chigwada et al., 2019). The literature highlighted that a librarian should become a partner to the researchers throughout the research lifecycle, and play an extra role in RDM, particularly in raising awareness and educating researchers on best practices in RDM (Bradley-Ridout, 2018; Conrad et al., 2017; Nitecki & Davis, 2019).

The open science movement has encouraged researchers to further recognize the importance of RDM activities. Undeniably, open science has become a key driver for institutions to develop RDM services (Chiware & Becker, 2018; Cox et al., 2017; Elsayed & Saleh, 2018; Rice, 2019). Both library and researchers need to build capacity in open science, in terms of knowledge and skills relating to research data to ensure the RDM will be well developed within the institution (Cox et al., 2017). However, open science services, such as data repositories, are less used because openness in research is not fully practised by the researchers (Rebouillat, 2017). This may be because researchers are not assessed by their institutions and funding agencies on the basis of their works' openness and transparency, such as open access and open data.

In Malaysia, the open science initiative began in November 2019 when the Ministry of Energy, Science, Technology, Environment and Climate Change (MESTECC) and the Academy of Sciences Malaysia (ASM) launched the Malaysia Open Science Platform (MOSP) (Abd Rahman, 2019). This platform has been envisaged to be a trusted data-sharing platform for Malaysian researchers, enabling the accessibility and sharing of research data. Although open science guidelines are already in place, the endeavour to manage research data, particularly data sharing in research performing institutions, is still in the early stages. One of MOSP's focus areas is capacity building and awareness, where the aim is to train 200 data stewards among librarians, research managers, and researchers. This training programme

aims to upskill these targeted groups, and the training module serves as a guideline for data stewards or data librarians to apply the RDM skills in their institution. Despite the optimistic concept and readiness of open data initiatives, open data sharing among Malaysian researchers is a significant issue that must be addressed (Abrizah, 2019).

Previous papers have highlighted issues relating to RDM activities and perceptions among researchers, in areas such as data creation, data storage, data sharing, data policy and data management skills (e.g., Aydinoglu et al., 2017; Bunkar & Bhatt, 2020; Lu & Ke, 2020; Majid et al., 2018; Tenopir et al., 2011). Data sharing issues are the most often mentioned, with negative perceptions on the part of researchers (Pinfield et al., 2014). It has been reported that researchers are uncertain and do not feel comfortable with data sharing due to reasons such as lack of interest and readiness for research data to be opened and shared (Cox et al., 2017; Flores et al., 2015; Plomp et al., 2019; Vela & Shin, 2019), despite the encouragement or requirement by policymakers, funding agencies, and publishers for data sharing. Researchers seem reluctant to share their research data with others, and one of the reasons is the fear that data would be misused or misinterpreted (Borgman, 2012; Fecher et al., 2015; Vilar & Zabukovec, 2019; Wallis et al., 2013). Studies on data sharing willingness among researchers reported that only a small percentage (12%) were willing (Patterton et al., 2018), and quite a substantial percentage (25%) were not planning (Vela & Shin, 2019) to share their data.

While many researchers have the intention to make their data open, some researchers assume the library will do data management for them (Higman et al., 2019). There are inconsistencies in the researchers' perceptions and attitudes around data sharing (Vela & Shin, 2019), and it is unlikely that researchers will see the values and benefits of managing their research data for data sharing or preservation and archiving (Kennan & Markauskaite, 2015). These issues could be considered a critical concern around RDM activities, and possibly reflect that data sharing will be delayed until RDM is regarded as a part of gauging research progress (Frederick & Run, 2019).

SITUATIONAL ANALYSIS

As research is becoming highly data and digital driven and with MOSP in place, the national funders of science such as the Ministry of Higher Education (MOHE) and the Ministry of Science, Technology and Innovation (MOSTI) would likely enforce the retention of data for its funded research and provide a framework for data management services by universities and research institutions. However, universities and research institutions in Malaysia (and most probably also other countries in Southeast Asia) are yet to fully implement data management services. RDM in the Malaysian context is a fairly new term for researchers. However, it is not new for academic librarians in research-intensive universities such as Universiti Sains Malaysia (USM). The USM Library established its Research Data Management Unit in 2018 as the first step in developing RDM services. The unit started with three senior librarians and two library support staff. RDM activities conducted in the first-two years were mainly in the form of awareness programmes, development of guidelines and informational materials, and research consultations. In 2020, the unit was upgraded as a new division under the name Research Support Division. The USM Library also set up an Open Science Committee, consisting of USM librarians who had successfully participated in the MOSP Capacity Building programme and chaired by the Chief Librarian.

The empirical literature on the RDM behaviours of Malaysian-based researchers is practically non-existent. Failure to understand the researcher's current practices in managing

research data would make it difficult for the library to initiate RDM services. Successful RDM services require researchers and university management to collaborate, and make their data accessible to the international community. As such, a study was conducted at USM with the objectives:

1. To discover the RDM behaviours and current practices of USM researchers;
2. To examine the RDM services provided by the academic library to support the researchers in managing research data.

This paper presents findings on important issues to consider when developing and implementing RDM services at a university, particularly in Malaysia. The paper also includes information on how to sustain the RDM service once it has been initiated.

METHOD

This is a case study involving both quantitative and qualitative data-gathering. Quantitative data were gathered through a survey, and qualitative data were obtained from interviews. USM academic researchers were invited to participate in this study. A survey instrument was developed incorporating three specific RDM themes: data creation, data storage and preservation, and data sharing, which are the three main stages in the data lifecycle management of any research project (Lefebvre et al., 2020), and detailed by Ashiq et al. (2020) in their systematic literature review on RDM practices and services. The formation and constructs of the questionnaire items were adapted from previous studies reporting researchers' practices and needs on RDM (Rowlands, 2018; Tenopir et al., 2011). The instrument was designed to assess participants' current practices on RDM activities regarding the background of research, data types, metadata, data storage, data preservation, data sharing, and data privacy. The survey included a video clip presenting brief information about RDM followed by a set of multiple-choice questions, Likert-type scales, and close and open-ended questions.

The survey was distributed online to USM researchers through several platforms such as staff e-mails and the library website. A total of 113 researchers responded to the survey. Semi-structured interviews were conducted with 12 USM researchers purposivefully sampled as informants by recruiting researchers from schools in USM from four broad disciplines, namely Health Sciences, Engineering, Science & Technology and Social Sciences. An interview guide was prepared for the interview sessions.

The data collected from the survey were analyzed descriptively, based on frequency counts and percentages. Qualitative data from the interview were analyzed using ATLAS.ti. The complementary use of quantitative and qualitative approaches provides a greater range of insights and perspectives, and permits triangulation or the confirmation of findings by survey and interviews. This may improve the overall validity of results, and makes the findings of better use to the library.

FINDINGS & DISCUSSION

Findings were obtained from a total of 113 researchers who participated in the survey, and they comprised researchers from four broad disciplines: Health Sciences (33, 29%), Engineering (34, 30%), Science & Technology (29, 26%) and Social Sciences (17, 15%). In addition, four researchers from each discipline participated in the interview, a total of 12

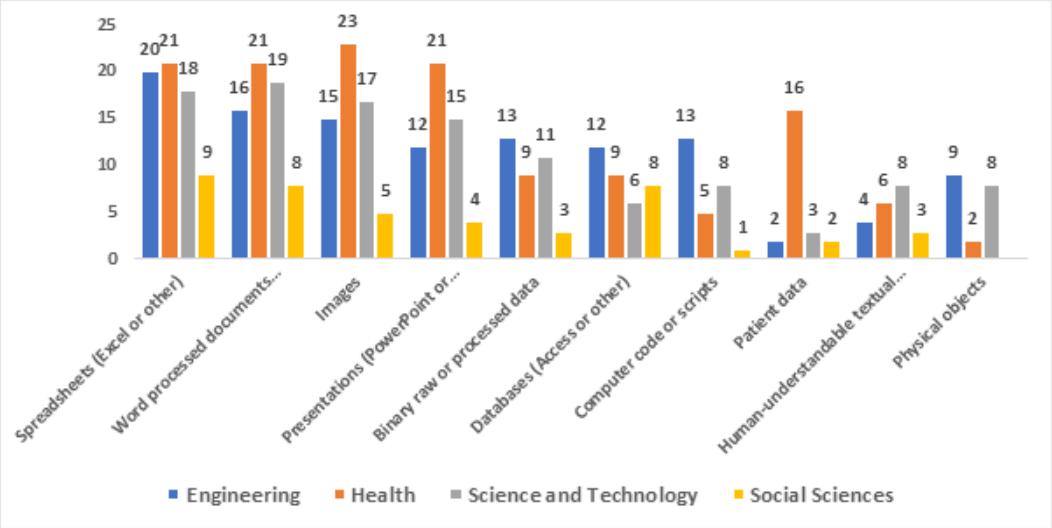

Figure 1. Types of research data produced by researchers based on broad disciplines

informants (R1-R12). Findings were structured around three specific RDM themes: data creation, data storage & preservation and data sharing as described in the next sub-sections.

Data creation

One focus of this study was the types of research data that USM researchers generated, along with the research activities done. Figure 1 shows that most of the respondents from the four broad disciplines mainly produced word-processed documents (64, 57%) and spreadsheets (63, 56%) as research data. Respondents from Engineering generated databases, binary raw or processed data and images, while respondents from Health Sciences produced presentations, images and patient data more than the rest. Science & Technology researchers mostly created presentations, images and binary raw or processed data. Social scientists mainly produced datasets in the form of word-processed documents and spreadsheets. Overall, this finding is consistent with a prior study by Vilar & Zabukovec (2019) who found that researchers mostly produced research data in various forms—either document (text or Microsoft Word files), spreadsheets and presentations.

Metadata creation is the main element in the initial stage of the data lifecycle. Thus, the respondents were asked about creating metadata records for their datasets. Most respondents from all broad disciplines stated that they did not have any metadata records for their research data (42, 37%). Some respondents reported having created the metadata manually (25, 22%), or created the metadata half manually/half automatically (22, 20%). Figure 2 details the responses based on the broad disciplines. The interviews revealed similar findings—all the informants stated that they created metadata records manually and did not adhere to any standards, as illustrated in the following verbatim statements:

> I just created the basic labelling manually for my data such as the title of the project, image name, date before and after data being analyzed. (R4, Health Sciences)

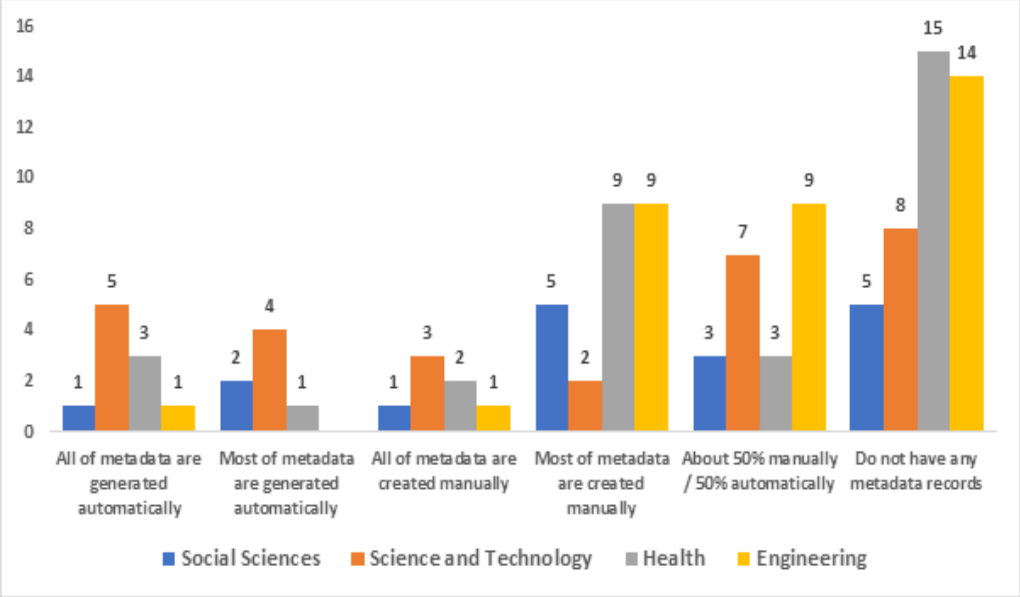

Figure 2. The metadata creation activity by researchers based on broad disciplines

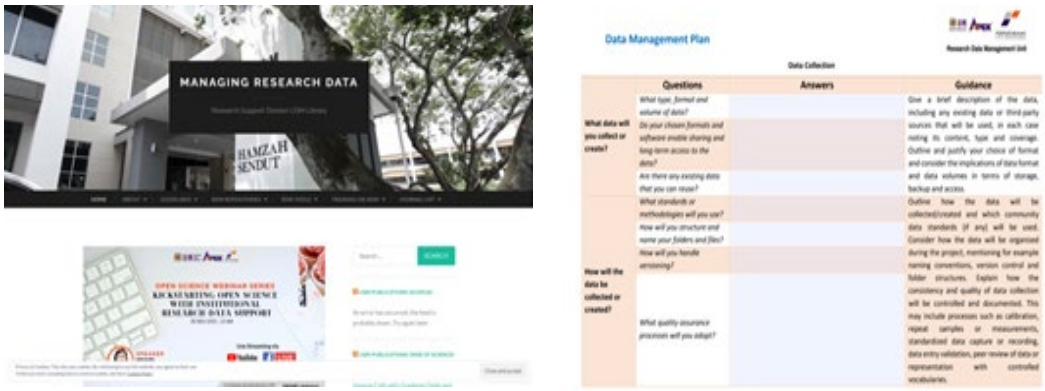

Figure 3. RDM USM Library website and Data Management Plan

They also acknowledged that they did not generate metadata records since they assumed other researchers in the same field would understand the research data when storing and sharing it:

> Folders are created manually and stored in the printed and electronic versions. Folders are created separately for each project. However, researchers in the same field may understand. (R2, Health Sciences)

This study has confirmed that there are researchers who do not create any metadata records for their research data. It was also found by Parsons (2013), Elsayed & Saleh (2018) and Vilar & Zabukovec (2019) that most researchers did not record the metadata for their research data, and only a few researchers created metadata according to the standard and guidelines for data documentation. Some researchers may not be familiar with data documentation, especially in creating metadata for research data, and have limited knowledge about creating metadata (Akers & Doty, 2013; Berman, 2017). Connaway & Dickey (2010) reported that the researchers did not have time to create accurate metadata for their research documents. This is probably true also for research data. Metadata standards are important for guiding the construction of more accurate and consistent metadata. As recommended by

Akers & Doty (2013) and Tripathi et al. (2017), the library could educate researchers in proper data documentation and metadata creation.

In connection with this, the USM Library has come up with a Data Management Plan (DMP), adapted from Digital Curation Centre (DCC, 2013), for the USM researchers. The purpose of the DMP is to provide a checklist to the USM research community for guidance, and the researchers may find helpful informational materials on the Library's RDM website (see Figure 3) when writing DMP, especially in data creation which includes metadata records and data documentation.

Data storage & preservation

Another important aspect of the data lifecycle is data storage and preservation. Researchers should plan beforehand the storage of research data during data collection, and in preserving the data after the research project is finished. The respondents were asked about the size of data storage for the short term (three years or less) and long term (more than three years). Based on Verbaan & Cox (2014), the main difference between short-term and long-term data storage is that short-term storage of research data is for an active project, while long-term storage is for a non-active project. Figure 4 illustrates the size of data storage required by the respondents in the short term. Most respondents from all disciplines stored their research data in less than 10 Gigabytes; Engineering (16, 47%), Health sciences (16, 48%), Science & Technology (12, 41%) and Social Sciences (7, 41%). Only one researcher from Science & Technology stored research data in 10 to 100 Terabytes, and another from Health Sciences stored more than 100 Terabytes. The results indicated that very few respondents stored their research data in 1 to 10 Terabytes.

The interviews revealed that all informants required less than 10 Gigabytes of research data storage. Two informants explained how they stored their research data:

> Our research does not require much space since students are keeping research data. Later on, when they graduate, a copy of research data together with the thesis will be sent to the school. (R6, Science & Technology)

> Data is stored in printed and electronic versions such as flash memories, hard disks and google drive. Apart from it, the lab book will be used as a backup. (R2, Health Sciences)

The results indicated that USM researchers were most likely to store their research data in less than 10 Gigabytes for the short term. However, they required between 10 to 100 Gigabytes for storing data in the long term. This result agrees with Adika & Kwanya (2020), Vilar & Zabukovec (2019) and Chen & Wu (2017) who found that most researchers responded that the size of their data storage need is primarily in Gigabytes. Significantly, very few respondents generated their research data in Terabytes. The estimation of data size can grow from small to larger size depending on the type of research and funding capacity (Abduldayan et al., 2021). It will be reflected in the size of data storage required by the researchers.

The respondents were also asked about data storage required in the long-term. This question was to gauge how the respondents utilized the size of storage for data preservation. Figure 4 shows that the majority of respondents from Science & Technology (11, 38%) and Health Sciences (16, 48%) stored their research data in 10 to 100 Gigabytes for their size of data storage, while the majority of respondents from Social Sciences (5, 29%) and

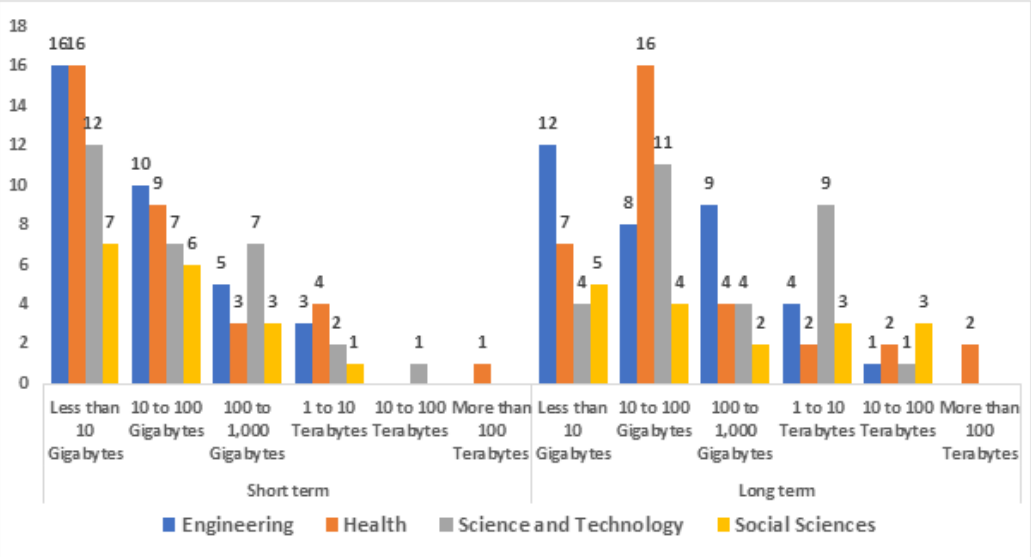

Figure 4. Data storage for the short and long-term

Engineering (12, 35%) utilized less than 10 Gigabytes for the same purpose. Only 6% (2) of the respondents (Health Sciences) utilized more than 100 Terabytes of long-term storage.

Interview findings revealed that the researchers did not dispose of their research data. They acknowledged having their specific ways of storing and preserving their research data:

> I keep all the research data from my first day of doing research until today and never dispose of them unless mentioned otherwise in the funder agreement. (R8, Engineering)

> We keep all data of our projects so far. This method is applied to recent data only (about ten years ago), whereas previous storage practices were not as stringent. So far, we do not find the need to eliminate them. (R2, Health Sciences)

The study found that an overwhelming majority of the respondents across all disciplines preferred an internal Do It Yourself (DIY) or personal solution (such as an external hard disk) to store their research data for active projects; Engineering (33, 80%), Health (31, 79%), Science & Technology (27, 77%) and Social Sciences (15, 68%). The majority of respondents from this study also preferred an internal Do It Yourself (DIY) or personal solution (such as an external hard disk) to store their research data for long-term data preservation; Engineering (32, 78%), Health (31, 78%), Science & Technology (27, 79%) and Social Sciences (15, 71%). Respondents also reported choosing other platforms for both purposes. These platforms are the external digital service or data repository such as Figshare and university-provided service such as Repository@USM. The study indicates that the respondents from all disciplines have stored their research data for active projects; Engineering (3, 7%), Health (5, 13%), Science & Technology (4, 11%) and Social Sciences (6, 27%) and preserved their research data; Engineering (5, 12%), Health (5, 13%), Science & Technology (4, 12%) and Social Sciences (5, 24%) in external digital service or data repository such as Figshare. Moreover, a similar percentage of respondents from all disciplines stored their research data for active projects (13, 9%) and preserved their research data (12, 9%) in their institutional repository such as Repository@USM. Figure 5 illustrates these findings.

Informants in the interview, especially from the sciences mentioned that they utilized multiple platforms such as their personal computers and external hard disks as the central

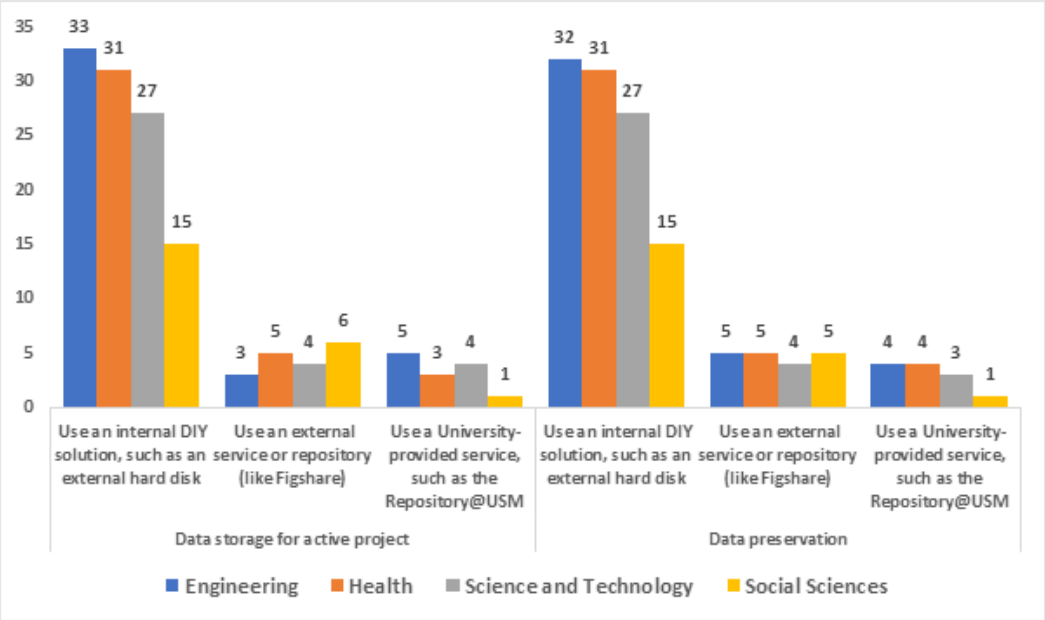

Figure 5. Types of data storage for active project and preservation platforms

platform to store and preserve their research data at the same time. Moreover, a few informants mentioned they did the transition in storing their research data from physical to electronic platforms:

> Due to sensitivity of research data, I just keep all research data in the Pendrive and make them to hardcopy to be used as a backup. (R1, Health Sciences)
>
> I keep all my research data in my computer and also cloud storage such as Google Drive and Dropbox as a backup. (R10, Social Sciences)
>
> At the beginning, I used my computer and hard disk to keep all my research data, and now I begin to use cloud storage such as Google Drive and Dropbox to preserve my data. (R3, Health Sciences)
>
> As a backup and also to preserve my data, I used hard disk, e-mail and cloud storage such as Dropbox and Onedrive. (R5, Science & Technology)

This study has shown that most respondents use personal solutions to store their research data for active projects, such as external hard disks, cloud storage, and personal computer. Very few respondents use the institutional repository and external digital repository such as Figshare. It is consistent with the findings from a recent study in Zimbabwe that most researchers used a personal solution in storing their research data in e-mails, laptops, external hard drives and cloud storage such as Google Drive and Dropbox (J. P. Chigwada, 2021). A recent study in Taiwan found that researchers use their own storage devices and cloud services for long-term research data storage (Lu & Ke, 2020). Buys & Shaw (2015) reported that researchers had chosen more than one solution to store their research data in various ways. Some researchers reported a preference to store their research outputs in cloud storage because of security and accessibility from anywhere as opposed to personal computers, which are more vulnerable to physical theft or damage (Whitmire et al., 2015). The results of this study show that very few researchers stored their research data in the repository, most

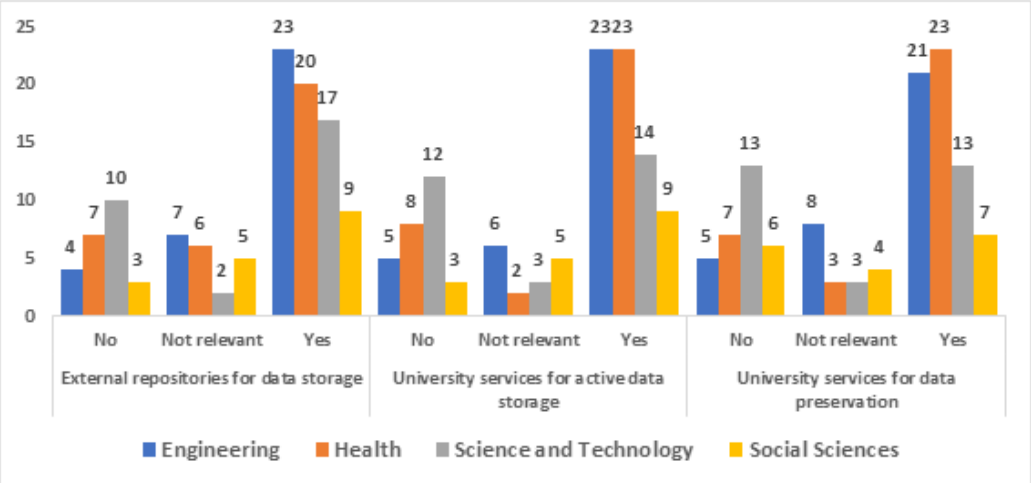

Figure 6. Data storage and preservation services required

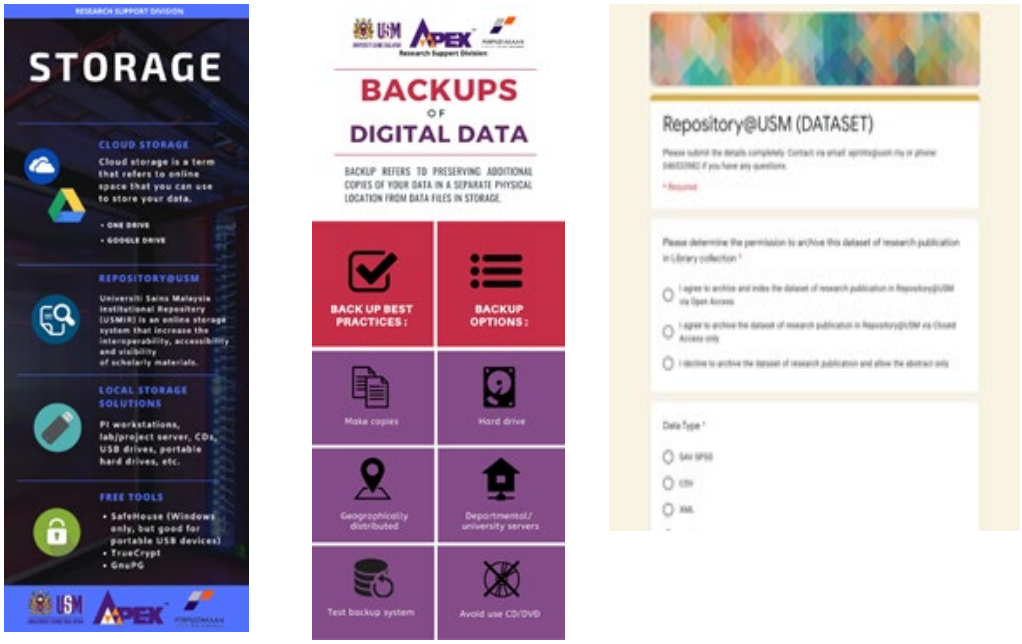

Figure 7. Guideline on data storage & preservation and the online form to self-deposit research data to Repository@USM

probably since RDM is relatively new in Malaysia. This situation may guide the library in providing the data repository as a data storage and preservation for the researchers.

In another question, the researchers were asked about data storage and preservation services that are required. Figure 6 shows that the majority of respondents from all disciplines required external repositories for data storage, university services for active data storage, and data preservation. The requirement for external repositories and university services for data received a similar number of responses (69, 61%), while university services for data preservation has a slightly lower response (64, 57%). Some respondents reported that they did not require any data storage services for storing and preserving their research data. Few respondents also indicated that data storage services were not relevant to them.

With respect to these findings, USM Library has provided a guideline and informational materials on the library RDM website to facilitate the researchers to store and

preserve their research data (see Figure 7). The Library offers the researchers options to choose any platforms they need to store their research data, whether a personal solution or the institutional repository services such as Respository@USM. Even though most researchers have chosen to store their research data in their personal storage, such as cloud storage and external hard drives, the USM Library still encourages them to properly store and backup their research data. If USM researchers need to store and preserve their research data in Repository@USM, the USM Library provides the digital space and guidelines for researchers to self-deposit their research data into the repository. The USM researchers could deposit their research data through the online form provided on the USM Library website, as shown in Figure 7.

Data sharing

Data sharing, an essential component in open science practice, has been discussed in previous literature as the main topic around RDM activities of the researchers. The primary goal of data sharing is to make research data FAIR: findable, accessible, interoperable and reusable. Barriers to data sharing such as sensitive data, privacy and confidentiality have always been voiced in the literature. The respondents in this study were asked about privacy and confidentiality issues associated with their research data. As shown in Figure 8, most respondents from all disciplines had privacy or confidentiality issues associated with their research data (76, 67%). At the same time, less than a quarter of the respondents (25, 22%) did not handle data with some privacy or confidentiality issue. Very few respondents (12, 11%) were unsure whether they handled research data associated with these issues. There was a 45% difference between the group of researchers who handled data that were private or confidential and those who did not.

Information gathered in the interview sessions with researchers highlighted some details about this issue that may discourage data sharing, as illustrated in the following verbatim statements:

> The research data produced from my research is sensitive to ethnic relations in Malaysia. We also have to abide by the conditions highlighted by the university's Human Research Ethics Committee, JPeM-USM. (R3, Health Sciences)

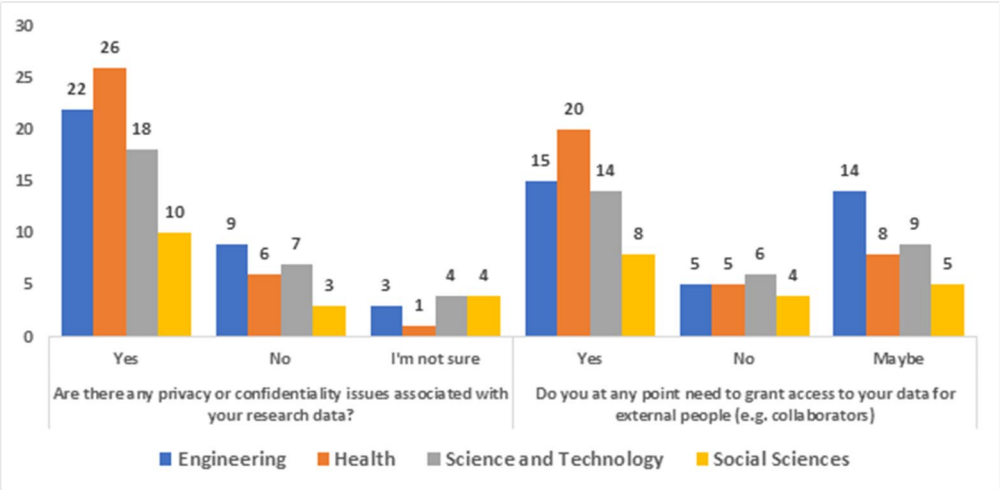

Figure 8. Data privacy or confidentiality and data access issues

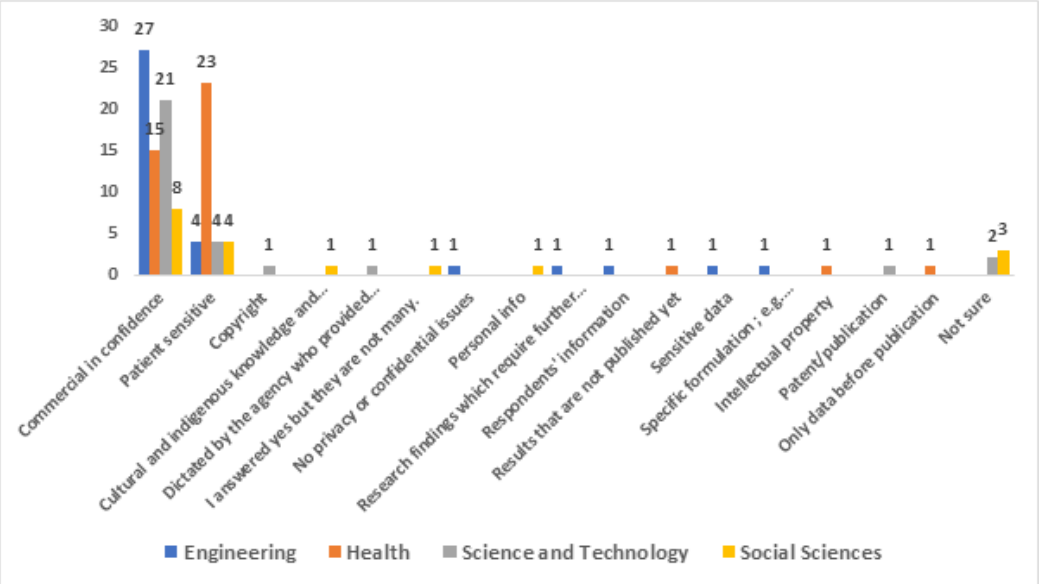

Figure 9. Data privacy and confidential issues associated with research data

> Our data is related to youth behaviour on social media, and it is very confidential and sensitive. (R10, Social Science)

> Data is privacy which is related to human samples. Confidentiality in individual identification data and some findings are proprietary. (R2, Health Sciences)

The respondents also answered a question on sharing or granting access to their research data with external colleagues and peers, such as their research collaborators. Most respondents from all disciplines acknowledged providing access to their research data to external people (57, 50%) (Figure 8). Only a few respondents responded that they did not need to do so (20, 18%). Those who were interviewed stated that they were aware of data access issues. They confirmed that access to the research data was provided only to project members and research collaborators, and that data sharing and access were controlled by the principal investigator.

> Data shared among partners and scientific community (obviously only for relevant information). Other than that, data only can be accessed by project members, research partners, and the scientific community. Restricted data among project members and partners with agreement when needed. (R2, Health Sciences)

> Data will be controlled by Principal Investigators (PI) and will be shared among collaborators only either through e-mail or also Google Drive. (R7, Engineering)

A researcher highlighted the restriction on data access, emphasizing that research data should not be freely accessed to abide by the ethical standards of the research ethics committee:

> Data created are related to patient sensitivity, and it will not be exposed as well as abide by JPeM-USM. I will ensure all my students attend the data handling and research ethics courses to ensure data is handled correctly. In my opinion, all research proposals must be submitted to the JPeM-USM and let them decide whether the research needs to be endorsed and ethically sound. (R1, Health Sciences)

The last question in the survey was on researchers' specific privacy or confidentiality issues associated with their research data. The question gave respondents the option to answer with at least one related issue, such as "patient sensitive" or "commercial in confidence". The respondents could enter any issue in their own words, rather than select from a set of options. Figure 9 shows that most respondents across all disciplines were dealing with commercial information that must be kept confidence (71, 57%), followed by patient sensitive data (35, 28%), which obviously have data privacy and confidentiality issues for researchers in Health Sciences. Other issues related to privacy and confidentiality mentioned were "specific formulation", "intellectual property", "patent", and "cultural & indigenous". Very few (5, 4%) were not sure about privacy or confidentiality issues related to research data.

The survey results show that most researchers were dealing with data privacy and confidentiality, and that they generated research data associated with patient data and commercial information provided in confidence. These types of data can be considered sensitive data. Chigwada et al. (2017) and Rolando et al. (2013) mentioned that research data's access depends on the data privacy classification and specific circumstances. Due to privacy and confidentiality concerns, data sharing of sensitive research data may be non-mandatory, and more stringent access control may be imposed on the data (Thelwall et al., 2020). The management of these data must abide by the ethical considerations of the

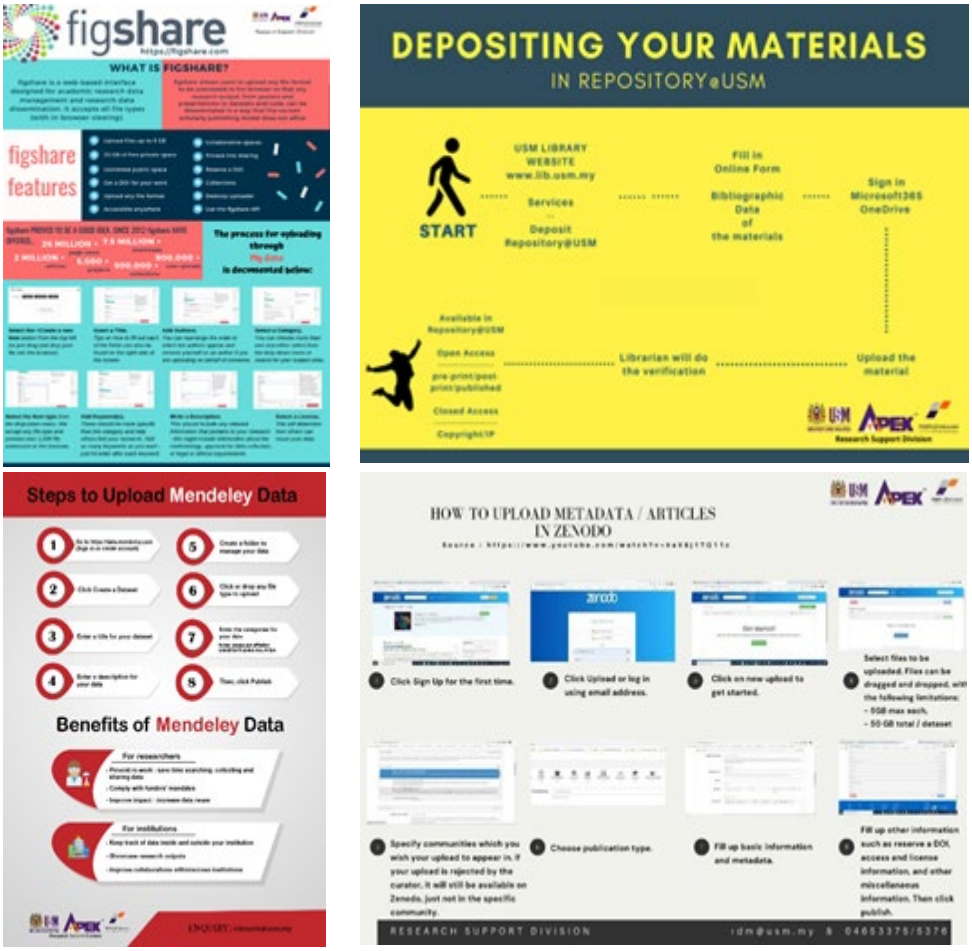

Figure 10. Guidelines of deposit research data in multiple data repositories

Figure 11. RDM and Open Science advocacy programs

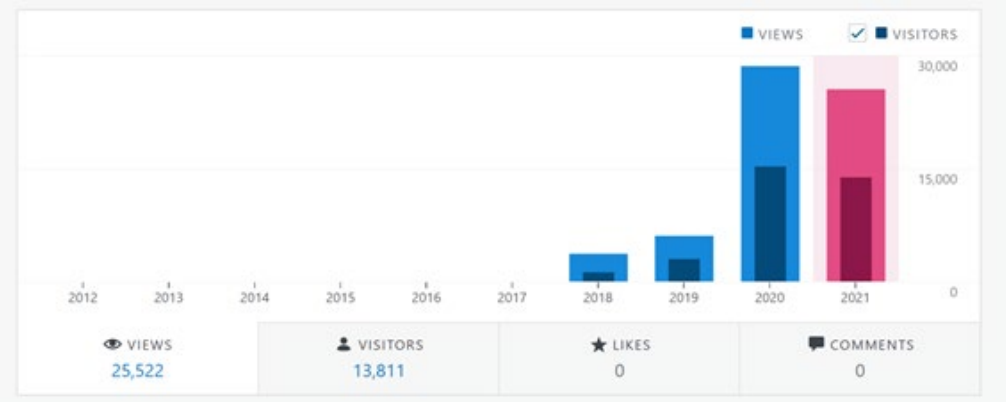

Figure 12. Views and visitors on the RDM USM Library website as of 12th August 2021

university. Indeed, ethical considerations such as security protocols involved and treatment of sensitive data are one of the areas requiring further consideration by researchers in RDM (Renwick et al., 2017). This issue has been identified as the primary reason for the researcher's refusal to share their research data (Elsayed & Saleh, 2018; Joo & Peters, 2020). Joo & Peters (2020) found that 39.3% of the respondents dealt with data that included confidential, proprietary or classified information. This study shows that most researchers wanted to share or grant access to their research data to external people, but were concerned about privacy or confidentiality issues, which are consistent with the issues that previous literature has raised.

Hodonu-Wusu et al. (2020) found that one of the main reasons Malaysian researchers were reluctant to share their research data was due to unclear information on data policy. The library can play a role in providing information on data policy, especially related to data sharing, and ensuring that data that cannot be shared are stored and preserved securely on relevant platforms. Some publishers require researchers to associate their research data with the publication. The USM Library is well aware of the benefits and issues of data sharing for

USM researchers, and responded to this through providing guidelines for USM researchers to utilize several data repositories as secured and trusted data storage and sharing platforms (see Figure 10). The USM Library encourages community-building data sharing initiatives and provides support to USM researchers to store their research data in Repository@USM or external data repositories such as Figshare, Mendeley Data and Zenodo. All these repositories provide open licenses and different types of access for their research data.

An immediate response: RDM advocacy programs

The USM Library has embarked on RDM awareness programs since establishing a new division that focuses on RDM services. An immediate response based on the results of this study was to strengthen the library's awareness activities with advocacy programmes that emphasize RDM benefits—to influence the research community's perspective of RDM as "good behaviour in science" (Chawinga & Zinn, 2019). The USM Library has conducted advocacy programmes such as talks, roadshows, symposiums and webinar series since the pandemic (see Figure 11), actively engaging RDM and open science thought leaders, experts and practitioners from Malaysia and other parts of the world to share their experiences and perspectives about RDM. The first open science symposium organized by the USM Library entitled Open Science @USM received huge participation and positive feedback from the USM researchers, administrative staff and librarians. Moreover, training in RDM has also been included in advocacy programmes, such as searching the research data in data repositories, citation and data documentation using Mendeley, and data analysis using SPSS and ATLAS.ti.

Other than advocacy programs, the Library also established a scholarly blog to provide information about RDM (https://rdmusm.wordpress.com/). This web resource, developed in 2018, is a subset of the main library website. Guidelines on managing research data related to data lifecycle is detailed in this website. There are also recordings of RDM webinars and information about data repositories, RDM tools and training. There are 25,522 views and 13,811 visitors to this website (see Figure 12), and the numbers are increasing.

CONCLUSION

RDM is seen as a new activity for researchers during research works as well as a new service that has the potential to be offered by libraries. Definitely, RDM represents the individual and organizational behavioural changes in managing research data. A long-term strategy for RDM would be required for research data's sustainability, discoverability and accessibility. Eliciting researchers' RDM practices and asking them to share their opinions with the library was an essential foundation for an effective RDM service. Overall, this study has provided a current snapshot of how an academic library establishes RDM services by conducting needs assessment of the academic research community. It demonstrates that researchers from different subject disciplines are all involved in RDM, from data creation to data sharing. This study, however, has also established that these researchers were not applying RDM best practices. Many of these researchers did not have enough experience when deciding on data storage and preservation platforms. They also did not have sufficient backup knowledge and had little experience in describing metadata. Concerns were also raised about privacy and confidentiality issues, although researchers would want to make data accessible and available to external colleagues and peers, including members of their research group.

The outcome of the study further confirms the need for advocacy programs that emphasize the anticipated benefits. Such programs should be included when planning and

implementing RDM services—to raise awareness within the university, and in the longer term result in more of the university's research data being made accessible to the international community. The library can play a central role in RDM services within the institution, and working with other stakeholders in supporting the researchers. This effort could also support the national agenda to make Malaysia's research data a valuable national asset (Malaysia Open Science Platform, 2020). Therefore, it is essential that the library remains actively involved by identifying affordable RDM services and resources that are available to individual researchers.

The study is limited to one case setting; as such, the findings should be treated with caution, and larger studies are needed to confirm these findings. In addition, the study has not covered the comprehensive RDM activities based on the data lifecycle management. Therefore, it is recommended that future studies explore all aspects of RDM activities among the researchers with varied perspectives, including the drivers and influencing factors. Further studies are needed to investigate the effectiveness and impact of RDM activities after implementing the different stages of RDM services by academic libraries.

ACKNOWLEDGMENTS

We thank the two anonymous reviewers whose comments/suggestions helped us improve and clarify this manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Abd Rahman, N. (2019). The need for open science. *Journal of Research Management & Governance*, *2*(1), 22–30. https://doi.org/10.22452/jrmg.vol2no1.3

Abduldayan, F. J., Abifarin, F. P., Oyedum, G. U., & Alhassan, J. A. (2021). Research data management practices of chemistry researchers in federal universities of technology in Nigeria. *Digital Library Perspectives*, *37*(1), 70–90. https://doi.org/10.1108/DLP-06-2020-0051

Abrizah, A. (2019). Malaysian researchers on open science readiness: Call for action. *Exploratory Discourse: Charting the Way Forward For MOSP*, *November*. www.akademisains.gov.my

Adika, F. O., & Kwanya, T. (2020). Research data management literacy amongst lecturers at Strathmore University, Kenya. *Library Management*, *41*(6–7), 447–466. https://doi.org/10.1108/LM-03-2020-0043

Akers, K. G., & Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation*, *8*(2), 5–26. https://doi.org/10.2218/ijdc.v8i2.263

Ashiq, M., Usmani, M. H., & Naeem, M. (2020). A systematic literature review on research data management practices and services. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-07-2020-0103

Aydinoglu, A. U., Dogan, G., & Taskin, Z. (2017). Research data management in Turkey: perceptions and practices. *Library Hi Tech*, *35*(2), 271–289. https://doi.org/10.1108/LHT-11-2016-0134

Berman, E. (2017). An exploratory sequential mixed methods approach to understanding researchers' data management practices at UVM: Findings from the quantitative phase. *Journal of EScience Librarianship*, *6*(1), e1098. https://doi.org/10.7191/jeslib.2017.1098

Borgman, C. L. (2012). Research data: Who will share what, with whom, when, and why? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1714427

Bradley-Ridout, G. (2018). Preferred but not required: Examining research data management roles in health science librarian positions. *Journal of the Canadian Health Libraries Association / Journal de l'Association Des Bibliothèques de La Santé Du Canada*, *39*(3), 138–145. https://doi.org/10.29173/jchla29368

Bunkar, A. R., & Bhatt, D. D. (2020). Perception of researchers & academicians of parul university towards research data management system & role of library: A study. *DESIDOC Journal of Library and Information Technology*, *40*(3), 139–146. https://doi.org/10.14429/djlit.40.3.15302

Buys, C. M., & Shaw, P. L. (2015). Data management practices across an institution: Survey and report. *Journal of Librarianship and Scholarly Communication*, *3*(2), 1225. https://doi.org/10.7710/2162-3309.1225

Chawinga, W. D., & Zinn, S. (2019). Global perspectives of research data sharing: A systematic literature review. *Library and Information Science Research*, *41*(2), 109–122. https://doi.org/10.1016/j.lisr.2019.04.004

Chen, X., & Wu, M. (2017). Survey on the needs for chemistry research data management and sharing. *The Journal of Academic Librarianship*, *43*(4), 346–353. https://doi.org/10.1016/j.acalib.2017.06.006

Chigwada, J., Chiparausha, B., & Kasiroori, J. (2017). Research data management in research institutions in Zimbabwe. *Data Science Journal*, *16*(0), 31. https://doi.org/10.5334/dsj-2017-031

Chigwada, J. P. (2021). Management and maintenance of research data by researchers in Zimbabwe. *Global Knowledge, Memory and Communication*, *ahead-of-p*(ahead-of-print). https://doi.org/10.1108/GKMC-06-2020-0079

Chigwada, J. P., Hwalima, T., & Kwangwa, N. (2019). A proposed framework for research data management services in research institutions in Zimbabwe. In *Research Data Access and Management in Modern Libraries* (pp. 29–53). https://doi.org/10.4018/978-1-5225-8437-7.ch002

Chiware, E. R. T., & Becker, D. A. (2018). Research data management services in Southern Africa: A readiness survey of academic and research libraries. *African Journal of Library, Archives and Information Science*, *28*(1), 1–16. https://www.ajol.info/index.php/ajlais/article/view/174148

Connaway, L. S., & Dickey, T. J. (2010). Towards a profile of the researcher of today : what can we learn from JISC projects ? Common Themes Identified in an Analysis of JISC Virtual Research Environment and Digital Repository Projects Major Themes from VRE and Digital Repository Projects. *Most*, 1–9.

Conrad, S., Shorish, Y., Whitmire, A. L., & Hswe, P. (2017). Building professional development opportunities in data services for academic librarians. *IFLA Journal*, *43*(1), 65–80. https://doi.org/10.1177/0340035216678237

Cox, A. M., Kennan, M. A., Lyon, L., & Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research data service maturity. *Journal of the Association for Information Science and Technology*, *68*(9), 2182–2200. https://doi.org/10.1002/asi.23781

DCC. (2013). *Checklist for a Data Management Plan. v4.0*. Edinburgh: Digital Curation Centre. http://www.dcc.ac.uk/resources/data-management-plans

Elsayed, A. M., & Saleh, E. I. (2018). Research data management and sharing among researchers in Arab universities: An exploratory study. *IFLA Journal*, *44*(4), 281–299. https://doi.org/10.1177/0340035218785196

Fecher, B., Friesike, S., & Hebing, M. (2015). What drives academic data sharing? *PLoS ONE*. https://doi.org/10.1371/journal.pone.0118053

Flores, J. R., Brodeur, J. J., Daniels, M. G., Nicholls, N., & Turnator, E. (2015). Libraries and the research data management landscape. In *The Process of Discovery: The CLIR Postdoctoral Fellowship Program and the Future of the Academy* (pp. 82–102). http://www.clir.org/pubs/reports/pub167/

Frederick, A., & Run, Y. (2019). The role of academic libraries in research data management: A case in Ghanaian university libraries. *OALib*, *06*(03), 1–16. https://doi.org/10.4236/oalib.1105286

Higman, R., Bangert, D., & Jones, S. (2019). Three camps, one destination: The intersections of research data management, FAIR and Open. *Insights: The UKSG Journal*, *32*. https://doi.org/10.1629/uksg.468

Hodonu-Wusu, J. O., Noorhidawati, A., & Abrizah, A. (2020). Malaysian researchers on open data: The first national survey on awareness, practices and attitudes. *Malaysian Journal of Library and Information Science*, *25*(2), 1–20. https://doi.org/10.22452/mjlis.vol25no2.1

Joo, S., & Peters, C. (2020). User needs assessment for research data services in a research university. *Journal of Librarianship and Information Science*, *52*(3), 633–646. https://doi.org/10.1177/0961000619856073

Kennan, M. A., & Markauskaite, L. (2015). Research data management practices: A snapshot in time. *International Journal of Digital Curation*, *10*(2), 69–95. https://doi.org/10.2218/ijdc.v10i2.329

Lefebvre, A., Bakhtiari, B., & Spruit, M. (2020). Exploring research data management planning challenges in practice. *It - Information Technology*, *62*(1), 29–37. https://doi.org/10.1515/itit-2019-0029

Lu, Y. C., & Ke, H. R. (2020). A study on scholars' perceptions and practices of research data management. *Journal of Library and Information Studies*, *18*(2), 103–137. https://doi.org/10.6182/jlis.202012_18(2).103

Majid, S., Foo, S., & Zhang, X. (2018). Research data management by academics and researchers: Perceptions, knowledge and practices. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11279 LNCS*(February 2019), 166–178. https://doi.org/10.1007/978-3-030-04257-8_16

Malaysia Open Science Platform. (2020). *What is Malaysia Open Science Platform?* https://www.akademisains.gov.my/mosp/about/what-is-malaysia-open-science-platform/

Nitecki, D. A., & Davis, M. E. K. (2019). Expanding academic librarians' roles in the research life cycle. *Libri*, *69*(2), 117–125. https://doi.org/10.1515/libri-2018-0066

Parsons, T. (2013). Creating a research data management service. *International Journal of Digital Curation*, *8*(2), 146–156. https://doi.org/10.2218/ijdc.v8i2.279

Patterton, L., Bothma, T., & 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*, *84*(February 2019), 15–27. https://doi.org/10.7553/84-2-1761

Pinfield, S., Cox, A. M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS ONE*, *9*(12), e114734. https://doi.org/10.1371/journal.pone.0114734

Plomp, E., Dintzner, N., Teperek, M., & Dunning, A. (2019). Cultural obstacles to research data management and sharing at TU Delft. *Insights: The UKSG Journal*, *32*. https://doi.org/10.1629/uksg.484

Rebouillat, V. (2017). Inventory of research data management services in France. In C. L. & L. F. (Eds.), *21st International Conference on Electronic Publishing, ELPUB 2017* (pp. 174–181). IOS Press BV. https://doi.org/10.3233/978-1-61499-769-6-174

Renwick, S., Winter, M., & Gill, M. (2017). Managing research data at an academic library in a developing country. *IFLA Journal*, *43*(1), 51–64. https://doi.org/10.1177/0340035216688703

Rice, R. (2019). Supporting research data management and open science in academic libraries: A data librarian's view. *Mitteilungen Der Vereinigung Österreichischer Bibliothekarinnen Und Bibliothekare*, *72*(2), 263–273. https://doi.org/10.31263/voebm.v72i2.3303

Rolando, L., Doty, C., Hagenmaier, W., Valk, A., & Parham, S. W. (2013). Institutional readiness for data stewardship: Findings and recommendations from the research data assessment. *Georgia Institute of Technology*, *June*, 1–32. http://smartech.gatech.edu/handle/1853/48188

Rowlands, I. (2018). *Your research data management needs: Research Data Management Survey results*. https://doi.org/https://doi.org/10.25392/leicester.data.7078127.v1

Syn, S. Y., & Kim, S. (2019). Professional and institutional support for RDM: A case of the National Institutes of Health (NIH). *Proceedings of the Association for Information Science and Technology*, *56*(1), 776–777. https://doi.org/10.1002/pra2.170

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE*, *6*(6), 1–21. https://doi.org/10.1371/journal.pone.0021101

Thelwall, M., Munafò, M., Mas-Bleda, A., Stuart, E., Makita, M., Weigert, V., Keene, C., Khan, N., Drax, K., & Kousha, K. (2020). Is useful research data usually shared? An investigation of genome-wide association study summary statistics. *PLoS ONE*, *15*(2), 1–11. https://doi.org/10.1371/journal.pone.0229578

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

Vela, K., & Shin, N. (2019). Establishing a research data management service on a health sciences campus. *Journal of EScience Librarianship*, *8*(1), e1146. https://doi.org/10.7191/jeslib.2019.1146

Verbaan, E., & Cox, A. M. (2014). Occupational sub-cultures, jurisdictional struggle and third space: Theorising professional service responses to research data management. *The Journal of Academic Librarianship*, *40*(3–4), 211–219. https://doi.org/10.1016/j.acalib.2014.02.008

Vilar, P., & Zabukovec, V. (2019). Research data management and research data literacy in Slovenian science. *Journal of Documentation*, *75*(1), 24–43. https://doi.org/10.1108/JD-03-2018-0042

Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If we share data , will anyone use them? Data sharing and reuse in the long tail of science and technology. *PLoS ONE*, *8*(7), e67332. https://doi.org/10.1371/journal.pone.0067332

Whitmire, A. L., Boock, M., & Sutton, S. C. (2015). Variability in academic research data management practices: Implications for data services development from a faculty survey. *Program*, *49*(4), 382–407. https://doi.org/10.1108/PROG-02-2015-0017