Journal of eScience Librarianship

Volume 8 | Issue 1 Article 3

2019-03-21

Establishing a Research Data Management Service on a Health Sciences Campus

Kathryn Vela
Washington State University

Et al.

ISSN 2161-3974

JeSLIB 2019; 8(1): e1146
doi:10.7191/jeslib.2019.1146

Journal of eScience Librarianship
putting the pieces together: theory and practice

Full-Length Paper

Establishing a Research Data Management Service on a Health Sciences Campus

Kathryn Vela and Nancy Shin

Washington State University, Spokane, WA, USA

Abstract

Objective: Given the increasing need for research data management support and education, the Spokane Academic Library at Washington State University (WSU) sought to determine the data management practices, perceptions, and needs of researchers on the WSU Spokane health sciences campus.

Methods: A 23-question online survey was distributed to WSU researchers and research support staff through the campus listserv. This online survey addressed data organization, documentation, storage & backup, security, preservation, and sharing, as well as challenges and desired support services.

Results: Survey results indicated that there was a clear need for more instruction with regard to data management planning, particularly as data management planning addresses the areas of metadata design, data sharing, data security, and data storage and backup.

Conclusions: This needs assessment will direct how RDM services are implemented on the WSU Spokane campus by the Spokane Academic Library (SAL). These services will influence both research data quality and integrity through improved data management practices.

Correspondence: Kathryn Vela: kathryn.vela@wsu.edu
Keywords: research data management, health science library, needs assessment, biomedical research

Disclosures: The authors report no conflict of interest.

Introduction

Washington State University (WSU) is a land-grant public university with four urban campuses located throughout the state (WSU, n.d.a). The WSU Spokane health sciences campus is home to the university's Colleges of Medicine, Nursing, and Pharmacy. As of spring 2017, there were 1,616 undergraduate and graduate students that were enrolled in Spokane programs, that were taught by 629 full- and part-time staff (WSU 2017).

In 2017, WSU launched its "Drive to Twenty Five" campaign whose mandate is to become one of the top 25 public research universities by 2030 (WSU, n.d.b). As a result, research support and output has been given a high priority. This is evident in the increasing amounts of grant-funded research both institution-wide and on the WSU Spokane campus specifically: In 2016, WSU researchers received over $197 million in both federal and non-federal grant awards. In the 2017 fiscal year, the WSU Spokane campus received $27.5 million (WSU Research, n.d.; WSU Spokane, n.d.).

The Spokane Academic Library (SAL) serves the health sciences students and faculty on the WSU Spokane campus, providing traditional information services, which include reference consultations and instruction. Despite the administrative emphasis on research and research support, SAL has not provided research data management (RDM) support as part of its information services. In terms of existing RDM infrastructure, RDM development at WSU is very basic; researchers have access to an institutional repository, a general RDM LibGuide on data management plans which was developed in 2018, and access to research consultations with a librarian. The limited nature of these RDM services is largely due to lack of staffing and experience; there are only three full-time librarians to support the entire WSU Spokane campus, and none of them had previous experience in research data management prior to this study.

Recently, WSU and SAL have taken steps to remove these obstacles and to improve RDM support services. In January 2018, an RDM librarian was hired on a one-year Sewell Fund fellowship grant to work in the College of Medicine's Technology Incubator (Elson S. Floyd College of Medicine, n.d.). At the same time, one of the existing full-time librarians took part in the inaugural cohort of the Biomedical and Health Research Data Management course for Librarians offered by the National Library of Medicine and the National Network of Libraries of Medicine Training Office (Zhao 2017). An informal environmental scan of RDM support services was performed by the librarian as part of this course, and it revealed a number of gaps in service that could be filled by SAL.

With increased support and improved knowledge, the SAL is now better equipped to provide RDM support to researchers on campus. In order to provide the most effective service model, the library conducted a needs assessment, which determined the current practices, perceptions, and needs of researchers on the WSU Spokane health sciences campus.

Literature Review

RDM has many definitions, but it generally refers to "the set of practices and skills that are needed to work effectively and efficiently with data throughout the research cycle" (Federer 2016, vii). This includes the creation, storage, security, preservation, sharing, and reuse of all

types of data generated by scientific research (Cox and Pinfield 2014). RDM principles and practices have become increasingly important over the last decade for a number of reasons, one of which is the growing ease with which complex data sets can be generated. As stated by Whitmire, Boock, and Sutton (2016), this collection of large data sets is outpacing the "development of the knowledge and skills that are necessary to properly manage them." This can create problems for preservation and retrieval because a failure to adhere to best practices can result in a complete loss of data and a waste of research effort.

Another reason for the increased interest in RDM is the ongoing development of RDM mandates and policies from federal funding agencies and publications. The 2013 Office of Science and Technology Policy (OSTP) memo directed federal agencies to develop policies and procedures to make research data generated by funded projects publicly available (White House Office of Science and Technology Policy 2013). Since then, 16 agencies have complied and released research data sharing policies, including the National Science Foundation (NSF) and the National Institutes of Health (NIH) (SPARC 2016). Many publishers are also starting to require that the underlying data for a publication must be made publicly available; PLOS, Nature, and BioMed Central are a few of the publishers with public access policies (PLOS 2014; Nature 2018; BioMed Central 2018). As Buys and Shaw (2015) point out, these expectations place additional pressure on researchers to engage in and to document RDM practices.

Academic libraries and librarians are well-suited to provide RDM services to researchers faced with these new and growing expectations about research data management and availability. They already possess "expertise and training in skills like metadata, searching, and discovery, archiving and preservation, and knowledge management" (Federer, Lu and Joubert 2016, 52). These are areas that have been identified as core strengths of the library that can be applied to RDM support services (Cox and Pinfield 2014; Faniel and Connaway 2018; Johnson, Butler, and Johnston 2012; Yoon and Schultz 2017). Typically, providing training through educational workshops and consultations is an easy and effective way for an academic library to start RDM services, given that these are activities in which most libraries are already engaged (Akers and Doty 2013; Henderson and Knott 2015; Johnson, Butler, and Johnston 2012; Parsons 2013).

A recurring theme in the literature is that there is no single model for RDM services, but that academic libraries should consider the unique needs of their researchers in conjunction with their own available expertise and resources (Akers and Doty 2013; Henderson and Knott 2015). It is important for libraries to be aware of and to be responsive to researcher RDM needs, especially across different areas of research and career stages (Akers and Doty, 2013; Federer, Lu, and Joubert, 2016). There are many case studies available that describe how different libraries have pursued the development of RDM services that could serve as a model for other institutions (Henderson and Knott 2015; Johnson, Butler, and Johnston 2012; Parsons 2013; Whitmire, Boock, and Sutton 2015; Wittenberg and Elings 2017). However, it is often advised that libraries seeking to build RDM services from scratch should start small, rather than attempting to implement a full suite of RDM services at once (Henderson and Knott 2015; Christensen-Dalsgaard et al. 2012).

Even with the unique circumstances found at each institution, there are several factors that facilitate the development and promotion of RDM services. Faniel and Connaway (2018) identified five factors: technical resources; human resources; researchers' perceptions about

the library; leadership support; and communication, coordination, and collaboration. Forming partnerships with other groups on campus, like campus IT or the Office of Research, often helps to provide a more cohesive RDM infrastructure (Cox and Pinfield 2014; Henderson and Knott 2015; Lyon 2012; Parsons 2013; Wittenberg and Elings 2017). For basic guidance, a 2012 working group from the Association of European Research Libraries designed a list of ten recommendations for libraries who are setting up RDM services on an academic campus.

Needs assessments are a common method of determining the state of RDM on a campus prior to implementing support services through the library because as Parsons says, "the creation of an RDM service is dependent upon understanding the landscape a research project inhabits," (2013, 147). Typically, assessments consist of an online survey that is distributed to faculty and staff who are involved in research on a specific campus or institution. The collected data reveals how researchers view RDM and the types of support they would like to receive (Akers and Doty 2013; Buys and Shaw 2015; Whitmire, Boock, and Sutton 2015). Some assessments will also include other measures like interviews, focus groups, and observational techniques (Anderson et al. 2007; Parsons 2013).

This exploratory study will examine the RDM practices, perceptions, and needs of health sciences researchers on the WSU Spokane health sciences campus. While many studies address the RDM needs of researchers in various disciplines, this needs assessment will survey only health sciences researchers about their RDM needs. The results will be used to design appropriate library RDM support services for the WSU Spokane campus which will be based on the capabilities of the library and the issues highlighted by the respondents.

Methods

Survey Design

The survey design was based on questions used by Buys and Shaw (2015), Read et al. (2015), Whitmire (2015), and DICE (2012). The survey was created using Qualtrics software and consisted of 23 questions regarding the RDM practices, perceptions, and needs of researchers and research support staff from the Colleges of Medicine, Nursing, and Pharmacy (n=307). Participants were asked about the research that they performed and their data management practices, including data management plans, data description, storage, and preservation. It also asked participants about their perceptions and their intentions for data sharing. Finally, participants were asked how they would like to improve their RDM practices and what types of library assistance they would prefer to use. The survey was reviewed by the Associate Vice Dean of Research for the WSU Spokane campus prior to distribution.

Survey Distribution

The survey was distributed through an anonymous Qualtrics link on the WSU Spokane campus announcements listserv. The survey invitation was posted in the announcements twice a week for five weeks, from April 5 to May 11, 2018. The survey was closed after five weeks. In addition, an email with the survey information was sent to the Associate Deans of Research for the Colleges of Medicine, Nursing, and Pharmacy for additional distribution.

Survey Non-Response

Out of the 64 submitted responses, 12 respondents did not supply any responses past accepting the consent form and were therefore excluded from the analysis. Of the remaining 52 responses, 24 respondents left one to three questions incomplete in their final submission. One approach to this issue is to delete the incomplete responses; however, this would reduce the total number of valid responses to a level that would limit the ability to draw valid conclusions. Instead, hot deck imputation was used to resolve the item non-response (Andridge and Little 2010; Cheema 2014; Fox-Wasylyshyn and El-Masri 2005). Hot deck imputation is a statistical method often used on social science research (Cheema 2014) and "involves replacing missing values with values from a 'similar' responding unit," (Andridge and Little 2010, 40). Given that the survey analysis would use descriptive statistics, this was a valid method because it allows for the use of the same sample weight for all survey questions (Fox-Wasylyshyn and El-Masri 2005).

Ethical Considerations

To reduce the possibility of respondent re-identification, the survey did not ask respondents to list their research area or specific lab; given the small size of some labs, this would have made it possible to identify an individual respondent.

The principal investigator received approval by the WSU Institutional Review Board to conduct the assessment under contract IRB No: 16863. Consent was obtained for all survey participants.

Results

The survey was distributed to 307 researchers and research support staff and received 64 responses. After eliminating 12 blank responses and using the hot deck imputation method in order to resolve responses with missing values, there were 52 valid responses for analysis. This resulted in a response rate of approximately 17%; other similar needs assessment surveys have had response rates between 6% and 20% (Akers and Doty 2013; Anderson et al. 2007; Buys and Shaw 2015; Whitmire, Boock, and Sutton 2015).

There was an issue with Question 16, which asked respondents why they did not intend to share their research data if they had indicated this in a previous question. No responses were collected for this question, which points towards a display logic error in the survey. Since no data was collected, this question was not included in the analysis.

Demographics

The majority of the survey respondents reported that they have been involved in research either at WSU or other institutions for 10 or more years (n=33, 63%). The distribution of the respondents between the colleges on the WSU Spokane campus was most concentrated in the College of Medicine, with 52% (n=27) of the respondents indicating that as their place of work; of the remaining respondents, 19% (n=10) work in the College of Pharmacy, and 17% (n=9) work in the College of Nursing. The remaining 12% (n=6) of responses indicated "Other." When reporting their position on campus, 65% (n=34) of the respondents indicated that they

were research faculty, while 12% (n=6) were post-doctoral fellows, 10% (n=5) were clinical faculty, 8% (n=4) were graduate students, and 2% (n=1) were adjunct faculty (Table 1).

Table 1: Demographics of survey respondents.

	n	%
Years of Experience (N=52)		
0-2 years	2	4%
3-5 years	9	17%
6-9 years	8	15%
10 or more years	33	63%
College (N=52)		
Medicine	28	54%
Pharmacy	10	19%
Nursing	8	15%
Other	6	12%
Position (N=50)		
Research Faculty	34	68%
Post-Doctorate Fellow	6	12%
Clinical Faculty	5	10%
Graduate Student	4	8%
Adjunct Faculty	1	2%

RDM Practices

Funding

The majority of the respondents received funding from federal sources like the NIH (n=47, 90%). Funding from non-profit foundations and associations and from WSU each account for 13% of respondents (n=7), while state and industry funding account for 8% (n=4) and 6% (n=3) of the respondents, respectively. 10% of the respondents indicated unspecified funding sources (n=5) and one respondent indicated receiving no funding (2%).

JeSLIB 2019; 8(1): e1146
doi:10.7191/jeslib.2019.1146

Planning

When asked if their research has a data management plan (DMP), 58% of the respondents indicated that they have a DMP in place for their research project (n=30). Of the remaining 42% of responses, 11 respondents indicated that they do not have a DMP for their research (21%), 2 respondents reported that they do not know if their project has a DMP (4%), and 9 respondents do not know what a DMP is (17%) (Figure 1).

The majority of the respondents indicated that they do not have a data safety monitoring board (DSMB) (n=27, 52%), and 31% of the respondents do not know if they have a DSMB (n=16). Only 17% of the respondents reported that they do have a DSMB (n=9). It should be noted that only certain types of human subject research require a DSMB.

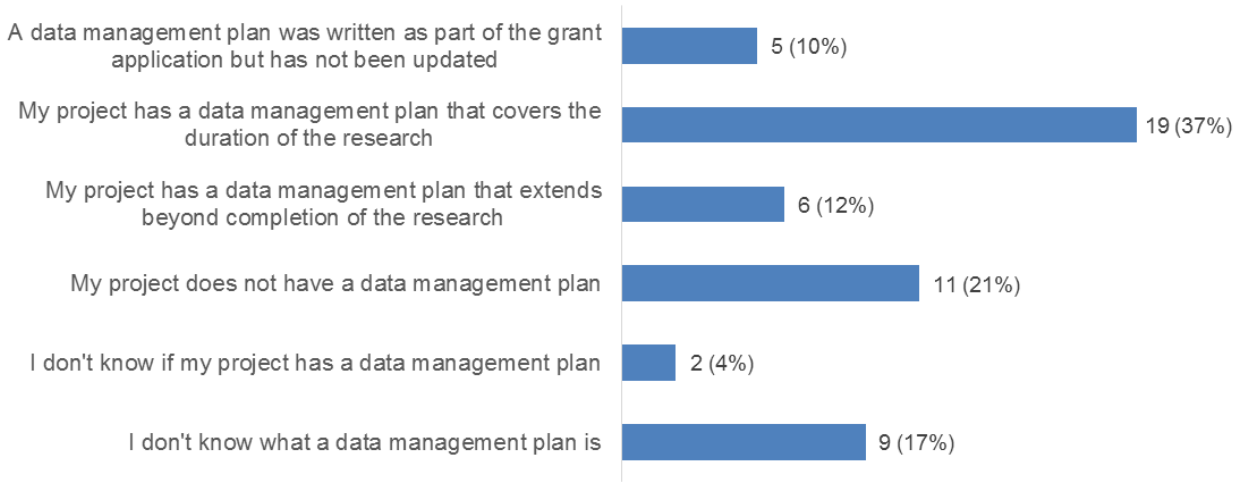

Figure 1: Presence of a data management plan (DMP).

Data Generation

Approximately one-third of the respondents were unsure how much research data they typically generated, or indicated that it did not apply to their research (n=18, 35%). The remaining respondents reported generating a wide range of data, from less than one gigabyte to multiple terabytes (n=34, 65%).

Many types of data were used by respondents (Figure 2), with spreadsheets and text-based data being the most commonly selected data format (n=47, 90%; n=45, 87%). Other types of data that were used by respondents included generic data (i.e. CSV files, etc.) (n=41, 79%), image data (n=34, 65%), statistical data (n=31, 60%), and scanned documents (n=31, 60%). All of the respondents used more than one type of data, with the majority using between four and eight different types of data for their research projects (n=34, 65%). Multiple answers were allowed for this question.

JeSLIB 2019; 8(1): e1146
doi:10.7191/jeslib.2019.1146

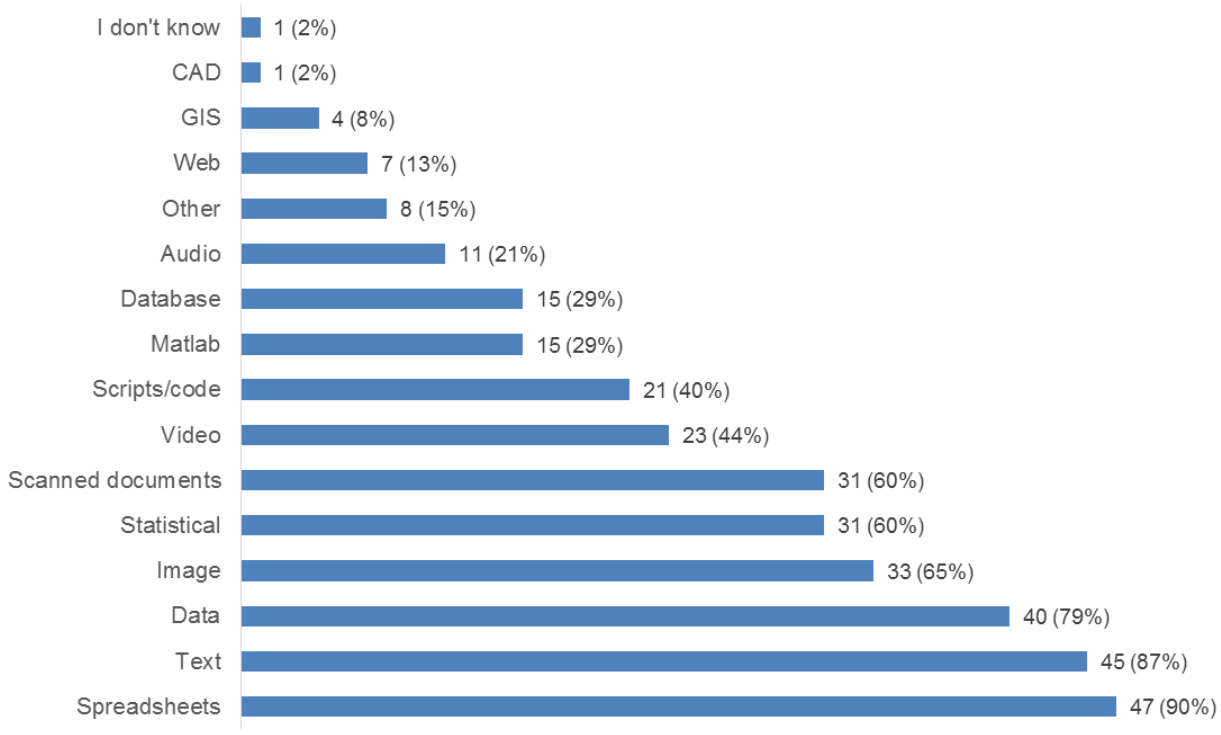

Figure 2: Data types/formats generated by researchers.

Organization

The tools used by survey respondents to record the data generated by their research varied, with 96% (n=50) using electronic documents like Word documents and Excel spreadsheets. More respondents indicated using paper lab notebooks than electronic lab notebooks (n=36, 69%; n=19, 37%). Database software like Microsoft Access was used by 37% of the respondents (n=19), and one of the least popular tools was electronic data capture systems like REDCap (n=16, 31%). Multiple answers were allowed for this question.

Documentation

The most common method used by respondents to describe their research data was through laboratory notebooks (n=32, 62%), followed by standard operating procedures (n=28, 54%). 33% of the respondents reported using codebooks to describe data (n=17), 29% reported using a data dictionary (n=15), and 23% reported using metadata schema (n=12). The remaining 10% of the respondents indicated using a method that was not listed (n=5) (Figure 3). Multiple answers were allowed for this question.

JeSLIB 2019; 8(1): e1146
doi:10.7191/jeslib.2019.1146

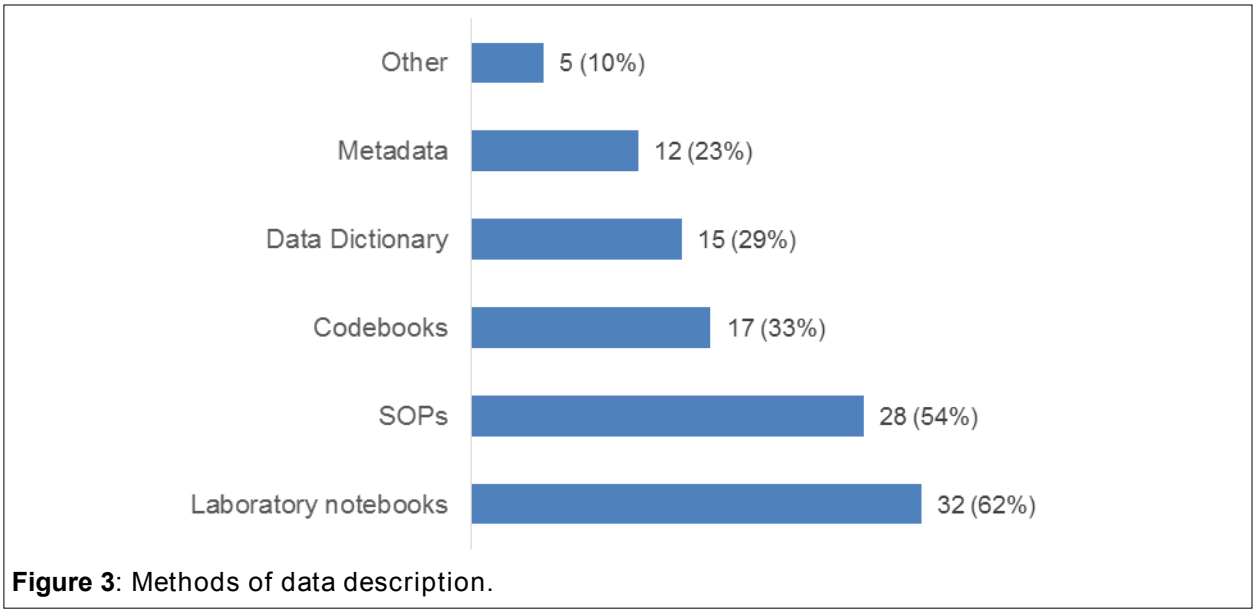

Figure 3: Methods of data description.

Storage

Figure 4 shows that when it comes to storing their research data, the majority of researchers use an institutional server (n=45, 87%), a flash drive or hard drive (n=37, 71%), a paper or electronic lab notebook (n=35, 67%), and/or a lab computer (n=33, 63%). Half of the respondents make use of cloud-based storage (n=26, 50%) and/or a personal computer (n=25, 48%). 17% of respondents reported using CD-ROMs, DVD-ROMs, and/or Blu-ray to store their data (n=9). The remaining 10% of respondents indicated using a storage method that was not listed, like a lab-based server or a dedicated off-site server (n=5). Nearly all of the respondents reported using more than one storage method for their data (n=49, 96%). Multiple answers were allowed for this question.

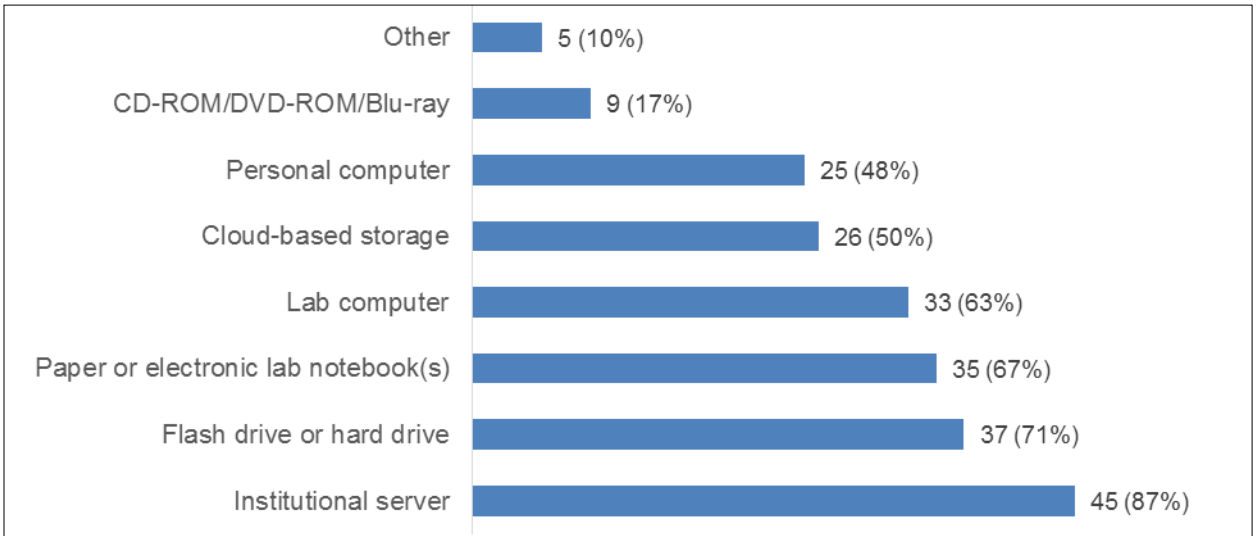

Figure 4: Data storage.

Data Sharing

When asked if they intended to make their research data publicly available at the conclusion of their research project, 38% of the respondents reported that they would share their data (n=20), while 37% indicated that they were unsure about sharing their data (n=19). The remaining 25% said that they do not plan to share their data (n=13).

Of the 38% of respondents who indicated that they do intend to share their data, the majority indicated that they prefer to do so via an institutional repository (n=16, 80%), either at their institution or at a collaborator's institution. Sharing data via email was the preferred option for 70% of the respondents (n=14) and depositing data in a discipline-specific repository was the preference of 45% of the respondents (n=9). Social media was the least preferred option (n=2, 10%). Multiple answers were allowed for this question.

Preservation

Most researchers indicated that keeping research data for 10 or more years after the research project is complete is acceptable in terms of access and usability (n=22, 42%). 21% of the respondents felt that 1-3 years was an appropriate duration (n=11), while 17% felt that 4-10 years was appropriate (n=18, 34%). Only 2% agreed that keeping your research data after research project completion for less than a year was an acceptable duration (n=1).

Security

Figure 5 shows that nearly all of the respondents indicated that they secure and protect their research data through password protection on computers and/or files at (n=51, 98%), while 88% of the respondents reported using access restrictions to computers and/or files (n=46) and 81% of respondents reported using network security (n=42, 81%). Encryption was a security method used by almost half of the respondents (n=24, 46%). 23% of the respondents reported using a physical safe (n=12). 4% of the respondents reported using a data security method that was not listed, like a "dedicated storage room" and "codes" (n=2). Most respondents indicated that they use more than one type of data security method (n=50, 96%), the most frequent combination being password protection, access restrictions, encryption, and network security (n=14, 27%).

RDM Services

When asked what existing data services respondents currently use at WSU Spokane, 19% reported having a librarian consultation at (n=10). 12% of the respondents reported using the WSU institutional repository (n=6) and 2% have used the LibGuide for Data Management Plans (n=1). 6% of the respondents reported using an option that was not listed (n=3); these responses indicated that they consider access to databases and software like PubMed and REDcap as RDM resources, as well. However, 71% of the respondents indicated that they have not used any of the existing RDM services at WSU Spokane (n=37).

JeSLIB 2019; 8(1): e1146
doi:10.7191/jeslib.2019.1146

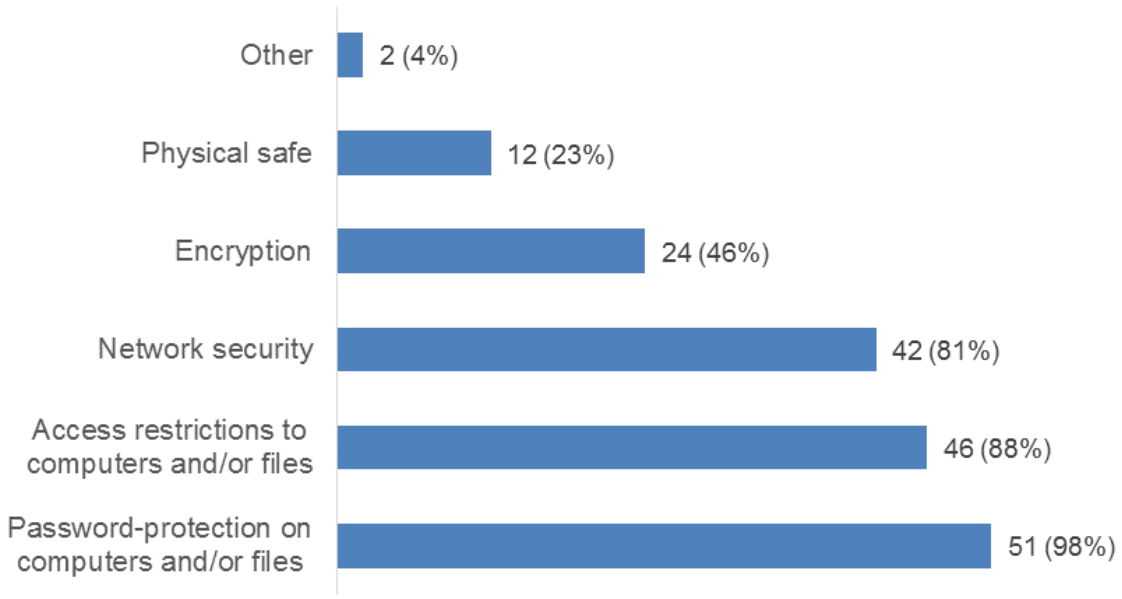

Figure 5: Methods of data security.

RDM Perceptions

When asked if they agree or disagree with whether their research project closely adheres to their DMP, 52% of the respondents indicated that they somewhat or strongly agree with that statement (n=27), while 44% of the respondents were neutral (n=23). Most respondents somewhat or strongly agreed that having established experimental workflows and/or protocols in place was good for the reproducibility of their research (n= 49, 94%). Nearly all of the respondents indicated that they somewhat or strongly agreed that someone should be responsible for data management during the research process (n= 45, 96%), and 93% indicated they somewhat or strongly agreed that someone should be responsible for data management after the research process (n= 44). Approximately two-thirds of the respondents somewhat or strongly agree that it is important to openly share their data with others (n= 34, 65%); 18% of the respondents somewhat or strongly disagree with this statement (n=9), and 17% are neutral (n=9) (Figure 6).

RDM Needs

The majority of the respondents selected “inconsistencies between researchers” and “lack of digital space” as their greatest challenges to storing and accessing data during their research (n=24, 46%, both). Other responses included “security” (n=20, 38%), “software obsolescence” (n=18, 35%), “interoperability” (n=16, 31%), “hardware obsolescence” (n=14, 27%), and “lack of physical space” (n=10, 19%). Four respondents indicated an “other” option that was not listed (8%); responses included “ability to share data,” lack of a clear institutional data policy, coordinating IT needs, and “speed of access” (Figure 7).

JeSLIB 2019; 8(1): e1146
doi:10.7191/jeslib.2019.1146

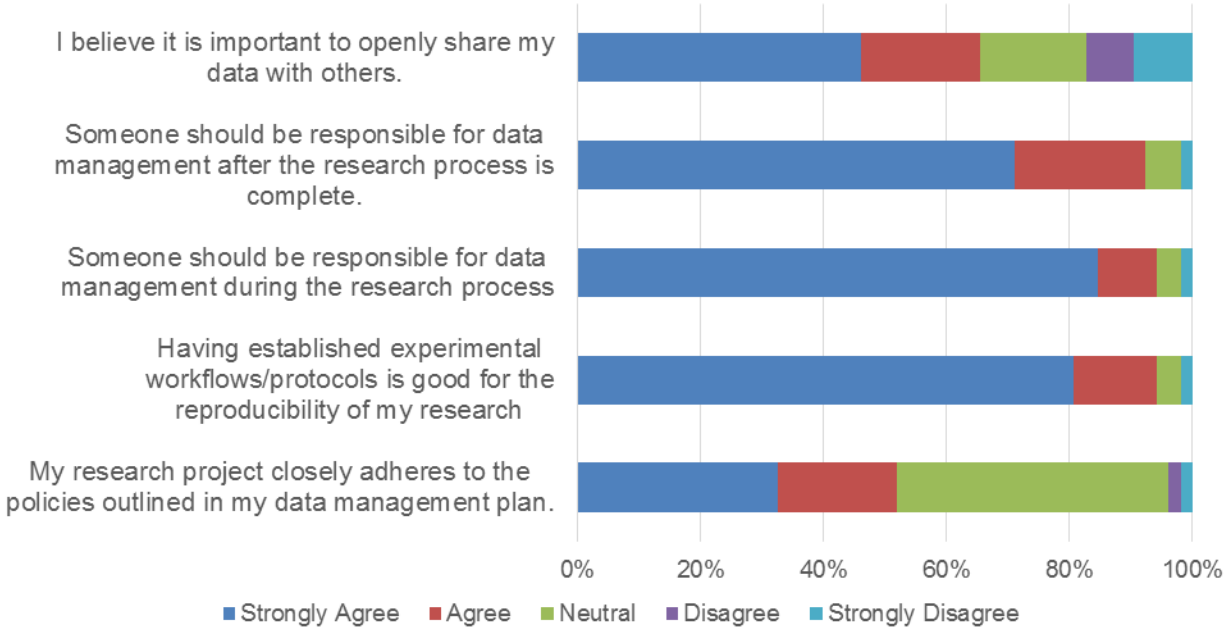

Figure 6: RDM perceptions and attitudes.

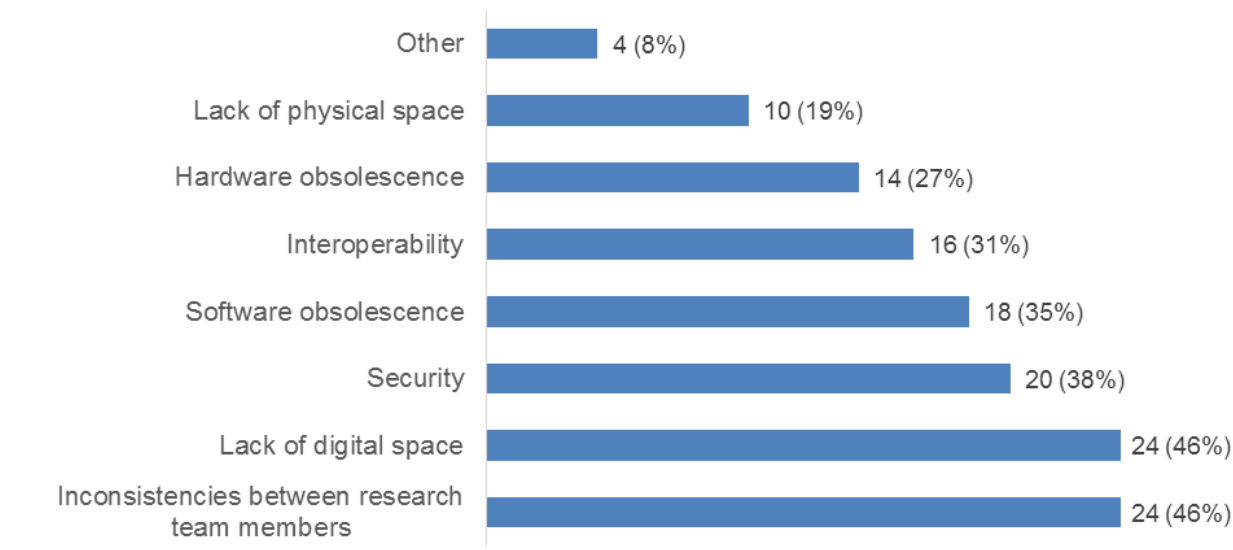

Figure 7: Challenges to storing and accessing data.

When asked in what area of RDM they would most like to improve, most respondents selected “data preservation and back-up” (n=41, 79%) and “data recording and processing” (n=41, 79%). Approximately half of the respondents would like to improve their practices in “data publication and sharing” (n=27) and 44% would like to improve in “data generation and collection” (n=23).

JeSLIB 2019; 8(1): e1146
doi:10.7191/jeslib.2019.1146

The training and support most likely to be used by respondents were one-hour online workshops on specific RDM topics (n=35, 67%), closely followed by data consultations at the beginning of a research project (n=33, 63%). Approximately half of the respondents indicated that they would be interested in an online self-guided series of training modules (n=28, 54%), an online self-guided tutorial (n=28, 54%), and/or one-on-one training (n=26, 50%). 44% of the respondents would be likely to participate in a structured series of online workshops (n=22), 37% would be interested in a structured series of in-person workshops (n=19), and one-third of the respondents were likely to attend 1-hour in-person workshops on specific topics (n=17). Respondents were not interested in half-day or full-day in-person workshops, with only 17% (n=9) and 12% (n=6) of the respondents indicating that they were likely to take advantage of that type of service (Figure 8).

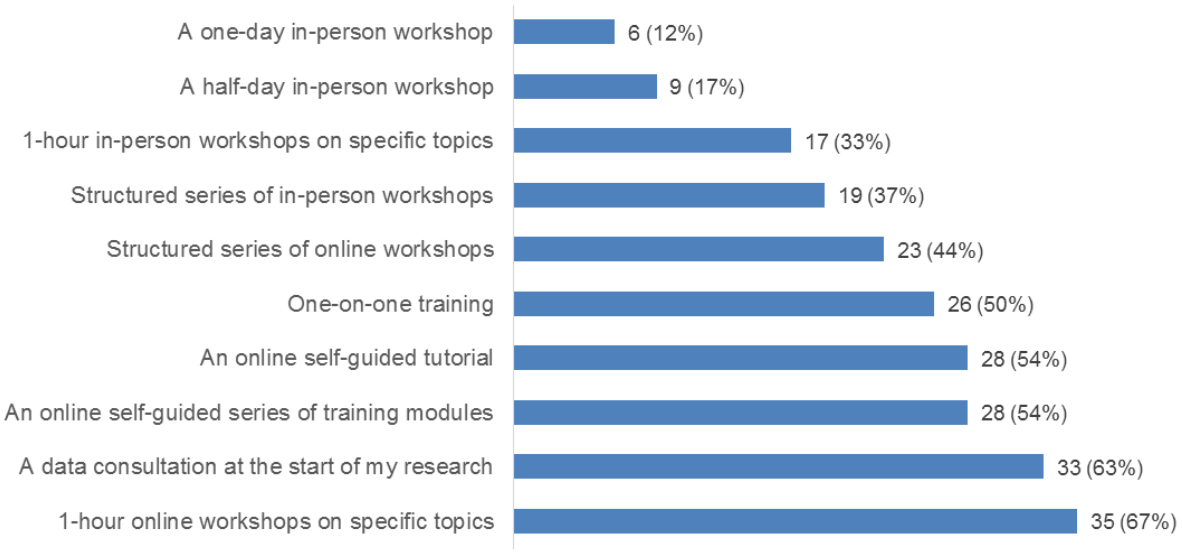

Figure 8: RDM training and support preferences.

Discussion

The results of this study provide valuable information about the RDM behaviors, perceptions, and needs of researchers on the WSU Spokane campus. A small majority of the researchers (58%) reported having DMPs for their research projects; this result is very similar to the study by Buys and Shaw (2015) which reported that 45% of their survey respondents having DMPs. Spreadsheets and text-based data were the most common data types generated by respondents, which is comparable to the results reported by Whitmire, Boock, and Sutton (2015) and Buys and Shaw (2015). Like Anderson et al. (2007), we also found widespread use of non-specialized data applications like Word and Excel to be the most popular method of data collection, which may be because of the relatively short learning curve and inexpensive costs. The greatest RDM needs that were indicated by respondents were data preservation and backup as well as data recording and processing (i.e. data documentation and metadata); this seems to be a common area for improvement (Akers and Doty 2015; Buys and Shaw 2015; Parsons 2013).

The data reveal some interesting inconsistencies in respondent behaviors and attitudes around data sharing. Almost all of the respondents reported having funding from federal sources like the NIH, yet only 38% indicated that they plan to share their data at the conclusion of their research project. This seems to agree with data collected from other RDM needs assessments; Buys and Shaw (2015) found that 34% of the researchers from the school of medicine on their campus share or plan to share their data, and approximately 30% of the medical sciences researchers surveyed by Akers and Doty (2013) planned to share their data. Additionally, when asked how much they agree that it is important to openly share their data with others, 65% of the respondents indicated that they somewhat or strongly agree, which is almost double the percentage of respondents who actually intend to make their data publicly available.

The NIH states that "data should be made as widely and freely available as possible" and requires that grants over $500,000 provide a Data Sharing Plan to describe how they intend to make their data publicly available or why they are unable to do so (NIH, 2003). It is possible that some grant funding that was received by survey respondents falls below that threshold, so they are not required to formally consider sharing their data. There are a number of acknowledged barriers to sharing data, including confidentiality concerns, dataset preparation, and lack of knowledge (Buys and Shaw 2015; Federer et al. 2015; Myers 2016, 60). Key among these barriers in health research is protected health information (PHI), which must be carefully managed under the HIPAA Privacy Rule (U.S. Department of Health & Human Services, 2018); the restrictions that are established by the HIPAA Privacy Rule can make it challenging to make research data publicly available, even if the researcher agrees with data sharing in principle.

Data security is an RDM topic that is not often addressed in the literature. The UK Data Service recommends addressing physical security, network security, and computer security in order to keep sensitive documents like patient information, consent forms, and interview transcripts confidential and secure (n.d.). We found that the majority of respondents secure and protect their research data through a combination of password protection on their computers/files, access restrictions, or network security; most respondents are using all three methods to secure their data. There is little formal data policy at WSU; indeed, this was pointed out by one respondent as a challenge to data storage and access. The WSU IRB manual recommends the use of encryption, but there are no explicit requirements. While it seems that researchers at WSU Spokane are taking steps to keep their data secure, it would be helpful to have a formal policy in place that outlined clear expectations. This is especially important because of the stringent security requirements put in place by the HIPAA Security Rule for protected health information in datasets (U.S. Department of Health & Human Services, 2017).

One component of RDM that emerged as a key area for support is the development of and adherence to DMPs. A large minority of respondents indicated that they did not have a DMP in place, or that they did not have any knowledge of DMPs in their research. This lack of knowledge appears to be causing confusion among researchers about how a DMP functions in a research project; 30 respondents reported having a DMP, yet 27 respondents indicated that they agreed with the statement that their research project closely adhered to their DMP and 23 respondents were neutral. Further analysis shows that of these 23 neutral respondents, 70% (n=16) had previously reported either not having a DMP in place for their research or that they did not know what a DMP is. This confusion seems to be a typical experience for researchers

(Akers and Doty 2013; Buys and Shaw 2015; Federer, Lu, and Joubert 2016; Parsons 2013), and it is a concern because it is recommended to have a DMP in place before research begins and to follow the workflows outlined therein throughout the project (Michener 2015).

Increased focus and support on the development of DMPs may also address some other RDM issues which were revealed by the survey results. One of these issues is inconsistencies between researchers, which translates to a lack of general lab-wide standards for workflow/ research protocol. Another issue identified by the results was a lack of digital space for data storage, which is a fairly common RDM issue across all disciplines (Anderson et al. 2007; Buys and Shaw 2015; Henderson and Knott 2015; Johnson, Butler, and Johnston 2012; Westra 2010). While nearly all respondents reported that they used some kind of documentation method for their research data, only 38% of respondents reported using a metadata schema to describe their data; this is lower than what has been reported in other similar studies (Whitmire, Boock, and Sutton 2015; Tenopir et al. 2012). Finally, a third of respondents indicated that they do not know the amount of data that they generate for their research; this is twice the number of researchers who provided a similar response for an RDM needs assessment conducted at Oregon State University (Whitmire, Boock, and Sutton 2015).

Based on these responses, the library can have the most impact by offering a DMP service where we review and help create DMPs with researchers (Cox and Pinfield 2014; Henderson and Knott 2015; Iwema, Ratajeski, and Ketchum 2016, 95-104; LIBER 2012; Tenopir et al. 2014). This service would encourage researchers to consider the benefits of having a DMP in place prior to the start of their projects like the following benefits: clear expectations, ease of access throughout the data lifecycle, and improved usability of research data (Michener 2015). Given the information that is typically required by a DMP (i.e. types of data, standards, access, etc.), researchers will be better prepared to address the issues that they reported in this needs assessment (DataONE, n.d.). This focused approach to RDM support services is ideal for SAL because of the impending loss of the Fellow at the end of their year of service. As Henderson and Knott say, “finding a focus and sticking to it when the staff and budget are small is important to make the service successful and establishing a niche” (2015).

It is also clear that researchers need further education on RDM practices and principles, and on DMPs in particular. When it comes to instructional delivery, online media were the clear preference; except for data consultations and one-on-one training, online instructional delivery was selected by the majority of the respondents over in-person delivery. This is likely because online delivery allows for more flexibility in scheduling, it provides the opportunity for recorded sessions, and it provides instruction that is more easily targeted to the time of need.

With these RDM needs and instructional preferences in mind, the SAL will provide data consultations to researchers by request and will promote an online LibGuide for Data Management Plans and RDM best practices as an educational resource. A one-hour online workshop will be delivered to researchers during the fall and spring terms; the NYU health sciences library provides resources for a one-hour “RDM Essentials” that will be adapted for use at WSU Spokane (Read and Surkis 2018). If evaluations provide positive feedback, additional online workshops on specific RDM topics may be developed. REDCap training can also be provided to researchers as this is a HIPAA-compliant data collection tool. RDM training can be provided to the other librarians on staff if these RDM services begin to exceed the capacity of a single librarian, in order to provide seamless support to researchers.

To supplement the online instruction, the SAL will also seek to develop partnerships with campus Information Technology department and with the WSU Spokane Office of Research. Since nearly three-quarters of survey respondents indicated that they have not used library RDM resources in the past, partnerships and promotion will be critical to making these RDM services a success. It is hoped that the development and promotion of these RDM services will serve to improve the quality of research on the WSU Spokane campus. We also hope that it will increase researchers' awareness of librarians' skills and create more opportunities for engagement and collaboration (Federer, Lu, and Joubert 2016).

Limitations

There are several study limitations to consider along with the conclusions provided by this data. Researchers were only asked about digital data generated by their research; it is possible that there are other types of physical data being generated on campus that would require different RDM considerations. The results of the survey favor research faculty in the College of Medicine, and do not equally address the RDM practices, perceptions, and challenges for other positions and colleges. The motivations and reasons behind these RDM behaviors and perceptions remain unclear; this information is likely to emerge as more conversations occur between the library and researchers about RDM, but this will occur on an ad hoc basis and will need to be integrated after RDM services have been developed.

Additionally, while needs assessments are a useful tool in determining the practices, perceptions, and needs of the researchers on one campus, the results may not directly apply to other institutions. This is due to the unique nature of each individual institution and to the convenience sample aspect of needs assessments. These results should be considered within the larger body of literature on the topic of developing RDM services on academic health sciences campuses.

Conclusion

This RDM needs assessment indicates that there is a need for improved RDM support for researchers on our health sciences campus. In particular, the need for education and guidance with DMPs emerged as a critical area for support. The all-encompassing design of DMPs implicitly solves multiple issues that arose from the needs assessment, including effective metadata instruction and guidance, data storage and backup, data security, and data sharing. Based on the insights provided by this needs assessment, the SAL will design services to help to improve the quality of research on the health sciences campus and to generate mutually-beneficial collaborations for researchers and librarians alike.

Supplemental Content

Appendix
An online supplement to this article can be found at http://dx.doi.org/10.7191/jeslib.2019.1146 under "Additional Files".

Acknowledgments

This work was supported by the Grace and Harold Sewell Memorial Fund.

References

Akers, Katherine G., and Jennifer Doty. 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

Anderson, Nicholas R., Sally E. Lee, J. Scostt Brockenbrough, Mark E. Minie, Sherrilynne Fuller, James Brinkley, and Peter Tarczy-Hornoch. 2007. “Issues in Biomedical Research Data Management and Analysis: Needs and Barriers.” *Journal of the American Medical Information Association* 14(4): 478-488. https://doi.org/10.1197/jamia.M2114

Andridge, Rebecca R., and Roderick J.A.Little. 2010. “A Review of Hot Deck Imputation for Survey Non-Response.” *International Statistical Review* 78(1): 40-64. https://doi.org/10.1111/j.1751-5823.2010.00103.x

Christensen-Dalsgaard, Birte, Marc van den Berg, Rob Grim, Wolfram Horstmann, Dafne Jansen, Tom Pollard, and Annikki Roos. 2012. “Ten Recommendations for Libraries to Get Started with Research Data Management.” *Association of European Research Libraries*. https://libereurope.eu/wp-content/uploads/The%20research%20data%20group%202012%20v7%20final.pdf

BioMed Central. 2018. “Open data.” Accessed May 27, 2018. https://www.biomedcentral.com/about/policies/open-data

Buys, Cunera M., and Pamela L. Shaw. 2015. “Data Management Practices Across an Institution: Survey and Report.” *Journal of Librarianship and Scholarly Communication* 3(2): eP1224. https://doi.org/10.7710/2162-3309.1225

Cheema, Jehanzeb. 2014. “A Review of Missing Data Handling Methods in Education Research.” *Review of Educational Research* 84(4): 487-508. https://doi.org/10.3102/0034654314532697

Cox, Andrew M. and Stephen Pinfield. 2014. “Research Data Management and Libraries: Current Activities and Future Priorities.” *Journal of Librarianship and Information Science* 46(4): 299-316. https://doi.org/10.1177/0961000613492542

DataONE. n.d. “Plan Data Management Early in Your Project.” Accessed May 20, 2018. https://www.dataone.org/best-practices/plan-data-management-early-your-project

DICE. 2012. “DICE Final Report.” https://lsedice.files.wordpress.com/2012/07/dice-final-report-v-1-2.pdf

Elson S. Floyd College of Medicine. n.d. “Technology Incubator.” Departments, Centers, & Programs. Accessed May 31, 2018. https://medicine.wsu.edu/research/dcp/technology-incubator

Faniel, Ixchel M., and Lynn Silipigni Connaway. 2018. “Librarians’ Perspectives on Factors Influencing Research Data Management Programs.” *College & Research Libraries* 79(1): 100-119. https://doi.org/10.5860/crl.79.1.100

Federer, Lisa. 2016. “Data Management Across the Research Data Life Cycle.” In *The Medical Library Association Guide to Data Management for Librarians*, edited by Lisa Federer, 91-93. Lanham: Rowman & Littlefield.

Federer, Lisa M., Ya-Ling Lu., and Douglas J. Joubert. 2016. “Data Literacy Training Needs of Biomedical Researchers.” *Journal of the Medical Library Association* 104(1): 52-57. https://doi.org/10.3163/1536-5050.104.1.008

Federer, Lisa, Ya-Ling Lu, Douglas J. Joubert, Judith Welsh, and Barbara Brandys. 2015. “Biomedical Data Sharing and Reuse: Attitudes and Practices of Clinical Scientific Research Staff.” *PLoS ONE* 10(6): e0129506. https://doi.org/10.1371/journal.pone.0129506

Fox-Wasylyshyn, Susan M. and Maher M. El-Masri. 2005. “Handling Missing Data in Self-Report Measures.” *Research in Nursing & Health* 28(6): 488-495. https://doi.org/10.1002/nur.20100

Henderson, Margaret E., and Teresa L. Knott. 2015. “Starting a Research Data Management Program Based in a University Library.” *Medical Reference Services Quarterly* 34(1): 47-59. https://doi.org/10.1080/02763869.2015.986783

Iwema, Carrie L., Melissa A. Ratajeski, and Andrea M. Ketchum. 2016. “Library Support for Data Management Plans.” In *The Medical Library Association Guide to Data Management for Librarians*, edited by Lisa Federer, 95-108. Lanham: Rowman & Littlefield.

Johnson, Layne M., John T. Butler, and Lisa R. Johnston. 2012. “Developing E-Science and Research Services and Support at the University of Minnesota Health Sciences Libraries.” *Journal of Library Administration* 52(8): 754-769. https://doi.org/10.1080/01930826.2012.751291

Lyon, Liz. 2012. “The Informatics Transform: Re-Engineering Libraries for the Data Decade.” *International Journal of Digital Curation* 7(1): 126-138. https://doi.org/10.2218/ijdc.v7i1.220

Michener, William K. 2015. “Ten Simple Rules for Creating a Good Data Management Plan.” *PLoS Computational Biology* 11(10): e1004525. https://doi.org/10.1371/journal.pcbi.1004525

Myers, Bethany. 2016. “Research Data as Record.” In *The Medical Library Association Guide to Data Management for Librarians*, edited by Lisa Federer, 31-48. Lanham: Rowman & Littlefield.

National Institutes of Health. 2003. “NIH Data Sharing Policy and Implementation Guidance.” Accessed May 28, 2018. https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

National Institutes of Health. 2015. “National Institutes of Health Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research.” https://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf

National Science Foundation. n.d. “NDF ENG Data Management Plan Requirements.” Accessed May 28, 2018. https://www.nsf.gov/eng/general/dmp.jsp

Nature. 2018. “Availability of data, material and methods.” Accessed May 27, 2018. https://www.nature.com/authors/policies/availability.html

Parsons, Thomas. 2013. “Creating a Research Data Management Service.” *International Journal of Digital Curation* 8(2): 146-156. https://doi.org/10.2218/ijdc.v8i2.279

Public Library of Science. 2014. “Data availability.” Accessed May 27, 2018. http://journals.plos.org/plosone/s/data-availability

Read, Kevin, and Alisa Surkis. 2018. “Research Data Management Teaching Toolkit”. *figshare*. https://doi.org/10.6084/m9.figshare.5042998.v6

Read, Kevin, Alisa Surkis, Catherine Larson, Aileen McCrillis, Alice Graff, Joey Nicholson, and Juanchan Xu. 2015. “Starting the Data Conversation: Informing Data Services at an Academic Health Sciences Library.” *Journal of the Medical Library Association* 103(3): 131-135. https://doi.org/10.3163/1536-5050.103.3.005

SPARC. 2016. “Browse Data Sharing Requirements by Federal Agency”. Accessed October 18, 2018. http://datasharing.sparcopen.org/data

Tenopir, Carol, Robert J. Sandusky, Suzie Allard, and Ben Birch. 2013. “Academic Librarians and Research Data Services: Preparation and Attitudes.” *IFLA Journal* 39(1): 70-78. https://doi.org/10.1177/0340035212473089

Tenopir, Carol, Robert J. Sandusky, Suzie Allard, and Ben Birch. 2014. “Research Data Management Services in Academic Research Libraries and Perceptions of Librarians.” *Library & Information Science Research* 36(2): 84-90. https://doi.org/0.1016/j.lisr.2013.11.003

UK Data Service. n.d. “Data Security.” Accessed May 30, 2018. https://www.ukdataservice.ac.uk/manage-data/store/security

U.S. Department of Health & Human Services. 2018. “Research.” Accessed October 18, 2018. https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html

U.S. Department of Health & Human Services. 2017. “The Security Rule.” Accessed October 18, 2018. https://www.hhs.gov/hipaa/for-professionals/security/index.html

Washington State University. 2017. “Campus Facts.” Accessed May 31, 2018. https://spokane.wsu.edu/about/campus-facts

Washington State University. n.d.a. “Quick Facts.” Accessed May 31, 2018. https://wsu.edu/about/facts

Washington State University. n.d.b. “Drive to Twenty Five.” Accessed May 30, 2018. https://wsu.edu/drive-to-25

Washington State University Research. n.d. “Facts and Figures.” Accessed May 30, 2018. https://research.wsu.edu/accomplishments/facts-and-figures

Washington State University Spokane. n.d. “Research on Campus.” Accessed May 29, 2018. https://spokane.wsu.edu/research

Westra, Brian. 2010. “Data Services for the Sciences: A Needs Assessment.” *Ariadne* 64. http://www.ariadne.ac.uk/issue64/westra

White House Office of Science and Technology Policy. 2013. “Increasing Access to the Results of Federally Funded Scientific Research.” Accessed October 18, 2018. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Whitmire, Amanda L., Michael Boock, and Shan C. Sutton. 2015. “Variability in Academic Research Data Management Practices: Implications for Data Services Development from a Faculty Survey.” *Program: Electronic Library and Information Systems* 49(4): 382-407. https://www.emeraldinsight.com/doi/pdfplus/10.1108/PROG-02-2015-0017

Wittenberg, Jamie, and Mary Elings. 2017. “Building a Research Data Management Service at the University of California, Berkeley: A Tale of Collaboration.” *IFLA Journal* 43(1): 89-97. https://doi.org/10.1177/0340035216686982

Yoon, Ayoung, and Teresa Schultz. 2017. “Research Data Management Services in Academic Libraries in the US: A Content Analysis of Libraries’ Website.” *College & Research Libraries* 78(7): 920-933. https://doi.org/10.5860/crl.78.7.920

Zhao, Shirley. 2017. “Biomedical & Health RDM Training for Librarians: Participants Applications.” *National Network of Libraries of Medicine Training Office*. https://news.nnlm.gov/nto/2017/10/11/biomedical-health-rdm-training-for-librarians-participant-applications