Scientists' Research Data Management Questions: Lessons Learned at a Data Help Desk

Bradley Wade Bishop and Rose M. Borden

abstract: This study investigates scientists' data needs to provide a basis for professionals in research data management (RDM) to tailor services to meet those needs. Eighty-one participants completed a survey after they had asked a question at a Data Help Desk staffed by data management professionals at one of two science conferences. The qualitative responses were coded for common themes, and the quantitative questions were tabulated to compare results between the two conferences. The combined results provide an overview of scientists' questions and training experiences on RDM. The study found that 70 percent of scientists in the survey had no prior RDM training, and the most common RDM need for both ecologists and geologists was storage. This study provides evidence that scientists need additional information about RDM. Traditional service desk models offer one way of assisting them with data needs, but the results indicate a need for more training to meet the expectations of data sharing. Academic and other libraries should consider expanding hours and services for RDM to provide help where and when scientists need it.

Introduction

Scientists have many varying data needs—whether they are aware of them or not. Several information services and tools assist scientists' data needs, but the increased adoption of open science and the FAIR Data Principles—a set of guidelines to make data findable, accessible, interoperable, and reusable—has added new requirements related to research data management (RDM) for scientists.[1] For those creating, collecting, and modeling data, RDM helps curate those digital outputs for reuse, which, in turn, requires discovery and evaluation of fitness for use. To encourage more public-facing scientific research, many funding agencies (86 percent of research councils in the United Kingdom and 63 percent of funding bodies in the United States) require

portal: Libraries and the Academy, Vol. 20, No. 4 (2020), pp. 677–692.

data management plans (DMPs) with grant proposals. These mandates have created more incentives to manage data and digital outputs across domains.[2] These new RDM considerations raise numerous questions for scientists, many of whom may need to complete a DMP for the first time.

…the increased adoption of open science and the FAIR Data Principles—a set of guidelines to make data findable, accessible, interoperable, and reusable—has added new requirements related to research data management (RDM) for scientists.

Although many systems and tools of the research cyberinfrastructure, including human resources, technology, and policies, already facilitate data sharing, individual scientists still might benefit from information services designed to answer their RDM questions. To provide static RDM help for scientists, information professionals have created a large volume of presentations, reports, and documents (http://dmtclearinghouse.esipfed.org/; https://www.dataone.org/education-modules). Several data professionals thought more outreach efforts at science conferences might help to market existing educational materials and assist scientists with their RDM questions. To address these needs, members of the RDM community, led by the Earth Science Information Partners (ESIP), launched a physical help desk, staffed with data experts, at science conferences. ESIP is a federation of organizations that works to make earth science data more discoverable, accessible, and useful to researchers, policy makers, and the public, through efforts by a community of professionals across sectors.

At the 2017 American Geophysical Union (AGU) Fall Meeting, ESIP hosted the first Data Help Desk in the conference exhibit hall. The Help Desk was staffed with volunteers who helped answer data reference questions and offered data-centric workshops as well as tool and platform demonstrations. The organizers deemed the service a success, despite no formal evaluation, and planned to expand the Help Desk to other science conferences. In 2018, the Data Help Desk debuted at the Ecological Society of America (ESA) Annual Meeting. Its partners included the Environmental Data Initiative, which promotes curation and reuse of environmental data; Integrated Digitized Biocollections (iDigBio), which works to make biological specimens available in electronic format; Data Observation Network for Earth (DataONE), which seeks to increase access to earth and environmental data; and the Arctic Data Centre, which handles data management for the Norwegian Meteorological Institute. The Help Desk also repeated at the 2018 AGU Fall Meeting but without the partnerships it had at ESA, due to different types of scientists attending the two conferences. Both Help Desks used a similar staffing model, and some of the same volunteers participated in staffing at both conferences. The volunteers represented many organizations and institutions and included data librarians and subject matter experts from a range of scientific fields.

The information services offered at the Help Desk presented a new take on an old reference service model. Scientists attending the conferences could stop and ask any data questions, much as they would at a traditional reference desk—the only point of service was physical and in person in the respective exhibit halls. The two 2018 Help

Desks included evaluation. This study surveyed scientists after they had asked a question to gather the types of questions posed and other data management-related queries. The analyses reported here will help to inform future staffing as well as facilitate the design of the most relevant RDM training, resources, and reference help for scientists.

Literature Review

Scientists' Data Management Practices and Needs

A data management plan (DMP) is a structured, formal document describing roles and responsibilities for maintaining and managing data during and after a research project.[3] Unlike more established aspects of the research enterprise, such as methods and theory, scientists at all levels of experience and training may need to complete a DMP but lack the necessary training to do so. A DMP may cover anything related to the who, what, where, and when of data, including formats, types, metadata used, storage solutions, access, security, sharing, licensing, roles and responsibilities, and budgeting for all those elements. Research has shown that many scientists have had no formal training in data management as part of their education and may not know where to find answers.[4] In one study, 81 percent of researchers have had no instruction in RDM.[5] This absence of formal RDM training may result in a lack of awareness about the benefits of proper data management. Scientists may not realize, for example, that good data management allows for reproducibility of findings, increases the impact and visibility of research, maximizes transparency, reduces the cost of duplicating data collection, and avoids data loss.[6]

…good data management allows for reproducibility of findings, increases the impact and visibility of research, maximizes transparency, reduces the cost of duplicating data collection, and avoids data loss.

A number of studies have looked at the data management practices, needs, and attitudes of scientists in various disciplines. Some studies focused on graduate students,[7] while others looked at needs across an institution including both faculty and students[8] or explored only the requirements of faculty or professional researchers.[9] The studies of graduate students found that they focus on the specific tasks they need to do and think of data management in terms of the software needed to complete that work.[10] Some specific areas of concern for graduate researchers were data documentation and organization, data sharing, long-term data management, and ownership of data.[11] Another study looked at graduate students' awareness and showed that some of the most common areas where they engage with data are data analysis, data visualization, data conversion, and data cleaning.[12] Graduate students may become so immersed in the details of data management that the bigger picture and in-perpetuity considerations for RDM might not be taught or might not be an immediate concern.

In these studies, most graduate students' knowledge of data management came from their peers or research groups, not from formal training. This inward focus makes

it difficult for most students to learn about standards and practices beyond their research group. Discipline-specific metadata standards already exist in many fields, however, and to share beyond individual projects and labs, these must be utilized. These same issues may also apply to faculty and research staff, since RDM training has not become standard in most science programs. As a result, scientists' knowledge may be limited to what they have learned on the job or sought out informally as needed.

Additional research has explored the effects of recent policy changes at the National Science Foundation and other U.S. federal agencies that require DMPs to be submitted with grant proposals.[13] While DMPs for these agencies must include certain pieces of information, there remains a range of degrees of understanding of RDM in the strategies for data storage and sharing.[14] A survey in 2011–2012 asked librarians and others at universities whether their institutions had policies on research data, whether they thought such a policy was important, and what their own data practices were.[15] Martin Halbert found that 54 percent of respondents kept their own research data on a local computer or external hard drive, storage practices that limit accessibility. Further study is needed to know if a lack of awareness of institutional repositories is a factor in these practices. Another survey asked what skills researchers need to manage their data as well as if their institutions provide the required infrastructure.[16] This study, by Anne Diekema, Andrew Wesolek, and Cheryl Walters, found that faculty were willing to share data and use good data management practices, but were generally unaware of institutional policies and of repositories that were available for them to use. Future work should explore the factors that result in the nonuse of institutional repositories, but outreach efforts like the Data Help Desk may start to offer answers for RDM questions.

…faculty were willing to share data and use good data management practices, but were generally unaware of institutional policies and of repositories that were available for them to use.

Scientists' Information-Seeking Behavior

The information-seeking behavior of scientists in various disciplines related to RDM has been studied increasingly within the last five to eight years.[17] With a variety of methods, including surveys, interviews, and direct observations, a better understanding has emerged of information seeking by scientists in several fields, including mathematicians, astronomers and astrophysicists, computer scientists, aquaculture researchers, chemists, natural science researchers, medical researchers, and engineers. Most of these studies focused on searching for information online and the changing needs of researchers.

In 2012, a survey of 2,036 academic researchers in natural science, engineering, and medicine was analyzed to find patterns of information-seeking behavior and the factors that affect their information-seeking choices.[18] This survey revealed that several different factors affect scientists' preferences when searching for information, including demographics, psychological factors, academic role, and environment. Some questions studied included whether researchers preferred electronic or print resources, what factors determine where they publish their work, and how many hours they spend reading papers in their field. The most important factor in determining information behavior was

found to be the researcher's academic position or role, such as master's or PhD student, assistant or associate professor, or postgraduate or research staff. This result indicates that those providing RDM assistance need to consider the academic role of the person they are assisting and the specific needs that individual may have.

Most studies focus on researchers in specific disciplines rather than surveying a range of fields. One study surveyed 288 Indian astronomers and astrophysicists regarding their information needs and information-seeking behaviors.[19] It found that the majority of researchers look for information informally, such as from e-mail lists or peer conversations, and prefer to find sources online. In 2018, 231 academic chemists were asked about their information-seeking behaviors and attitudes.[20] The results showed that chemists can be overwhelmed by the amount of information in their field and feel they lack the time to keep up with it all but wish they were more aware of new technologies in their fields. A survey of geoscientists at the Geological Survey of Norway in 2018 showed that they used Web searching and asking colleagues more often than bibliographic databases.[21] The reliance on colleagues as trusted and available resources for help is a common theme, but future work should study scientists' awareness that data librarians may assist with their RDM needs.

Data Management Training and Support for Scientists

A number of papers have looked at the needs of graduate students and best practices for training them in data management. Some of these papers describe workshops and efforts by librarians to educate graduate students at their universities.[22] In some cases, regular for-credit classes were developed and offered to graduate students, often involving librarians along with other faculty members.[23] Other training formats included workshops, seminars, and embedded librarian services to help with RDM queries within departments.

…data management training has begun to be recognized as important for researchers at all levels in higher education.

While it is important to provide training for researchers beginning at the graduate level, many scientists in other roles, such as faculty members and research staff, need to receive training after they have completed their formal higher education. At some institutions, librarians provide RDM instruction and services to researchers within certain disciplines through collaboration as subject liaisons, or to the whole university through workshops offered at the library.[24] These efforts show that data management training has begun to be recognized as important for researchers at all levels in higher education.

Beyond academic institutions, professional societies and organizations within disciplines also recognize the need for data management training and have started offering their own workshops to help scientists gain data management skills.[25] The studies of the need for and delivery of RDM training and education address many questions scientists now have about DMPs, but little research has investigated the RDM questions scientists have outside such formalized training settings. This article addresses this gap in the literature by gathering not only the types of questions asked of the Data Help Desk but also scientists' other data management-related queries.

Methodology

The purpose of this study was to assess scientists' data management needs through the questions they asked related to RDM at a Data Help Desk. After scientists had asked a question, the authors administered a voluntary, Institutional Review Board-approved survey to the questioners. All volunteers staffing the Data Help Desk were trained to direct those with questions to the survey.

The survey questions provided feedback on the information service, questions asked, and details about educational attainment, RDM needs, and data management training. Everyone who asked a question at the Data Help Desk was directed to the survey, given informed consent, and if they agreed to participate in the evaluation were asked to complete these survey questions:

1. What was your question for the Data Help Desk?
2. What is your highest level of education? PhD; master's; bachelor's
3. What is your current job title?
4. In your job, what are your top three research data management needs?
5. If you have received training, what types of data research management training did you receive?
6. If you have not received training, what types of data research management training would you be most interested in?
7. What delivery formats for training do you prefer (for example, half day workshop, webinar, graduate coursework)?
8. What was your level of satisfaction with your service at the Data Help Desk?

The last question used a Likert scale on which individuals could express their degree of satisfaction.

Question-askers who agreed to take part in the study completed the survey on mobile devices near the Data Help Desk. The individuals helping the scientists with their questions did not administer the survey. The surveys were conducted at a distance from the information service workers to reduce any influence on their responses. With all types of question-askers, from students to early career scientists to more senior researchers, the data provided some context to the questions resulting from data management mandates and more general queries about data management.

Data Analysis

Eighty-one participants completed the survey after asking a question. Of these responses, 43 were collected at the ESA meeting in New Orleans, Louisiana, in August 2018. The ESA 2018 exhibit hall was open seven and a half hours each day for four days (30 hours total). The AGU annual meeting in Washington, D.C., in December 2018 yielded the other 38 responses. The authors collected replies over five days, starting with a two-hour kick-off opening of the exhibit hall, followed by three full days with seven and a half hours, and ending with a shortened last day of only four hours (28.5 hours total). Although the Data Help Desk was included in conference programs and promoted through social media, most conference attendees explored the exhibit hall without prior knowledge of the Help Desk. The idea of data help inside an exhibit hall might take multiple attempts to increase awareness.

Analysis of the data collected was done by first coding the answers to qualitative or open-ended questions for common themes or keywords. The responses to quantitative or multiple-choice questions were tabulated to compare the results. Results were compared between the two conferences, as well as combined to get an overall picture of geologists' and ecological scientists' questions and training experiences with RDM.

Limitations of the Study

The survey was conducted inside the exhibit halls near the Data Help Desk. More people stopped by the desk at ESA and AGU than participated in the survey. Because the desk was within an exhibit hall among other exhibitors, many people wanted only stickers or other swag and not substantial help. An exhibit hall includes booths representing research organizations, university programs, funding agencies, publishers, retailers, and a variety of other exhibitors, most of whom provide information that relates to a product, tool, degree program, or service. The Data Help Desk was unique in that it provided generalized data help like that given by a library reference desk. The data collection within an exhibit hall and the inherent limitations of operating a reference desk in that venue complicated the study methods. AGU took place in the Walter E. Washington Convention Center, and ESA in the Ernest N. Morial New Orleans Convention Center, both large exhibit halls. Therefore, direct observation of the question-answer negotiation was not possible given the difficulty of hearing transactions, multiple transactions occurring at once, and the challenge of accurately counting the attendees passing through the exhibit hall. Finally, although all volunteers staffing the Data Help Desk were aware of the research project and asked to promote survey participation to scientists at the end of their transaction, some volunteers were more consistent in recruitment than others. After nearly 60 hours of data collection, at two different science conferences, a snapshot of RDM questions was captured. The results are not generalizable but may be informative for ecologists and geologists and those helping them with RDM.

Results

The most common questions asked at the Data Help Desk involved the various organizations represented by ESIP, as well as questions related to general data management, data sharing, and data storage (see Table 1). Scientists also wanted to know about data analysis, finding and accessing data, and what software is available to help with data management. Many did not have specific questions and benefited most from hearing about RDM and existing resources.

The most common education level at each conference was a PhD at approximately 47 percent (36). About 20 percent of participants had a master's degree (15), and 34 percent a bachelor's degree (26). One participant did not have any degree. The most common job title was student (35 percent), then research scientist (28 percent) and professor (13 percent). Students, scientists, and professors were not probed for more specificity as this was a self-administered online survey on a mobile device. There were a few postdocs at ESA, but none at AGU, or at least none who identified themselves as such (see Table 2). Several more specific job titles were provided that do not map well to the broader categories: a state Department of Environmental Quality, head of R&D,

Table 1.

Types of questions asked at the Data Help Desk

Topic	Ecological Society of America	American Geophysical Union	Combined
Organizations	10	7	17
Data management	9	0	9
Sharing data	6	4	10
Data storage	6	5	11
Data analysis	4	1	5
Finding data	3	5	8
Other	3	9	12
Accessing data	2	4	6
Software	0	3	3
Total	**43**	**38**	**81**
Examples of other questions	General information; Can I help by hosting a data workshop?	What are you doing here? What are data? How to get involved	

Table 2.

Title	Ecological Society of America	American Geophysical Union	Combined
Student	12	15	27
Research scientist	11	11	22
Postdoc	6	0	6
Professor	6	4	10
Other	5	7	12
Total	**40**	**37**	**77**

Table 3.

Ecological Society of America Category	1	2	3	Weighted	American Geophysical Union Category	1	2	3	Weighted
Data storage	8	10	4	16.0	Data storage	11	2	0	12.3
How to do data management	7	5	4	11.7	Data accessibility	5	6	0	9.0
Data access	6	2	1	7.7	Interoperability (readability, standardization)	2	3	4	5.3
Data sharing	5	1	2	6.3	Sharing data	1	5	3	5.3
How to do data analysis	2	3	5	5.7	Data organization	0	1	1	1.0
Reproducibility	0	3	1	2.3	Finding data	3	0	0	3.0
Cleaning data	2	0	0	2.0	Metadata	1	2	1	2.7
Metadata	0	2	1	1.7	Databases	1	0	1	1.3
Resources	1	1	2	2.3	Collecting data	1	0	0	1.0
Quality control	1	0	0	1.0	Data processing	0	1	0	0.7
Visualization	0	0	1	0.3	Data availability	0	1	0	0.7
Privacy	0	0	1	0.3	Visualization	0	0	1	0.3
					Restricted data	0	0	1	0.3
Total	**32**	**27**	**22**		**Total**	**25**	**21**	**12**	
Others	4	4	5		Others	6	7	2	

biological science technician, manager, program coordinator, public affairs specialist, president of First Peoples' Council, program officer, project coordinator for Cooperative Weed Management Area, policy fellow, principal systems engineer, and manager of departmental computing.

The most common data management need at both ESA and AGU was data storage. Participants wanted more information on where they can deposit their data. Data accessibility and data sharing were also in the top five for both conferences. General data management information was higher at ESA, while scientists at AGU more commonly mentioned interoperability of data as a need. Data analysis was also a concern at ESA. This likely resulted from ESA including iDigBio and other interactive talks that focused on how to analyze data and left listeners wanting to know more.

Needs with fewer responses at ESA were reproducibility, cleaning data, metadata, resources, quality control, visualization, and privacy. A number of responses were not easily categorized and are listed under "Other" in Table 3. Needs with fewer responses at AGU included data organization, finding data, metadata, databases, collecting data, data processing, data availability, visualization, and restricted data. Other responses did not easily fit into a category. Table 3 presents the listed data management needs, and those needs listed first were weighted the most.

Although 3 were unsure, 54 of 79 (68 percent) of the respondents reported receiving no prior data management training. Of those who had, the most common format was a workshop (7). Other common formats were seminars at AGU or in graduate school, and being self-taught. The only general topics of these trainings mentioned more than once each were Data Carpentry, a program that provides data skills instruction (4), and R, a free statistical programming language (2). Survey participants received a long list of training topics from which to choose: image data time series; spatial data; graduate seminar; data life cycle; research group on the job at Jet Propulsion Laboratory (JPL); Python, a general-purpose programming language; modeling; data management basics; Science Base; iDigBio; Unidata committees, which promote the sharing of geoscience data; programming; reproducibility; metadata; data management; statistics; GitHub, which provides version control; provenance; automating scripts; and software. Without probing, it is impossible to know what knowledge, skills, and abilities the participants retained from these trainings. The variety of responses to this survey question showed that participants lacked a clear understanding of what RDM training entails.

When asked what format they would like for future trainings, the most common response at AGU was webinar (13), with workshop (6) listed second most. At ESA, the most common responses were workshop (5) and online (5), which could be construed to include webinars. Very few listed face-to-face or written tutorials as continuing education preferences. None suggested information services or data consultations like the Data Help Desk.

Table 4 presents the preferred topics of future trainings. The most common responses at AGU were data archiving, database management, data organization, and documentation/nomenclature. The most common responses from ESA were data management, data archiving, DataONE, data analysis, and databases. Listed only once were data processing; background research; coding; data cleaning; Python; version control; available services; data access; big data; finding data; and science data communities.

Table 4.

Topic	Ecological Society of America	American Geophysical Union	Combined
Data archiving	4	5	9
Data management	6	2	8
Database management	0	5	5
DataONE	4	0	4
Unsure	0	4	4
Data analysis	3	0	3
Databases	3	0	3
Sharing data	1	2	3
Data organization	0	3	3
Documentation/nomenclature	0	3	3
Total	**26**	**31**	**57**

Overall, 88 percent of respondents said they were highly satisfied with the help they received, and another 8 percent were somewhat satisfied. One respondent was neither satisfied nor dissatisfied, and two were somewhat dissatisfied. Future evaluation should explore expectations as well as satisfaction.

Discussion

Scientists have become more aware of RDM issues and more invested in improving data findability, accessibility, and interoperability for data reuse. This study shows that many visitors to the Data Help Desk had received no formal RDM training. The questions asked and top RDM needs presented in the survey show many scientists still do not know what data management is or need more information to understand its complexity. This discussion will explore the future roles of information professionals in data help services, the educational implications resulting from these new roles, and how library research on reference may inform future data help services.

…many scientists still do not know what data management is or need more information to understand its complexity.

Types of Questions

Overall, most users learned about ESIP and other partners staffing the Data Help Desk, even if they did not have a specific data management question. These outreach efforts may not immediately bear fruit but may increase awareness of RDM and benefit those

involved. The concept of someone in an exhibit booth providing immediate and free help is unusual in scientific conference exhibit halls, where most exhibitors try to sell services or products. It might take years of effort and substantial word of mouth to have the Data Help Desk become an integrated and known entity within an exhibit hall.

A majority of the questions were basic, suggesting that it would be good to provide brief handouts or a rolling video screen explaining and extolling the benefits of RDM. Many students and researchers still need education at this fundamental level. Although open science policies and culture have begun to move toward adoption of open data and FAIR Data Principles to increase sharing, the offering of help may be ahead of its time. Based on these survey responses, RDM training has not spread widely in the disciplines represented at these two conferences and likely in others worldwide. In short, the Data Help Desk helped reveal information needs that most conference attendees did not realize they have or should have.

Top RDM Needs and Training Interests

Most question-askers wanted information on where they can deposit their data. This points to the need to make all researchers more aware of institutional and subject-specific repositories for long-term storage of data. Funders may have specific requirements for the repositories and long-term storage plans for data collected with their funds. In addition, many scientists need more information about the qualities that make for a trustworthy repository, why they need a long-term storage option, or which of the repositories available to them would be the best choice for their data set. Data storage concerns and solutions would be an important priority for future data management trainings for researchers and a common question asked of any data information service.

Most question-askers wanted information on where they can deposit their data. This points to the need to make all researchers more aware of institutional and subject-specific repositories for long-term storage of data.

Question-askers at both conferences were also interested in data sharing and accessibility. This shows that researchers have a growing awareness of the importance of making data open for everyone but may not know the best way to make their own data easily accessible. It is important to support efforts to share data, because reproducibility of data and results is increasingly necessary among all scientific research today.

…researchers have a growing awareness of the importance of making data open for everyone but may not know the best way to make their own data easily accessible.

Around 70 percent of scientists in the survey had no prior RDM training. RDM topics of all kinds should be in demand, and despite the volumes of training available, more may be required. At least, scientists must be made aware of the educational materials that already exist. Question-askers suggested webinars or in-person workshops. Webinars

could easily be set up and hosted by information organizations or libraries, and workshops could also be organized in libraries for local researchers or at conferences like AGU and ESA.

Conclusion

The Data Help Desk is a much-needed information service, but a single point of access will not meet demand. Comprehensive RDM training is needed for all researchers from undergraduate and graduate students through senior-level scientists based on the lack of training reported by the question-askers at these two science conferences. As demand increases, there is a growing role for academic and other libraries to expand hours and services for RDM. Like the history of reference desks, RDM and data-intensive sciences have rapidly increased the volume and variety of questions, and subject specializations have begun to emerge across information professionals working in several sciences.[26] One desk with volunteers cannot address this growing need for information regarding DMPs, and static documents and trainings do not reach users where and when they have questions. An exhibit hall might not be the ideal location for offering data help, but it is a start, and lessons learned from this study about the types of data questions asked by scientists can inform future staffing to increase data sharing.

Acknowledgments

The authors would like to acknowledge support from the Earth Science Information Partners (ESIP), administered by the University of Tennessee, Knoxville; and those volunteers staffing the Data Help Desk, as well as the study participants and graduate teaching assistant Hannah Armendarez.

Bradley Wade Bishop is an associate professor in the School of Information Sciences, University of Tennessee, Knoxville; he may be reached by e-mail at: wade.bishop@utk.edu.

Rose M. Borden is a data engineer at Sandia National Laboratories in Albuquerque, New Mexico; she may be reached by e-mail at: @rborden16.

Notes

1. Sarah Higgins, "The Lifecycle of Data Management," chap. 2 in *Managing Research Data*, ed. Graham Pryor (London: Facet, 2012), 57–61; Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. "The FAIR [findability, accessibility, interoperability, reusability] Guiding Principles for Scientific Data Management and Stewardship," *Scientific Data* 3 (2016), https://www.nature.com/articles/sdata201618.
2. Nicholas Smale, Kathryn Unsworth, Gareth Denyer, and Daniel Barr, "The History, Advocacy and Efficacy of Data Management Plans," *bioRxiv* (2018), https://doi.org/10.1101/443499.
3. Bradley Wade Bishop and Carolyn F. Hank, "Curation, Digital," in *International Encyclopedia of Human Geography*, 2nd ed., ed. Audrey Kobayashi (Amsterdam, Neth.: Elsevier, 2020), *https://doi.org/10.1016/B978-0-08-102295-5.10531-1.*

4. Bradley Wade Bishop and Carolyn Hank, "Earth Science Data Management: Mapping Actual Tasks to Conceptual Actions in the Curation Lifecycle Model," in *Transforming Digital Worlds. iConference 2018. Lecture Notes in Computer Science*, vol. 10766, Gobinda Chowdhury, Julie McLeod, Val Gillet, and Peter Willett, eds. (Cham, Switz.: Springer, 2018), 598–608, *https://doi.org/10.1007/978-3-319-78105-1_67*.
5. Yurdagül Ünal, Gobinda Chowdhury, Serap Kurbanoğlu, Joumana Boustany, and Geoff Walton, "Research Data Management and Data Sharing Behaviour of University Researchers," *Information Research* 24, 1 (2019).
6. William K. Michener, James W. Brunt, John J. Helly, Thomas B. Kirchner, and Susan G. Stafford, "Nongeospatial Metadata for the Ecological Sciences," *Ecological Applications* 7, 1 (1997): 330–42, *https://doi.org/10.2307/2269427*.
7. Jake Carlson and Marianne Stowell-Bracke, "Data Management and Sharing from the Perspective of Graduate Students: An Examination of the Culture and Practice at the Water Quality Field Station," *portal: Libraries and the Academy* 13, 4 (2013): 343–61, https://*doi:10.1353/pla.2013.0034*; Sarika Sharma and Jian Qin, "Data Management: Graduate Student's Awareness of Practices and Policies," *asis&t* (*Proceedings of the Association for Information Science and Technology*) 51, 1 (2014): 1–3, *https://doi.org/10.1002/meet.2014.14505101130*; Christie A. Wiley and Erin E. Kerby, "Managing Research Data: Graduate Student and Postdoctoral Researcher Perspectives," *Issues in Science and Technology Librarianship* 89 (2018).
8. Cunera M. Buys and Pamela L. Shaw, "Data Management Practices across an Institution: Survey and Report," *Journal of Librarianship and Scholarly Communication* 3, 2 (2015), *https://doi.org/10.7710/2162-3309.1225;* Travis Weller and Amalia Monroe-Gulick, "Differences in the Data Practices, Challenges, and Future Needs of Graduate Students and Faculty Members," *Journal of eScience Librarianship* 4, 1 (2015) *https://doi.org/10.7191/jeslib.2015.1070*.
9. Anne R. Diekema, Andrew Wesolek, and Cheryl D. Walters, "The NSF/NIH [National Science Foundation/National Institutes of Health] Effect: Surveying the Effect of Data Management Requirements on Faculty, Sponsored Programs, and Institutional Repositories," *Journal of Academic Librarianship* 40, 3–4 (2014): 322–31, *https://doi.org/https://doi.org/10.1016/j.acalib.2014.04.010*; Mary Anne Kennan, "Data Management: Knowledge and Skills Required in Research, Scientific and Technical Organisations," in *Proceedings of IFLA WLIC* [International Federation of Library Associations and Institutions/World Library and Information Congress] *2016: Connections, Collaboration, Community*, *http://library.ifla.org/1466/1/221-kennan-en.pdf;* Merinda McLure, Allison V. Level, Catherine L. Cranston, Beth Oehlerts, and Mike Culbertson, "Data Curation: A Study of Researcher Practices and Needs," *portal: Libraries and the Academy* 14, 2 (2014): 139–64, *https://doi.org/10.1353/pla.2014.0009*; Jaana Pinnick, "Exploring Digital Preservation Requirements: A Case Study from the National Geoscience Data Centre (NGDC)," *Records Management Journal* 27, 2 (2017): 175–91; Christie Wiley and William H. Mischo, "Data Management Practices and Perspectives of Atmospheric Scientists and Engineering Faculty," *Issues in Science and Technology Librarianship* 85 (2016), *http://istl.org/16-fall/refereed3.html*.
10. Wiley and Kerby, "Managing Research Data."
11. Carlson and Stowell-Bracke, "Data Management and Sharing from the Perspective of Graduate Students."
12. Sharma and Qin, "Data Management."
13. Diekema, Wesolek, and Walters, "The NSF/NIH Effect"; Carolyn Bishoff and Lisa Johnston, "Approaches to Data Sharing: An Analysis of NSF Data Management Plans from a Large Research University," *Journal of Librarianship and Scholarly Communication* 3, 2 (2015): 1–27, *https://doi.org/10.7710/2162-3309.1231*; Martin Halbert, "The Problematic Future of Research Data Management: Challenges, Opportunities and Emerging Patterns Identified by the DataRes Project," *International Journal of Digital Curation* 8, 2 (2013): 111–22, *https://doi.org/10.2218/ijdc.v8i2.276*.
14. Bishoff and Johnston, "Approaches to Data Sharing."

15. Halbert, "The Problematic Future of Research Data Management."
16. Diekema, Wesolek, and Walters, "The NSF/NIH Effect."
17. Ian D. Gordon, Patricia Meindl, Michael White, and Kathy Szigeti, "Information Seeking Behaviors, Attitudes, and Choices of Academic Chemists," *Science & Technology Libraries* 37, 2 (2018): 130–51, *https://doi.org/10.1080/0194262X.2018.1445063*; Daryl L. Superio, Jessica B. Canaman, Janet P. Jaco, and Melda L. Estember, "The Information-Seeking Behavior of Aquaculture Researchers at the Southeast Asian Fisheries Development Center," *Information Development* 35, 3 (2019): 397–412, *https://doi.org/10.1177/0266666917754127;* Hemant Kumar Sahu and Surya Nath Singh, "Information Seeking Behaviour of Astronomy/Astrophysics Scientists," *Aslib* [Association for Information Management] *Proceedings* 65, 2 (2013): 109–42; Xi Niu and Bradley M. Hemminger, "A Study of Factors That Affect the Information-Seeking Behavior of Academic Scientists," *Journal of the American Society for Information Science and Technology* 63, 2 (2012): 336–53, *https://doi.org/10.1002/asi.21669*; Kumaripaba Athukorala, Eve Hoggan, Anu Lehtiö, Tuukka Ruotsalo, and Giulio Jacucci, "Information-Seeking Behaviors of Computer Scientists: Challenges for Electronic Literature Search Tools," *Proceedings of the American Society for Information Science and Technology* 50, 1 (2013): 1–11, *https://doi.org/10.1002/meet.14505001041;* Knut Alstad and Morten Hertzum, "Information Seeking by Geoscientists: An Update on [Julie] Bichteler and [Dederick] Ward," *Journal of Documentation* 74, 2 (2018): 447–60; Susan Wellings and Biddy Casselden, "An Exploration into the Information-Seeking Behaviours of Engineers and Scientists," *Journal of Librarianship and Information Science* 51, 3 (2019): 789–800, https://doi:*10.1177/0961000617742466;* Remigiusz Sapa, Monika Krakowska, and Małgorzata Janiak, "Information Seeking Behaviour of Mathematicians: Scientists and Students," *Information Research* 19, 4 (2014): 301–20.
18. Niu and Hemminger, "A Study of Factors That Affect the Information-Seeking Behavior of Academic Scientists."
19. Sahu and Singh, "Information Seeking Behaviour of Astronomy/Astrophysics Scientists."
20. Gordon, Meindl, White, and Szigeti, "Information Seeking Behaviors, Attitudes, and Choices of Academic Chemists."
21. Alstad and Hertzum, "Information Seeking by Geoscientists."
22. Jessica Adamick, Rebecca C. Reznik-Zellen, and Matt Sheridan, "Data Management Training for Graduate Students at a Large Research University," *Journal of eScience Librarianship* 1, 3 (2012): 130–88, https://doi:10.7191/jeslib.2012.1022; Jake Carlson, Lisa Johnston, Brian Westra, and Mason Nichols, "Developing an Approach for Data Management Education: A Report from the Data Information Literacy Project," *International Journal of Digital Curation* 8, 1 (2013): 204–17, *http://hdl.handle.net/1794/17990;* Bonnie L. Fong and Minglu Wang, "Required Data Management Training for Graduate Students in an Earth and Environmental Sciences Department," *Journal of eScience Librarianship* 4, 1 (2015): *https://doi.org 10.7191/jeslib.2015.1067;* Joseph H. Holles and Larry Schmidt, "Implementing a Graduate Class in Research Data Management for Science/Engineering Students" (presentation at 2018 ASEE [American Society for Engineering Education] Annual Conference & Exposition, Salt Lake City, UT, June 24–27, 2018); Lawrence Schmidt and Joseph Holles, "A Graduate Class in Research Data Management," *Chemical Engineering Education* 52, 1 (2018): 52–59; Joanna Thielen, Sara M. Samuel, Jake Carlson, and Mark Moldwin, "Developing and Teaching a Two-Credit Data Management Course for Graduate Students in Climate and Space Science," *Issues in Science and Technology Librarianship* 86 (2017), *http://dx.doi.org/10.5062/F42Z13HQ;* Maura Valentino and Michael Boock, "Data Management for Graduate Students: A Case Study at Oregon State University," *Practical Academic Librarianship* 5, 2 (2015): 77–91; Amanda L. Whitmire, "Implementing a Graduate-Level Data Information Literacy Curriculum at Oregon State University: Approach, Outcomes and Lessons Learned" (poster session, University of Massachusetts and New England Area Librarian e-Science Symposium, Worcester, MA, April 9, 2014), *https://doi:10.13028/351s-g605.*

23. Holles and Schmidt, "Implementing a Graduate Class in Research Data Management for Science/Engineering Students"; Schmidt and Holles, "A Graduate Class in Research Data Management."
24. Kerstin Helbig, "Research Data Management Training for Geographers: First Impressions," *ISPRS* [International Society for Photogrammetry and Remote Sensing] *International Journal of Geo-Information* 5, 4 (2016): 40; Lisa Johnston, Meghan Lafferty, and Beth Petsan, "Training Researchers on Data Management: A Scalable, Cross-Disciplinary Approach," *Journal of eScience Librarianship* 1, 2 (2012), https://doi:10.7191/jeslib.2012.1012; Robyn B. Reed, "Diving into Data: Planning a Research Data Management Event," *Journal of eScience Librarianship* 4, 1 (2015), https://doi:10.7191/jeslib.2015.1071.
25. Chung-Yi Hou, "Meeting the Needs of Data Management Training: The Federation of Earth Science Information Partners (ESIP) Data Management for Scientists Short Course," *Issues in Science and Technology Librarianship* 80 (2015), https://doi:10.5062/F42805MM; William Michener, Dave Vieglais, Todd Vision, John Kunze, Patricia Cruse, and Greg Janée, "DataONE: Data Observation Network for Earth—Preserving Data and Enabling Innovation in the Biological and Environmental Sciences," *D-Lib Magazine* 17, 1–2 (2011): 12, https://*doi:10.1045/january2011-michener;* Birgit Schmidt, Birgit Gemeinholzer, and Andrew Treloar, "Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey," *PLOS* [Public Library of Science] *One* 11, 1 (2016): 1–29, *https://doi.org/10.1371/journal.pone.0146695*; Shelley Stall, Kerstin Lehnert, Eric Robinson, Mark A. Parsons, Brooks Hanson, Joel Cutcher-Gershenfeld, and Brian Nosek, "Enabling FAIR and Open Data—the Importance of Communities on Influencing Change" (presentation at AGU [American Geophysical Union] Fall Meeting, New Orleans, LA, December 14, 2017); Carly Strasser and Patricia Cruse, "The DMPTool and DataUp: Helping Researchers Manage, Archive, and Share Their Data" (presentation at Research Data Management Implementations Workshop, Arlington, VA, March 13–14, 2013).
26. David A. Tyckoson, "Issues and Trends in the Management of Reference Services: A Historical Perspective," *Journal of Library Administration* 52, 6–7 (2012): 581–600, *https://doi.org/10.1080/01930826.2012.707953*.