The Effects of Research Data Management Services: Associating the Data Curation Lifecycle with Open Research Output

Nicolas Pares and Peter Organisciak

This study seeks to understand the relationship between research data management (RDM) services framed in the data curation life cycle and the production of open data. An electronic questionnaire was distributed to US researchers and RDM specialists, and the results were analyzed using Chi-Square tests for association. The data curation life cycle does associate with the production of open data and shareable research, but tasks like data management plans have stronger associations with the production of open data. The findings analyze the intersection of these concepts and provide insight into RDM services that facilitate the production of open data and shareable research.

Introduction

The increase of digital services and digital content in academia and libraries, due in part to the COVID-19 pandemic and continued technological change, has shifted researchers and RDM providers toward more networked and open digital paradigms. This study evaluates the relationship between research data management (RDM) services and the production of open data and shareable research outputs, from the perspective of institutional RDM services and researchers in the United States.

A survey was administered to collect data on the relationship between RDM practices, resources, and services framed in the data curation life cycle and the researcher's abilities and beliefs in the production of open data and shareable research outputs. This study focused on the following research questions:

1. Do institutional contexts such as location of RDM services, availability of RDM services, and institutional roles affect researchers' ability and sense of importance to produce shareable and open primary research data?
2. Do institutional services framed in the stages of the Data Life Cycle affect researcher's ability and sense of importance to produce shareable and open primary research data?

In the remainder of this paper, we describe the survey design, analysis, and outcomes, our results, and the role of data management plans in supporting open data production. We also note the importance of RDM education for researcher buy-in.

Nicolas Pares, University of Denver, email: nicolas.pares@du.edu; Peter Organisciak, University of Denver, email: Peter.Organisciak@du.edu.

Background

Open and shareable data has incredible value for scholarly communication and scientific growth.[1] When data and research can be reused for secondary research, findings are reproducible and more easily validated, major research projects spanning many years have continuity, or when researchers avoid unnecessary duplication scholarly communication advances more efficiently.[2] In addition, data sharing can lead to more collaboration, which makes research more beneficial to scholars.[3] This project studies the extent to which the availability of research data services facilitates the production of open and shareable data.

The need for open and shareable research data aligns with the public-service mission of most U.S. higher education institutions. This need is punctuated by the Open Data policies of the US government.[4] The Committee on Science, Engineering, and Public Policy recommends that developing policies, standards, and infrastructure needed to ensure the integrity, accessibility, and stewardship of research data is a critically important task.[5] Grechkin et al. explained that open data is a vital pillar of open science and a key enabler for reproducibility, data reuse, and novel discoveries.[6]

The State of the Scholarly Record

Scholarly communication is growing to incorporate more types of research and its outputs. Today, a wide range of research outputs, produced during initial collection as well as after formal publication, is being stewarded as part of the permanent documentation of scholarship, including research data, pre-prints, computer code, and more.[7] The management of a variety of research outputs and data becomes even more complex when specific approaches to their curation, management, metadata descriptions, and preservation are needed.[8] One would think that institutions would be acting to implement RDM services based on this growing problem, but a 2013 survey found that 82 percent of the respondent's home institutions had not implemented any institutional policy or services to address institutional RDM needs.[9]

This disconnect between institutional services and meeting the RDM needs of today becomes even more apparent as journals and funding sources increasingly require data sharing. Projects are being undertaken to identify research data associated with publications that should be openly available. For example, the Wide-Open project is a text mining system for detecting datasets that were referenced in published papers but are kept private. After parsing over 1.5 million open access publications, Wide-Open has identified hundreds of datasets overdue for publication; 400 of them were then released within one week.[10] With efforts like the Wide-Open project and increased use of digital materials, sharing must become a practice of academia and the institutional services supporting that practice.

Given the value proposition of open data, the argument for data sharing is even stronger when we consider that by increasing access to research data, the amount of scholarly communication and scientific growth increases.[11] There have been significant shifts by many stakeholders like journals and funders, but studies show that most researchers have not received any training in RDM such as data management planning, metadata, or file naming. However, most researchers would welcome formal training in different aspects of RDM.[12]

Research Data Management Services Today

RDM models have become common in many academic and research institutions but are highly diverse in scope and range.[13] The OCLC systematically analyzed and categorized the

primary methods of service delivery, detailing three general categories of service: education, expertise, and curation (see Figure 1).

FIGURE 1
"RDM Service Categories"*

Credit: The Realities of Research Data Management,[14] CC BY 4.0.
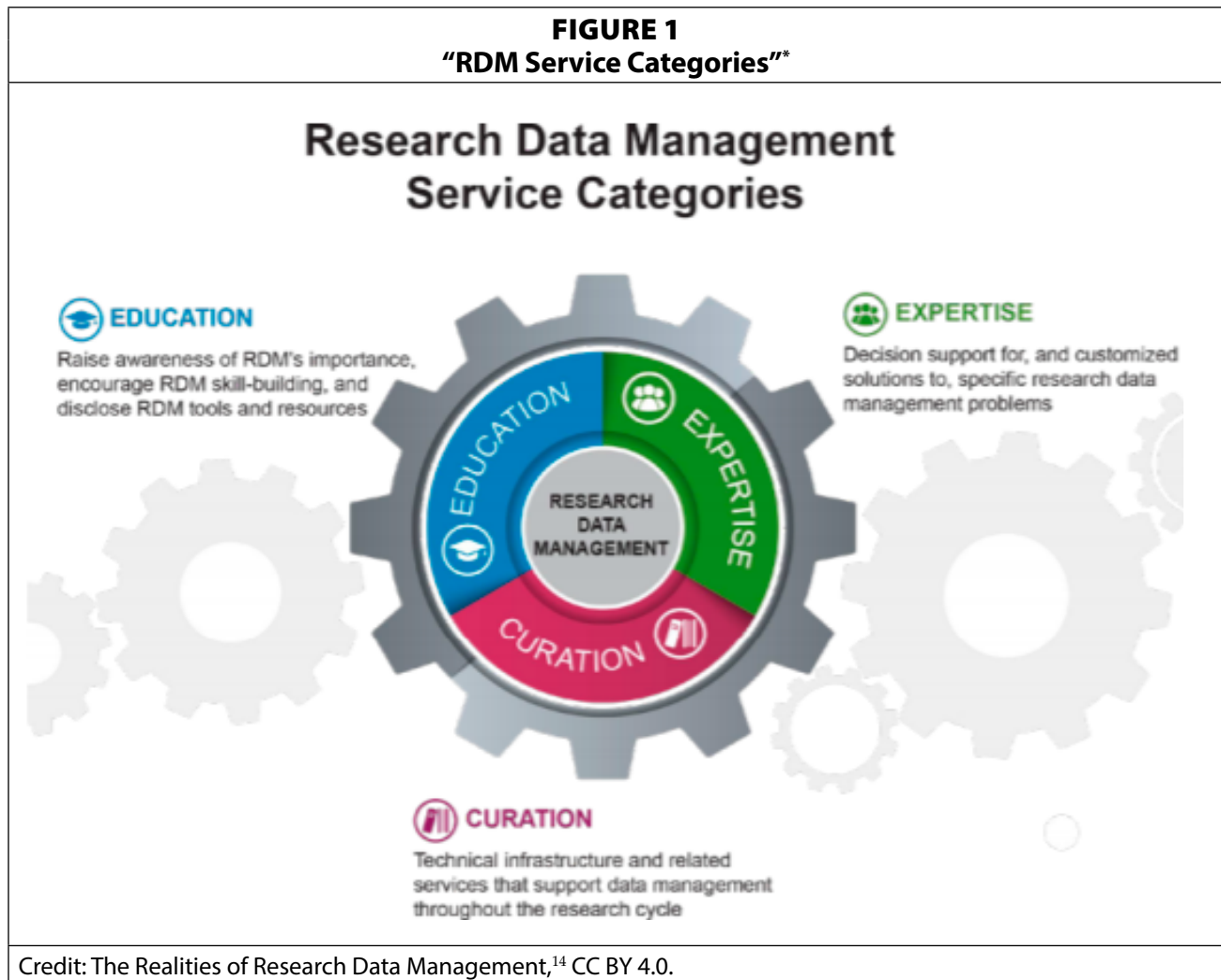
These three different methods for delivery of service do not account for the actual content or specific support service that is provided but do categorize the models analyzed by the OCLC.[15] Although this visual above simplifies the RDM space, RDM crosscuts many departments, functions, and sectors of academia, government, and industry,[16] and can look significantly different between disciplines.[17]

Data Lifecycle Model as Conceptual Model for RDM Services

Although RDM services have common models for delivery, the types of services differ. One conceptual model for framing RDM services is the data life cycle. According to Charalabadis et al., there are several data life cycle models being used globally, but the most comprehensive model is the Data Curation Centre (DCC) Data Curation Lifecycle Model,[18] since it includes administrative and managerial processes. DCC describes its Curation Lifecycle Model as a tool to help you "define research data management (RDM) workflows and associated roles and responsibilities within your organization," providing a "holistic approach to RDM infrastructure development and optimization [that] can be used to help organizations map research data

management activities and support across functional and operational units."[19] This generic but comprehensive design makes it ideal for developing RDM services at an institution that might not know its initial institutional needs. Molloy & Snow support this, noting that the majority of core RDM skills were generic across disciplines at the postgraduate level.[20]

This study adopts the DCC Data Curation Lifecycle Model.[21] as the basis for its survey of RDM services and the production of open and shareable research data. In summary, this choice was motivated by the fact that the model is

- applicable to a breadth of disciplines and domains of scholarly communication,
- holistic in considering both administrative and management design, and
- easily mappable to researcher tasks.

So how do RDM services based on the stages of the data curation lifecycle affect researchers' ability and their sense of importance for producing shareable and open primary research data?

Approach

To explore how RDM services affect researcher affordances for producing open data, a study of institutional services and researcher production of research was needed. Therefore, the current study targeted researchers who have used or have knowledge of their institution's RDM services and the RDM specialists providing support. The sample would need to be distributed among as many institutions as possible to get a sample size that represents as many different RDM service models as possible.

There have been many large-scale research studies conducted on data sharing and preservation in higher education. These studies have looked at the delivery methods for RDM, journal requirements and their instructions on data sharing, and other studies measuring the awareness and capability of researchers. This approach complements the current research and looks more specifically at a gap in RDM service designs as it relates to supporting open data and the production of shareable research outputs.

Situating This Study in the Current Research

When looking across the current research, there are many studies and articles highlighting different stakeholder needs and perspectives inside and outside of the US. Tenopir et al. provide insights into the institutional culture surrounding the research, process being the biggest barrier to sharing and preserving data.[22] Vasilevsky et al. found that 65.7 percent of the journals that they analyzed required reproducible data sharing but did not provide guidance on how to make research and data available and reusable.[23] Within this space of journals requiring data sharing, the Wide-Open project has pushed journals to honor their data sharing requirements, and of 473 datasets identified by Wide-Open by February 2017, 455 have been released.[24] Additionally, the Ünal et al. study demonstrated a clear gap in awareness and understanding of managing and sharing research data.[25] Together, these studies begin to tell a story with a missing stakeholder, the RDM services at institutions.

When looking at the research and literature into RDM service development in the United States, librarians have developed curriculums[26] and thoroughly identified the challenges, training, and research data management roles they now fulfil.[27] However, librarians as RDM service providers are only one model of delivery defined by the OCLC's "The Realities of Research Data Management: A Tour of the Research Data Management (RDM) Service Space."[28]

The OCLC thoroughly describes three types:
- Education—educating researchers and other stakeholders on the importance, and in some cases the necessity, of responsibly managing their data and making arrangements for its long-term curation
- Expertise—providing decision support and customized solutions for researchers working through specific RDM problems
- Curation—supplying technical infrastructure and related services that support data management throughout the research cycle

These three types of RDM service models are comprehensive but lack specific service types, i.e., data preservation or data analysis.[29] RDM services cover a lot of territory, and it is difficult to draw firm boundaries around this service space. Bryant et al. explain that the specific services offered within categories varied from institution to institution.[30] Although this is an excellent analysis of service models and how the service is delivered, it does not consider what stage in the data life cycle the service supports.

Problem Statement

Having evaluated the current literature and identified a study methodology that could provide an exploratory view into this gap in the research, this study aimed to answer the following problem statement: Do institutional RDM services framed in the stages of the data curation lifecycle affect the researcher's ability and sense of importance for producing open and shareable research data?

Research Questions

The following research questions will be tested to develop a better understanding of the variables that might be related to the problem statement above.
1. Do institutional aspects like location of RDM services, availability of RDM services, and institutional roles affect researcher's ability and sense of importance to produce shareable and open primary research data?
2. Do institutional services framed in the stages of the Data Life Cycle affect researcher's ability and sense of importance to produce shareable and open primary research data?

Methods

To address the research questions, a survey method was used. Fourteen survey items were crafted to reflect the stages of the DCC Data Curation Lifecycle Model[31] stages, and the respondent was asked to consider services and resources at each stage. Additionally, the survey collected background information and measures of confidence and beliefs in relationship to open data and sharing research. These items and scales were then analyzed using the chi-square test of association. The results of those statistical tests were then combined to provide a more holistic view of each research question.

Survey Sampling

This study utilized non-probabilistic, purposive sampling for two primary reasons. First, the study and survey were intended for a target population within higher education that has experience with RDM and/or academic publishing. This population can be difficult to reach geographically and in a timely fashion. The second reason for the purposive sampling was to

collect responses from as many different institutions as possible, which required individual solicitation of survey participation.

The survey was distributed electronically using Qualtrics, an electronic survey tool, to a series of relevant research-minded library and information science email listservs with a strong US representation. These included the American Library Association's Scholarly Communication community, Research Data Access and Preservation, the California Association for Institutional Research, the Research Assessment and Metrics Interest Group from ACRL, the Open Data Research Interest Group of ALA, the OCLC Research Support community, and the Colorado Academic Library Association.

Survey Items

The survey included fourteen items or questions that gathered specific measures in relationship to the research questions. The items are nominal or categorical variables, or they are ordinal variables i.e., Likert scales capturing belief or traits. The survey items are listed below and include their short name in parenthesis, which is used throughout the remainder of the study. This survey used the following nominal measures:

- What is your experience with academic research at your institution? (Role)
- Please provide the name of your institution. (Institution)
- Where are research data management services located on your campus? Select all that apply. (Location)
- Do you have access to research data services like training, consultations, or tools that support the planning of research data collection, preservation, and analysis? (Planning)
- Do you have access to research data services like training, consultations, or tools that support the collection of research data? (Collection)
- Do you have access to research data services like training, consultations, or tools that support the description of data, like documentation that describes the data, using metadata standards? (Describe)
- Do you have access to research data services like training, consultations, or tools that support data analysis efforts like NVivo, Stata, SPSS, etc.? (Analysis)
- Do you have access to research data services that support the organization, cleaning, and management of research data? (Management)
- Do you have access to research data services, training, or tools that support the development of data management plans? (DMP)
- Are there training, courses, or certifications available at your institution that support research data management practices? (Education)

    The following ordinal scale items were used to measure traits and beliefs:

- Do the research data management services and tools available to you support the creation and maintenance of data that would be shareable and openly publishable? (Support)
- How important is sharing research data? (Important)
- Can you produce/support research data that is shareable or could be made publicly available? (Ability)
- Is the success of scientific or academic publications dependent upon the availability of research data management services and tools? (Success)

Data Analysis

The study utilized nominal and ordinal variables that limited the selection of statistical analysis to the chi-square test for association. The chi-square test for association measures the relationship between nominal variables and ordinal variables; specifically, do the values of one variable depend on the other nominal or ordinal variable? This type of significance either supported or rejected the null research questions and gave some degree of insight into the research questions. The research questions are presented in the findings section and are written in the form of the null research question or research question that was supported by the chi square test.

Findings

The study's findings address the research questions and the relevant measures of association between survey items. Chi-square tests for association were used to accept or reject each null research question. The data met all assumptions of the chi-square test for association. Additionally, some descriptive statistics about the sample are presented. The survey items are referenced by their shorthand title, which is listed in the survey item section above.

Sample Description

The study sample consisted of N=46. The survey completion rate was 46/108 = 42 percent. The overall response rate is not known as the total eligible population reached via email listserv is uncertain. The final sample had representation from private, public, community college, four-year, and graduate American higher education institutions and represents a diverse geographic sample (see figure 2).

The survey sample had almost equal responses of researchers and RDM support personnel (see figure 3). The survey provided the option to select both role types, but no respondents identified as both.

FIGURE 2
Geographic Location of Participant Institutions
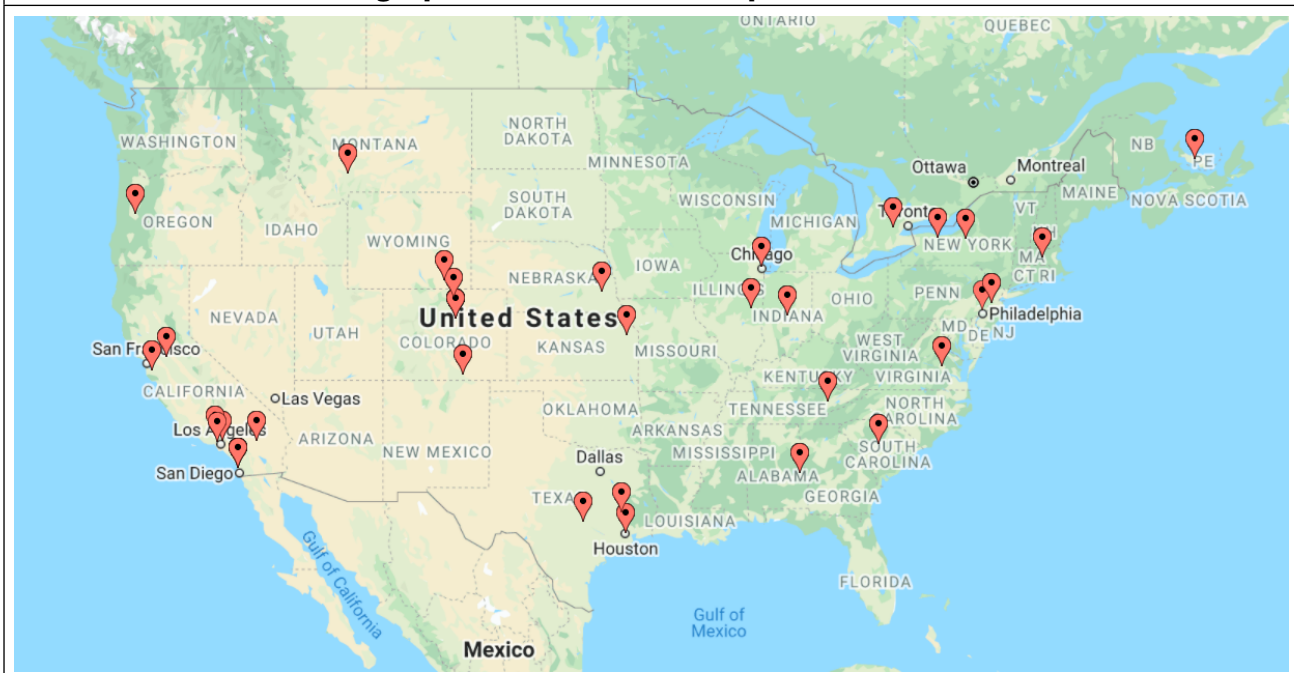
Research Question 1

Research question 1 explored institutional aspects like location of RDM services, availability of RDM services, and institutional roles that affect researcher's ability and sense of importance to produce shareable and open primary research data.

There was an association between location and importance.

A chi-square test of association was performed to examine the relationship between location of

FIGURE 3
Survey Participant Roles Count

- Research Data Management Support
- Academic Researcher

19, 43%

25, 57%

RDM and a researcher's belief in the importance of open data and sharing research. The relationship between these variables was significant, X^2 (4, N = 44) = 11.973, p = .018. When services are located at the library, they are associated with more importance.

There was an association between location and ability.

A chi-square test of association was performed to examine the relationship between location of RDM and a researcher's ability to produce open data and share research. The relationship between these variables was significant, X^2 (3, N = 44) = 9.469, p = .024. The association of services at the library had a relationship with their ability.

There was no association between role and ability.

A chi-square test of association was performed to examine the relationship between a researcher's role and a researcher's ability to produce open data and share research. The relationship between these variables was not significant, X^2 (6, N = 44) = 6.387, p = .381.

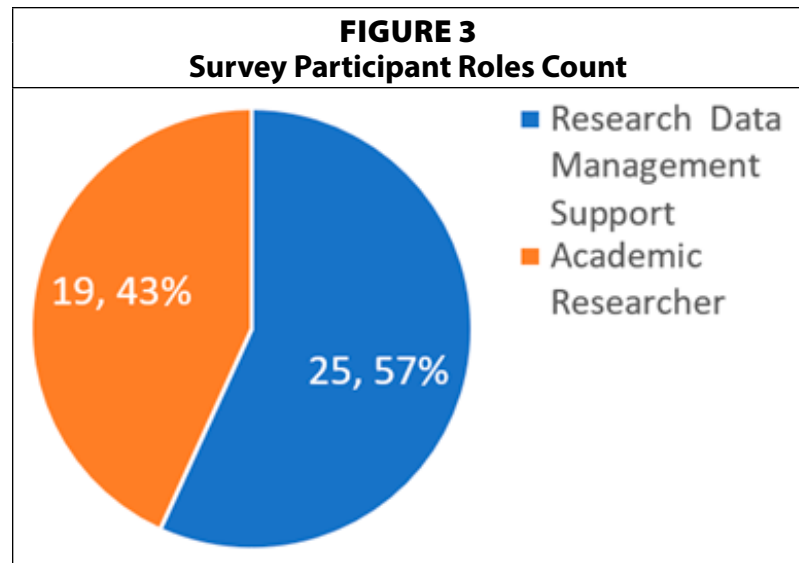### There was no association between role and importance.

A chi-square test of association was performed to examine the relationship between a researcher's role and a researcher's belief in the importance of open data and sharing research. The relationship between these variables was not significant, X^2 (4, N = 44) = 5.351, p = .361.

There was an association between availability and importance.

A chi-square test of association was performed to examine the relationship between location of RDM and a researcher's belief in the importance of open data and sharing research. The relationship between these variables was significant, $X2$ (4, N = 44) = 11.973, p = .018. When services are made available, there is an association with importance.

There was an association between education and importance.

A chi-square test of association was performed to examine the relationship between the availability of RDM education and a belief in the importance of open data and sharing research.

The relationship between these variables was significant, X^2 (2, N = 44) = 6.515, p = .038. The presence of educational services had an association with a researcher's ability.

Research Question 2

Research question 2 explored institutional services framed in the stages of the Data Life Cycle, and if they affect researcher's ability and sense of importance to produce shareable and open primary research data.

There was no association between services across the data curation life cycle and ability.

A series of chi-square tests for association were performed to examine the relationship between RDM services across the data curation life cycle and a researcher's ability to produce open data and shareable research. The relationship between these variables was not significant at any intercept.

There was an association between support and services across the life cycle.

There was a relationship between access to services at each stage and those services supporting the production of open and publishable primary research data. The relationship between these variables was significant, with a minimal significance of X^2 (2, N = 44) = 6.093, p = .048 found at the data collection stage.

There was an association between data management plans and support of open data and shareable research.

A chi-square test of association was performed to examine the relationship between data management plans and a researcher's ability to produce open data and shareable research. The relationship between these variables was significant, X2 (1, N = 44) = 7.801, p = .005.

Discussion

The survey data provided insight into the impact on the production of open and shareable research data from two contexts: the availability and home of RDM services in an institution, and the framing of services through the Curation Data Lifecycle Model. It also provided direction for future areas of study.

The Availability of RDM Services as an Aspect of Awareness

The study showed a significant association between the availability of RDM services and the importance of open data and shared research. It also showed a significant association between the availability of RDM services and researcher ability to produce open data and shareable research. These associations mean that when an institution has invested in and created RDM services, faculty have more confidence and belief of the importance for sharing data and research.

The significance of this finding is that RDM services not only help support good data practices at an institution, but their very availability serves an advocacy role, raising awareness of open data and research-sharing policies and their benefits. For information professionals planning RDM services, it is worth considering that they respond to existing institutional needs for open and shareable data.

Current RDM Services and Open, Shareable Research

When it came to finding significant associations between RDM services modeled after the Data Curation Lifecycle and open data or shareable research production, the survey produced mixed results. The survey did not yield significant associations between a researcher's ability to produce open research and the stages of the data curation life cycle. This needs to be analyzed further as there are likely further factors motivating a researcher's ability to produce open data and shareable research beyond the availability of services at each stage. The possible variables could be incentives, researcher motivation, technology need, etc.

There were significant associations between services provided at the library at each stage of the data curation lifecycle and support that leads to open and shareable research data. When RDM services are located at the library, they have a stronger association with producing open data and shareable research. It's not apparent why this is so, though it may be a factor associated with the academic library's traditional role as a service provider between units, which positions them well for advocacy or, at a minimum, awareness of their services. As Heidorn argues, RDM activities align well with the infrastructure and traditional skills of libraries.[32]

Indeed, the survey found that a majority of the RDM services at institutions represented in the sample were found or provided through the library in the US. One of the most significant trends is for libraries to work in conjunction with other units in their institutions, for example information technology units and research offices, to support RDM.[33] Since 80 percent of respondents identified RDM services at the library, the libraries should be the primary location for these types of services.

In practice, centralizing RDM services at libraries, even when offering those services through partnerships between units, helps promote open and shareable research at the institution.

Data Management Plans as a Promising Practice

RDM services supporting the development of data management plans had a strong association with a researcher's ability to produce open data and shareable research. In addition, data management plans might bring insights into the needs of researchers and institutional service design.[34] Williams et al.'s research into data management plans did find increased data sharing, but not necessarily research that was reproducible.[35] Further research into what could be considered a comprehensive data management plan should be undertaken as federal mandates and many grant funding opportunities begin to require data management plans.[36]

Limitations

The first limitation of the study is the representativeness of the greater American higher education landscape. The survey could have collected more background information and been circulated to a more diverse population. For example, no two-year institutions were included in the sample. The non-probabilistic, purposive sampling method would need to be extended further and more participants would need to be identified and contacted to obtain a more diverse sample.

A second limitation is a potential for confirmation bias in some of the survey items. Certain nominal survey items did not include an "unsure" option when describing available RDM services at their respective institutions. While the survey was distributed to a very specific population with knowledge of RDM services, this presumed a level of familiarity with the full

institution's services that may not be entirely representative of the roles. Further, researchers with multiple RDM roles were either not represented or did not identify as such in the sample.

Another limitation of the study was that it did not look at motivational aspects for sharing research and producing open data. Motivation—whether it be promotion, incentives, or grant requirements—could provide additional clarity on researcher's belief of importance in sharing research and producing open data.

Future Directions

The first recommendation for continued research would be to combine this data curation lifecycle-framed model of RDM services with a needs analysis. Assessment data and data management plans should significantly influence the selection or addition of RDM services and policies that guide their implementation.[37] Starting with a holistic data curation lifecycle-framed approach to RDM services and then iterating and refining to meet an institution's need through tracking of service use could lead to an effective RDM service model.

The second recommendation for continued research would be to modify this survey for a specific academic discipline. The data could be used to identify missing or wasted RDM services. There is some research in certain disciplines like the social sciences in the United Kingdom[38] or social and economic sciences in Germany[39] where growing awareness of data sharing and publishing of reproducible research is occurring. However, further research into available RDM services and their ability to support discipline-specific researchers in production of open data and shareable research is still needed.

The third recommendation for continued research is the effects of funding requirements or incentives on a researcher's belief in the importance and ability to produce open data and shareable research. Questions of researcher motivation are not addressed in this study but could affect the perceived importance of open data and sharing research.

The final recommendation for further research is the implications of the 2020 COVID-19 pandemic and institutional shifts to digital academic resources. This survey was administered during the pandemic, which may skew its data, but the potential impact is currently unknown. With an increase of digital content use, it remains to be seen if researchers are becoming more familiar with open content.[40] This new digital demand might have even changed budgetary concerns[41] and brought a renewed vigor for open data and shareable research. The impacts of this digital shift on the needs of RDM services should be researched further.

Conclusion

This study into research data management (RDM) services and the production of open data and shareable research was intended to fill a gap in the current RDM literature by addressing the problem statement: How do institutional RDM services framed in the stages of the data curation lifecycle affect researcher's ability and sense of importance for producing shareable research and open data? To address this problem, a set of research questions were tested using a survey methodology and statistical analysis.

The research question that addressed the data life cycle and how it might affect researcher's ability and sense of importance to produce shareable and open primary research data was partially rejected when there was no significant relationship between researcher's ability and access to data curation lifecycle services. However, there was a relationship between access to services at each stage of the life cycle and the belief that those services would support the

production of open and publishable primary research data. The rejection of one research question and the acceptance of another was an indicator that more work needs to be done on applying the data curation life cycle.

The research question that addressed institutional aspects like location and availability of RDM services and institutional roles and how much they affect researcher's ability and sense of importance was partially supported. The association that was significant was the location of services that did have significant relationship with ability and importance. The second association between availability of services and researcher's ability was found to have significant associations. The third association between availability of services and researcher's sense of importance for open data and shareable research was found to have significant associations. However, role association with ability or importance had no significant associations. While not all associations were accepted, the first three were found to be significant and will need additional research.

When we return to the problem statement, there is no clear answer to the use of the data curation lifecycle as a frame for RDM service models in the US. However, there were several interesting findings, such as the promising potential for data management plans and libraries as key locations for housing these RDM services. These findings have only led to more questions to explore at the intersection of open data and RDM service design.

Appendix A. Survey Questionnaire

Open Primary Research Data Survey

This survey references research data management, which is the care and maintenance of data during a research cycle. Funding agencies are increasingly requiring data management plans and research data management practices that would support dissemination of research to collaborators, evaluators, or other parties.

For this survey you will be asked to consider your institution's, college's, or university's research data management services. These services can often be provided through the library, institutional research, or an office of data analytics. These services might also be hosted by individual departments. When answering these questions, use that information to complete the following questions.

What is your experience with academic research at your institution?
- ☐ I support or provide research data management services.
- ☐ I produce academic research.
- ☐ I have experience using research data management services and tools but am not published.

Please provide the name of your institution.

_____

Where are research data management services located at your institution?
- ☐ The library
- ☐ A data services office
- ☐ other

The following questions will ask about available research data services like consultations, training, or tools that train and support researchers to generate primary research data. For additional information please visit Data Curation Life Cycle description and consider all research supports at your institution when answering.

Do you have access to research data services like training, consultations, or tools that support the planning of research data collection, preservation, and analysis?
- ☐ Yes
- ☐ Yes, but quantitative and statistical data only
- ☐ Yes, but for qualitative data only
- ☐ No

Do you have access to research data services like training, consultations, or tools that support the collection of research data (i.e., survey tools, recording software, or tools where observations are made either by hand or with sensors or other instruments and the data are placed an into digital form)
- ☐ Yes
- ☐ Yes, but quantitative and statistical data only

    □   Yes, but for qualitative data only
    □   No

Do you have access to research data services like training, consultations, or tools that support the description of data, like documentation that describes the data, using metadata standards?
    □   Yes
    □   Yes, but quantitative and statistical data only
    □   Yes, but for qualitative data only
    □   No

Do you have access to research data services like training, consultations, or tools that support data analysis efforts like NVivo, Stata, SPSS, etc.?
    □   Yes
    □   Yes, but quantitative and statistical data only
    □   Yes, but for qualitative data only
    □   No

Do you have access to research data services that support the organization, cleaning, and management of research data?
    □   Yes
    □   Yes, but quantitative and statistical data only
    □   Yes, but for qualitative data only
    □   No

Do you have access to research data services, training, or tools that support the development of data management plans?
    □   Yes
    □   Yes, but quantitative and statistical data only
    □   Yes, but for qualitative data only
    □   No

Are there training, courses, or certifications available at your institution that support research data management practices?
    □   Yes
    □   No

Does the research data management services and tools available to you support the creation and maintenance of data that would be shareable and openly publishable?
    □   Yes
    □   No

You are almost done! This final set of questions will ask you about your confidence in, the importance of, and ability to create or use shareable/open primary research data.
How important is sharing research data?
    □   Extremely important
    □   Very important
    □   Moderately important

    ☐ Slightly important
    ☐ Not at all important

Can you produce/support research data that is shareable or could be made publicly available?
    ☐ Strongly agree
    ☐ Agree
    ☐ Somewhat agree
    ☐ Neither agree nor disagree
    ☐ Somewhat disagree
    ☐ Disagree
    ☐ Strongly disagree

Is the success of scientific or academic publications dependent upon the availability of research data management services and tools?
    ☐ Strongly agree
    ☐ Agree
    ☐ Somewhat agree
    ☐ Neither agree nor disagree
    ☐ Somewhat disagree
    ☐ Disagree
    ☐ Strongly disagree

Notes

1. Institute of Medicine, National Academy of Engineering, National Academy of Sciences, Engineering Committee on Science, and Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, "Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age," Washington, D.C: National Academies Press, 2009.

2. Ibid.

3. Heather A. Piwowar and Todd J. Vision, 2013, "Data Reuse and the Open Data Citation Advantage," *PeerJ* 1 (October): e175. https://doi.org/10.7717/peerj.175; P. Bryan Heidorn, 2011, "The Emerging Role of Libraries in Data Curation and E-Science," *Journal of Library Administration* 51 (7–8): 662–72, https://doi.org/10.1080/01930826.2011.601269.

3. Maxim Grechkin, Hoifung Poon, and Bill Howe, "Wide-Open: Accelerating Public Data Release by Automating Detection of Overdue Datasets," *PLoS Biology* 15, no. 6 (2017), https://doi.org/10.1371/journal.pbio.2002477.

4. Veerle Van den Eynden and Louise Corti, "Advancing Research Data Publishing Practices for the Social Sciences: From Archive Activity to Empowering Researchers," *International Journal on Digital Libraries* 18, no. 2 (2017): 113–21.

5. Piwowar and Vision, "Data Reuse and the Open Data Citation Advantage."

6. Grechkin, Poon, and Howe, "Wide-Open: Accelerating Public Data Release by Automating Detection of Overdue Datasets."

7. Brian Lavoie, Eric Childress, Ricky Erway, Ixchel Faniel, Constance Malpas, Jennifer Schaffner, and Titia van der Werf (2014), *The Evolving Scholarly Record*, (Dublin, Ohio: OCLC Research), https://doi.org/10.25333/C3763V.

8. Joachim Schöpfel, Stéphane Chaudiron, Bernard Jacquemin, Hélène Prost, Marta Severo, and Florence Thiault, "Open Access to Research Data in Electronic Theses and Dissertations: An Overview," *Library Hi Tech* 32, no. 4 (2014): 612–27.

9. Martin Donnelly, "Review: Research Data Management: Practical Strategies for Information Professionals," *International Journal of Digital Curation* 9, no. 2 (2014): 1–5.

10. Grechkin, Poon, and Howe, "Wide-Open: Accelerating Public Data Release by Automating Detection of Overdue Datasets."

11. Benedikt Fecher, Sascha Friesike, and Marcel Hebing, "What Drives Academic Data Sharing?" *PloS One* 10, no. 2 (2015): e0118053, https://doi.org/10.1371/journal.pone.0118053.

12. Kevin B. Read, Catherine Larson, Colleen Gillespie, So Young Oh, and Alisa Surkis, "A Two-Tiered Curriculum to Improve Data Management Practices for Researchers," *PloS One* 14, no. 5 (2019): e0215509, https://doi.org/10.1371/journal.pone.0215509.

13. Rebecca Bryant, Brian Lavoie, and Constance Malpas, (2020, May 4), "The Realities of Research Data Management," n.p., n.d.

14. Rebecca Bryant, Brian Lavoie, and Constance Malpas, 2017, "Scoping the University RDM Service Bundle," *The Realities of Research Data Management Part 2* (Dublin, OH: OCLC Research), https://doi.org/10.25333/C3Z039.

15. Ibid.

16. Martin Donnelly, "Review: Research Data Management: Practical Strategies for Information Professionals." *International Journal of Digital Curation* 9, no. 2 (2014): 1–5.

17. K. G. Akers and J. Doty, "Disciplinary Differences in Faculty Research Data Management Practices and Perspectives," *International Journal of Digital Curation* 8 (2013): 26.

18. Yannis Charalabidis, Anneke Zuiderwijk, Charalampos Alexopoulos, Marijn Janssen, Thomas Lampoltshammer, and Enrico Ferro, "The Multiple Life Cycles of Open Data Creation and Use," *The World of Open Data*, 11–31 (Cham: Springer International Publishing), 2018.

19. Sarah Higgins, "The DCC Curation Lifecycle Model," *International Journal of Digital Curation* 3, no. 1 (2008): 134–40.

20. Laura Molloy and Kellie Snow. "The Data Management Skills Support Initiative: Synthesising Postgraduate Training in Research Data Management," *International Journal of Digital Curation* 7, no. 2 (2012): 101–9.

21. Higgins, "The DCC Curation Lifecycle Model."

22. Carol Tenopir, Suzie Allard, Kimberly Douglass, Arsev Umur Aydinoglu, Lei Wu, Eleanor Read, Maribeth Manoff, and Mike Frame, "Data Sharing by Scientists: Practices and Perceptions," *PloS One* 6, no. 6 (2011): e21101, https://doi.org/10.1371/journal.pone.0021101.

23. Nicole A. Vasilevsky, Jessica Minnier, Melissa A Haendel, and Robin E Champieux, "Reproducible and Reusable Research: Are Journal Data Sharing Policies Meeting the Mark?" *PeerJ* (San Francisco, CA) 5 (2017): e3208, https://doi.org/10.7717/peerj.3208.

24. Grechkin, Poon, and Howe, "Wide-Open: Accelerating Public Data Release by Automating Detection of Overdue Datasets."

25. Yurdagül Ünal, Gobinda Chowdhury, Serap Kurbanoğlu, Joumana Boustany, and Geoff Walton, (2019), "Research Data Management and Data Sharing Behaviour of University Researchers," *Information Research*, 24.

26. Kevin B. Read, et al., "A Two-Tiered Curriculum to Improve Data Management Practices for Researchers."

27. Rong Tang and Zhan H,. "Providing Research Data Management (RDM) Services in Libraries: Preparedness, Roles, Challenges, and Training for RDM Practice," *Data and Information Management* 3, no. 2 (2019): 84–101.

28. Bryant, Lavoie and Malpas, "The Realities of Research Data Management."

29. Ibid.

30. Ibid.

31. Higgins, "The DCC Curation Lifecycle Model."

32. Heidorn, "The Emerging Role of Libraries in Data Curation and E-Science."

33. Mary Williams, Jacqueline Bagwell, and Meredith Nahm Zozus, "Data Management Plans: The Missing Perspective," *Journal of Biomedical Informatics* 71 (2017): 130–42.

34. Susan Wells Parham, Jake Carlson, Patricia Hswe, Brian Westra, and Amanda Whitmire, "Using Data Management Plans to Explore Variability in Research Data Management Practices across Domains," *International Journal of Digital Curation* 11, no. 1 (2016): 53–67.

35. Williams, Bagwell, and Nahm Zozus, "Data Management Plans: The Missing Perspective."

36. A. R. Diekema, A. Wesolek, and C D. Walters, "The NSF/NIH Effect: Surveying the Effect of Data Management Requirements on Faculty, Sponsored Programs, and Institutional Repositories," *The Journal of Academic Librarianship* 40 (2014): 331.

37. Tyler O. Walters. "Data Curation Program Development in U.S. Universities: The Georgia Institute of Technology Example," *International Journal of Digital Curation* 4, no. 3 (2009): 83–92.

38. Van den Eynden and Corti, "Advancing Research Data Publishing Practices for the Social Sciences."

39. Monika Linne and Wolfgang Zenk-Möltgen, "Strengthening Institutional Data Management and Promoting Data Sharing in the Social and Economic Sciences," *LIBER Quarterly* 27, no. 1 (2017).

40. Christopher Cox, "Academic Libraries Will Change in Significant Ways as a Result of the Pandemic," *Inside Higher Ed,* June 4, 2020.

41. Geoffrey Boulton, Michael Rawlins, Patrick Vallance, and Mark Walport, "Science as a Public Enterprise: The Case for Open Data," *The Lancet* (British edition) 377, no. 9778 (2011): 1633–35.