Unfolding Research Data Services: An Information Architecture Perspective*

Hammad Rauf Khan
University of North Texas
Denton, TX
United States
hammadkhan@my.unt.edu

Hsia-Ching Chang
University of North Texas
Denton, TX
United States
hsia-ching.chang@unt.edu

Jeonghyun Kim
University of North Texas
Denton, TX
United States
jeonghyun.kim@unt.edu

ABSTRACT

With the exponential growth in digital research data, libraries are beginning to find opportunities to assist researchers with planning, maintaining, sharing, and accessing data through research data services. Using a content analysis with the lens of information architecture, this study sought to better understand how these services are organized in North American academic library websites and to what extent the research data lifecycle is supported within research data services. 50 academic library websites were studied and results yielded three provisions that make up research data services: Information Access, Technical Support, and Personalized Consultation. The data lifecycle was found to be strongly supported in research data services for planning, data curation, and data access stages.

KEYWORDS

Research Data Services (RDS), Research Data Management (RDM), Data Lifecycle

ACM Reference format:

H. Khan, H.-C. Chang, J. Kim. 2018. Unfolding Research Data Services: An Information Architecture Perspective. In *Proceedings of The 18th ACM/IEEE Joint Conference on Digital Libraries, Fort Worth, Texas, USA, June 3-7 2018 (JCDL'18)*, 2 pages. <https://doi.org/10.1145/3197026.3203887>

1 INTRODUCTION

Scientists are attempting to address the challenges of the data deluge, while librarians are finding opportunities to help researchers plan, manage, store, preserve, and share data. Many funding agencies, including the National Science Foundation and National Institutes of Health, mandate that applicants submit a data management plan (DMP), which addresses funder policy, legal and ethical considerations, organization and documentation, data backup

and security, data sharing and archiving, and data access. In response to these requirements from the funding agencies, academic libraries have begun to implement research data services (RDS), which support researchers' data management needs and challenges. In essence, research data management "consists of a number of different activities and processes associated with the data lifecycle" [2].

As the volume and complexity of digital research data increases, so does the need to address the challenges of managing it. Data management has been within the purview of librarians, as they have actively engaged in providing access to data repositories and offering services, such as data curation and now RDS. In previous research, Tenopir et al. [5] identified two major categories of RDS: informational RDS and technical RDS.

However, no significant work has been done on how RDS is organized in different academic libraries and how much of the research data lifecycle is supported within RDS. The purpose of this study is to better understand how RDS is aiding researchers in North American academic libraries. Particularly, the following two research questions guide this exploration:

RQ1. What are the provisions that make up RDS?

RQ2. To what extent does RDS support the research data lifecycle?

2 METHOD

This study adopted an innovative approach in examining RDS websites through content analysis with an information architecture (IA) perspective.

2.1 Data Collection

We focused on examining RDS websites for North American academic libraries. The sample was derived from Association of Research Libraries (ARL) member libraries that participated in the 2013 ARL survey [3]; it included a total of 50 libraries.

2.2 Data Analysis

Doing a content analysis of what is featured through the lens of IA of RDS websites could provide insight into how these services are organized and labeled. Data was recorded through the process of studying the IA, which encompasses search, organization, labeling, and navigation within the website. Using IA principles [4], we

*Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

JCDL '18, June 3-7, 2018, Fort Worth, TX, USA

© 2018 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-5178-2/18/06.

<https://doi.org/10.1145/3197026.3203887>

labeled how each RDS website was organized and the services offered. We also labeled how the services support each stage and/or phase of the research data lifecycle.

To assess the intercoder reliability, two independent judges coded a quarter of randomly chosen cases. The intercoder reliability was assessed by Cohen's kappa, because each judge reached coding decisions in a qualitative and nominal manner. Based on the criteria suggested by Cohen [1], substantial agreement beyond chance was obtained between the two judges ($k = 0.821$ for RQ1 and $k = 0.781$ for RQ2).

3 RESULTS AND DISCUSSION

By reviewing the contents and structure of the RDS websites, this study identified the following three emerging themes in RDS.

3.1 Information Access RDS

Information Access RDS is intended to provide access to information resources. Such RDS includes data management guides, tutorials for best practices to help researchers to manage and share data, links to funding agency requirements, links to external RDS websites, and data citation guides.

Approximately a quarter of libraries in this study have organized their RDS website as a format of LibGuide, which pulls many different types of resources on a subject or topic together in one place. Across many academic libraries, such LibGuides provide a similar organization and template design; they are often organized under categorized tabs. Many of these LibGuides heavily borrowed content and pages from other RDS LibGuides.

3.2 Technical Support RDS

Technical Support RDS involves the availability of data repositories, including institutional repositories and disciplinary repositories, as well as tools and software to assist users with creating DMPs, depositing data, and analysing/visualizing data. Technical support is extremely limited in RDS to the DMPTool, with all of the samples offering this to users, followed by 96% offering a data repository. Only 34% of the sample offers technical support for data analysis and visualization.

3.3 Personalized Consultation RDS

Personalized Consultation RDS refers to in-person assistance and/or instruction on various aspects of data management and curation. Almost all RDS websites (98%) provide some personalized consultation option. However, the consultation descriptions on the websites were vague and did not state who would provide such consultation. The RDS websites that offer specific information indicate that consultation is provided by either subject librarians (12%) or a data management team (14%).

It should be noted that 22% of libraries provide geographic information systems (GIS) within RDS. These libraries offer the option to consult specifically with a geospatial data librarian who is in charge of maps and geospatial data collection. They also provide their users the option of consulting on general research data management or the option of contacting the Library Maps, Data, or GIS department for consultation.

3.4 Data Lifecycle Alignment in RDS

Under the three RDS provisions discussed in the previous section, it is noticeable that RDS supports several stages of the research data lifecycle. Upon consulting many data lifecycle models reported in the literature, it was determined that the majority of them consider the research data lifecycle to include data planning, data acquisition, data curation, data access, and data processing.

The results showed that only 30% of the sample of North American academic libraries that provide RDS support data acquisition; the areas for data acquisition include data capture, data collection, and ownership of data. It was found that 34% of the institutions provide some form of data processing, such as tools and tips for data pre-processing, data analysis, and visualization. We also found that more than half of our sample supports data curation process and practices, such as data documentation, metadata, long-term data storage, and data deposit into repositories, whereas three-fourths of the institutions support data access, which includes licenses for research data and data citation.

From the IA investigation and content analysis of RDS provisions, this study discovered that technical support remains limited in the number of services offered. All of our samples provide data management planning, which is the highest representation of a data lifecycle stage within RDS. Additionally, this study found that the scope of current RDS activities is not comprehensive enough to provide full support for the entire research data lifecycle.

4 CONCLUSION

Research data services are critical to the progress of 21st century science in which data must be properly collected, stored, and made accessible for long-term use. This preliminary study is beneficial in understanding the composition of services provided in RDS. Information Access, Technical Support, and Personalized Consultation are all different ways in which libraries provide RDS. Based on the definitions and labeling of the three provisions of RDS, this research found that 98% of libraries provide all three options for RDS. Technical support RDS is underrepresented in the number of features offered when compared to Information Access RDS. In the future, it will be interesting to see if Technical Support RDS begins to implement more support to researchers in the analysis and visualization of data.

REFERENCES

- [1] Cohen, J. 1960. A Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46.
- [2] Cox, A.M., & Pinfield, S. 2014. Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*, 46(4), 299-316.
- [3] Fearon, D., Guina, B., Pralle, B., Lake, S., & Sallans, A. 2013. Research Data Management Services, SPEC KIT 334. Washington, D.C.: Association of Research Libraries.
- [4] Rosenfield, L., Morville, P., & Arango, J. (2015). *Information architecture: For the webs and beyond* (4th ed.). Sebastopol, CA: O'Reilly Media.
- [5] Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*, 36(2), 84-90.