ISSN 2161-3974
Book or Product Review

Journal of eScience Librarianship 11(2): e644
DOI: https://doi.org/10.7191/jeslib.644

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

Book Review: Research Data Management and Data Literacies

Jiebei Luo, New York University, New York City, NY, US, jl14614@nyu.edu

Abstract

This review offers a critical outline of Koltay Tibor's Research Data Management and Data Literacies. Tibor's insights on the current landscape of data ecologies and key components in research data management provide a creative connection between data governance and data management. The review identifies several key challenges in building comprehensive data ecologies and echoes Tibor's recommendations for librarians on how to address these challenges and provide effective research data services in academic libraries worldwide.

Received: December 8, 2022 **Accepted**: December 14, 2022 **Published**: December 19, 2022

Competing Interests: The author declares that they have no competing interests.

As indicated by its title, *Research Data Management and Data Literacies* focuses on two key concepts: research data management (RDM) and data literacies. By introducing the current landscape of data ecologies and key components in data management, this book creatively connects data governance with data management, outlines insufficiency in existing policy support for data sharing, and discusses the importance of data quality. The book also offers potential applications for building data literacies for information and data professionals, which can be particularly valuable for academic librarians providing research data services in their institutions.

The first chapter follows Borgman (2016) and provides a clear definition of the scope of research data as "entities used as evidence of phenomena for the purpose of research or scholarship" (Borgman 2016, pg. 29). This definition not only captures the traditional and relatively narrow concept of data, i.e., the arbitrary collection of items existing in some medium, but also incorporates a much broader sense of data that extends to information, texts, and documents. For example, 'data' can be something that a computer can read, transform, and analyze. Such extended scope is reflected throughout the rest of the book, centering on how research data are managed and how the corresponding literacies should be structured.

The next chapter outlines a few vital steps in cultivating the formation of the culture and ecologies of research data. For example, the National Science Foundation (NSF) in the U.S. required the preparation of data management plans (DMPs) for all grantees beginning in 2011, and the League of European Research Universities (LERU) proclaimed a "culture change in the way stakeholders in the research, education, and knowledge exchange communities create, store and deliver the outputs of their activity" (LERU 2018, pg. 3), and the adoption of the Fair Access to Science also fostered broader open science and open data initiatives. These efforts in turn offer numerous opportunities for academic libraries to contribute to building the open data ecosystem.

Meanwhile, Koltay also discusses several key challenges in continuing building comprehensive data ecologies in this chapter, especially in terms of data sharing and reuse. Different from open access (OA) of journal articles, the recursive nature of data allowing its reuse in various ways leads to a more complex situation for open data to overcome. For dataset owners, the main obstacle to sharing data comes from the underlying incentives. Current open data deposits are mainly contributed by article submissions (as required by many academic journals), requirements for research funding proposals, and social norms across different disciplines. Yet, Koltay also points out that the key to building data ecologies lies in establishing infrastructures to support researchers through financial and personal incentive mechanisms. From the perspective of data repositories including academic libraries, challenges of building such infrastructures come from several different aspects. For instance, data repositories (librarians) will need to determine the value of the deposited data, ensure its quality, create effective documentation and metadata for long-term access and preservation, which are also highly dependent on the knowledge and data handling skills in the relevant disciplines.

Despite these challenges, today's academic libraries possess the potential to play an integral role in research data management and data openness. The book provides a set of insightful suggestions with more practical guidelines offered in Chapter 4.

In particular, the book quotes ten recommendations for librarians, offered by the Association of European Research Libraries (LIBER), on how to provide research data services. These recommendations shed important light on how RDM services can be developed in academic libraries worldwide, and they are as follows (Koltay 2021, pg.78):

1. Offer research data management support, related to data management plans and intellectual property rights, among other factors;
2. Provide metadata services for research data and engage in the development of metadata and data standards;
3. Create data librarian posts and develop professional staff skills for data librarianship;
4. Actively participate in the institutional development of research data and open data policies;
5. Collaborate and build partnerships with researchers, research groups, data archives, and data centers;
6. Provide services for storage, discovery, and permanent access throughout the life cycle of research data;
7. Promote research data citation by applying persistent identifiers to research data;
8. Provide an institutional Data Catalog or Data Repository, depending on available infrastructure;
9. Get involved in subject-specific data management practice;
10. Offer or meditate secure storage for dynamic and static research data in cooperation with institutional IT units and / or seek exploitation of appropriate services.

In addition, this chapter identifies an important gap between providing quality RDM services and librarians' current capabilities in the form of discipline-specific knowledge and data-handling skills. As noted by Perrier, Blondal, and MacDonald, "librarians had already been collecting and managing research data, but due to the insufficient discipline knowledge and data management skills, providing a high level of support to researchers had not been a regular part of their activities" (Perrier, Blondal, and MacDonald 2018 pg. 177-178).

For LIS professionals, especially librarians who regularly provide research data services in their institutions and those who are interested in data librarianship and developing education programs, the last two chapters of the book offer practical guidelines and important insights about the necessary skill sets and learning goals in the presence of research data.

Focusing on the research data life cycle, RDM services can be further broken down into three areas, including expertise, education, and curation. Accordingly, Chapter 6 suggests that librarians need to possess a conceptual understanding of data, be knowledgeable about the surrounding issues, be able to perform data-related tasks, e.g., extracting, cleaning, analyzing, visualizing, etc., and be familiar with related lag requirements and ethical considerations. Regarding education program development, two focal points are RDM and data literacy education. Specifically, RDM education concerns researchers, teaching staff members, students, and librarians, and aims to develop the ability of data management questions, build and participate in an RDM community of practice, and understand data management needs in their research domain. In the data literacy education part, Koltay shifts the focus to meeting the needs of the target audience that future data librarians will serve. Although past studies have come to various recommendations for developing data literacy education programs, a few key elements suggested by different studies are worth noting, including the knowledge of research data lifecycle, the ability to evaluate, manage, and describe research data, as well as the importance of connecting data literacy with other forms of literacies, such as visual literacy and media literacy.

Competing Interests

The author declares that they have no competing interests.

References

Borgman, Christine L. 2016. *Big data little data no data: scholarship in the networked world*. MIT Press.

LERU. 2018. Open science and its role in universities. A roadmap for cultural change. Retrieved from https://www.leru.org/files/LERU-AP24-Open-Science-full-paper.pdf.

Perrier, Laure, Erik Blondal, and Heather MacDonald. 2018. "Exploring the experiences of academic libraries with research data management: A meta-ethnographic analysis of qualitative studies." *Library & Information Science Research* 40(3-4): 173-183. https://doi.org/10.1016/J.LISR.2018.08.002.