



Review

Book(s) Review: The Medical Library Association Guide to Data Management for Librarians and Data Management: A Practical Guide for Librarians

T. Scott Plutchak

University of Alabama at Birmingham, Birmingham, AL, USA

The Medical Library Association Guide to Data Management for Librarians

Edited by Lisa Federer,

Rowman & Littlefield

Copyright 2016 by Medical Library Association

[978-1-4422-6428-1](tel:978-1-4422-6428-1)

Data Management: A Practical Guide for Librarians

By Margaret E. Henderson

Rowman & Littlefield Publishers, 2017

[978-1-4422-6439-7](tel:978-1-4422-6439-7)

Correspondence: T. Scott Plutchak: tscott@uab.edu

Keywords: book review, data, research data, data management, data curation

Rights and Permissions: Copyright Plutchak © 2017



All content in Journal of eScience Librarianship, unless otherwise noted, is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Paradoxically, somewhere around the time that James Gleick published *The Information: A History, A Theory, A Flood* (Pantheon Books 2011), we moved beyond the information age and into the world of Big Data. For the general public, this means that vast amounts of personal data gathered by technology companies as we roam the internet, data sold and analyzed and resold, and reanalyzed and then used to predict our wants and needs, giving marketers astonishing new capabilities for selling us things and government agencies diabolical new tools for tracking our movements and behaviors.

For librarians, however, Big Data means Research Data, and that open data will lead to astonishing advances in scientific discovery as the raw data underlying the peer-reviewed literature becomes widely available. The opportunity for librarians is that curating this data – organizing it, archiving it, making it widely shareable and usable – requires specialists who can understand and master the complexities involved in taming it. They are rushing to fill the gap.

As librarians see the need for developing new services to assist investigators, they also recognize the need to advance their own skill sets and to think strategically about the best ways to connect with the members of their communities. A slew of books have been appearing with the intention of assisting librarians in expanding their knowledge base and thinking creatively about how to move forward.

Two recent entries offer complementary approaches. *The Medical Library Association Guide to Data Management for Librarians*, edited by Lisa Federer, provides a broad-brush overview of data management issues. Margaret Henderson's *Data Management: A Practical Guide for Librarians* is a step-by-step manual for thinking about, planning for, and implementing research data management services. Federer's book is most useful in helping readers gain a better understanding of the broad range of interrelated issues that fall under the general heading of data management. For librarians who have made the commitment to develop RDM services, or are contemplating taking that step, Henderson's book provides an invaluable roadmap.

The *MLA Guide* contains 15 chapters, arranged in three parts: Theory and Foundations, Data Management Across the Research Data Life Cycle, and Data Management in Practice. Federer is a highly regarded research data informationist at the National Institutes of Health (NIH) Library, and she has assembled a stellar group of authors. Anyone who has spent time investigating the world of research data management will recognize many of them as among the pioneers, leading the way as librarians and their libraries take on these new roles.

As expected in a book with many authors, the chapters vary in style and quality, but overall it's successful in painting a picture of the broad complexity of issues involved in dealing with research data. There are excellent chapters on the rise of data journals, the principles of data science, the relationship of traditional archival practices to the challenges of data curation, and the tools and practices for effectively visualizing data. There are, however, a few lapses. In the chapter on How to Learn More, a digression into self-directed learning theory seems a little superfluous for a book in which each chapter is accompanied by extensive bibliographies and references. The chapter on Library Infrastructures for Scholarship at Scale buries itself under theoretical jargon to finally make the obvious points that different disciplines have different data management needs and library support for scholarship needs to involve more than just data management. They're important points, but could have been made more economically.

The book is particularly effective in making the case that librarians in libraries of all types and sizes, including hospital libraries with perhaps only one professional librarian and a staff member, can and should become involved in data management services. The case study from the Health Sciences Library at NYU Langone Medical Center shows how an opportunistic entrepreneurial approach can lead to effective engagement when librarians are listening carefully to the expressed needs of the members of their communities. Readers of the MLA Guide will be reminded frequently that developing effective services requires this kind of listening and engagement along with collaboration with other players across the institution. Librarians cannot do this work in isolation.

That message is even more explicit in Henderson's *Data Management*, as she emphasizes throughout that "Librarians working in research data management need to add value, not work, to a researcher's project," and "Librarians need to remember that researchers don't think about their work the way librarians do." Henderson, who is the director of research data management at the Virginia Commonwealth Libraries, has a long history of work in the field and it shows in her no nonsense methodical step-by-step approach to developing services.

Data Management is much more tightly organized than the *MLA Guide*. While each of her chapters could stand alone as it addresses its topic, each also leads logically to the next, beginning with an understanding of the nature of research data, to the principles involved in its effective management, to the analysis of the needs of one's own institution and the development of services that are tailored to address those needs. She includes an excellent glossary and is very good at defining terms, essential in an area that is developing rapidly. As she says in the first sentence of Chapter One, "When you mention data to people, it is important to be clear what you are talking about."

Henderson is particularly good at making the case that librarians should take the lead in their institutions in taking on this work. While recognizing the need for librarians to expand their knowledge and skills, she emphasizes how traditional librarian skills are essential in underpinning the work that data librarians do. Her discussion of the ways in which the traditional reference interview provides a theoretical and practical grounding for the data interview is particularly compelling. A detailed description of how to talk with researchers about their projects is a high point.

Both books start out emphasizing the potential importance of data management services for the future of libraries. In the last few years many articles have appeared making a similar case. As sourcebooks, both volumes can serve as excellent reference materials. With Henderson's many practical examples and extensive bibliographies in particular, it is difficult to imagine a question that a data librarian (or aspiring data librarian) might have that they won't be able to find some help with here.

At the core of both books is the recognition that effective data management services can only be developed when the librarians involved are deeply engaged with their communities – listening closely to the expressed needs, and tailoring their offerings to the unique circumstances of their institutions. Too often, as librarians have moved into various areas of scholarly communication support, they have developed the services they think faculty and students ought to have and then had to work hard to market those services. Sadly, that approach often leads to wasted effort and frustrated librarians. The authors of these books

bring an important corrective to that, encouraging librarians to educate themselves first, to engage deeply with the members of their communities, and then to develop carefully tailored services to address clearly identified needs. Regardless of where a librarian may be in their own grasp of data management issues, they will surely benefit from spending time with either or both of these books.

Disclosure

The author reports no conflicts of interest.

References

Gleick, James. 2011. *The Information: A History, A Theory, A Flood*. New York: Pantheon Books.