Book Review:

The Data Librarian's Handbook

Robin Rice and John Southall, eds. London: Facet Publishing. 192 pp. £54.95. ISBN 978-1783300471

Ann Glusker

In the acknowledgements for this volume, editors Rice and Southall note that they were worried, in the course of writing, that "the field of data librarianship was changing faster than we could even fix our knowledge onto the page." (p.IX). They also note, in their preface, the growing number of books written recently about research data support, as librarians, particularly in academia, are increasingly called to provide this service. With the speed of change and the number of competing volumes available in the area of data librarianship, is it worth obtaining and delving into this one? The answer is a resounding YES. There are still very few books which take a hands-on approach to data librarianship (another excellent example being Margaret Henderson's *Data Management: A Practical Guide for Librarians*), so this volume makes an important contribution to the field.

The chapters included are as follows: Data librarianship: responding to research innovation; What is different about data?; Supporting data literacy; Building a data collection; Research data management service and policy: working across your institution; Data management plans as a calling card (including 8 vignettes from varied institutions and disciplines); Essentials of data repositories; Dealing with sensitive data; Data sharing in the disciplines; and Supporting open scholarship and open science. An extensive list of references and an index are also included.

The audience for this book is intended to be both practicing data librarians/research data support professionals, and teachers and students in library and information schools. As such, it's conceived as a primer, and able to be used for instruction. It includes very convenient "Key take-away points" at the end of each chapter, as well as "Reflective questions" that can be used with students. The tone of the writing is straightforward, engaging and accessible, and there are practical suggestions throughout the text, including some very useful sets of bullet-pointed lists. An advantage is that the authors have experience in and address both the US and European/British contexts, which gives this work a more international scope than some others.

Another helpful feature is that, rather than having a list of suggested URLs at the end of chapters, they are embedded as they occur in the text, and so as the reader engages with the material there are many side explorations available which enhance understanding (or illuminate a point—try the authors' suggestion of searching Twitter for "Lost USB"). In fact, they suggest that readers use the community resource "Open Research Glossary" as a companion to the book (http://www.righttoresearch.org/resources/openresearchglossary/).

While the authors cover a wide range of topics, in a well-delineated framework, they also engage directly with the complexity of each topic, demonstrating first-hand insight into the research data life cycle and its challenges. An example is this one, on the challenges of assessing denominators for evaluation:

> "If one-third of a new petabyte store is filled within a couple of months is that a good or bad rate of uptake? If three-quarters of departments have created collections in an open data repository after the first year, is that good or bad? Is it more important to get all the departments using it or to get more collections from existing departments? Does 30 downloads in a month mean a dataset is popular? If three principal investigators have sought advice from the library about a DMP for a research proposal in a month, is that a good rate of uptake? How many others have written plans without consulting the library or IT service; does it matter? How well are referrals working; are there gaps in the referral network? How many researchers or research groups are following and updating their plans after they receive funding?" (p.82)

The authors don't, of course, answer all of these questions, but they give resources and ideas for how to approach them (for example, discussing the use of benchmarking and standards and identifying some relevant web sites), and treat the many other topics in the book with similar thoughtfulness and thoroughness.

In fact, in the course of dealing with these aspects of data librarianship, the authors provide some very interesting and even unusual insights, based on their combined 30+ years as research data support professionals. An example is the ways in which the work of data archives and academic research user services are converging: "…as data archives have sought to emulate the reader services role of academic libraries, the latter have also begun to emulate the role of user support" (p.14). Another is the identification of drivers of support services as being influenced by "top-down drivers" (funders) as opposed to "bottom-up drivers" (institutions); taking into account this dichotomy can help determine strategy (pp. 69-73). The most interesting to me was the section on understanding how researchers view their research (p. 122). The idea that researchers have an emotional bond to their work, and may not be psychologically ready to share their data, to which they may have personal and deep attachments after the intense work required to create it, was enlightening to me, and is rarely discussed. This (often unacknowledged) attachment can be a crucial barrier to the relationship with the data librarian/professional, and as such we need to understand it, and how to work with and honor it.

One topic I would have liked to see discussed in this volume is strategies for continuing professional development, such those outlined by Goben and Raszewski in their (mainly US-focused) chapter "Data 101: Learning and Keeping Current in Data Management Skills". This is in large part because I'd be so interested in hearing Rice and Southall's ideas, while I admit that the purview of their book is more about the direct practice of providing research data support than data librarianship as a professional (and ever-changing) path. Nevertheless, the authors make an important contribution to those of us on that path. Since training programs in data science and data librarianship are still in their early stages, most of us in this field have had to pick up training on the job (and for many, without any background in creating or using research data). It can often feel like learning a language by ear; you can get quite good at communicating, but without understanding the grammar and structure of the language, there's only so far that you can advance. This volume provides the equivalent structure, acting as a textbook for the language of data librarianship, and filling in details which enhance our fluency.

In sum, this book provides a valuable reference source both for the beginner and the more experienced practitioner, giving background and suggestions for practice that may be new to them. I highly recommend it!

References

2/3 Glusker, Ann (2017) Book Review: The Data Librarian's Handbook, *IASSIST Quarterly* 41 (1-4), pp. 1-3.
DOI: https://doi.org/10.29173/iq14

Goben, A. & Raszewski, R. (2016). "Data 101: Learning and Keeping Current in Data Management Skills". In Federer, L., ed. *The Medical Library Association Guide to Data Management for Librarians*. Lanham, MD: Rowman & Littlefield.

Henderson, M. (2017). *Data Management: A Practical Guide for Librarians*. Lanham, MD: Rowman & Littlefield.

Ann Glusker PhD MPH MLIS
glusker@uw.edu
Librarian/Research & Data Coordinator
National Network of Libraries of Medicine, Pacific Northwest Region
University of Washington
Seattle, Washington, USA

DOI: https://doi.org/10.29173/iq14