



*Evidence Summary*

**Data Librarians' Skills and Competencies Are Heterogeneous and Cluster into Those for Generalists and Specialists**

**A Review of:**

Federer, L. (2018). Defining data librarianship: A survey of competencies, skills, and training. *Journal of the Medical Library Association* 106(3), 294–303. <https://doi.org/10.5195/jmla.2018.306>

**Reviewed by:**

Scott Goldstein

Web Librarian

Appalachian State University Libraries

Boone, North Carolina, United States of America

Email: [goldsteinsl@appstate.edu](mailto:goldsteinsl@appstate.edu)

**Received:** 29 Oct. 2018

**Accepted:** 27 Dec. 2018

© 2019 Goldstein. This is an Open Access article distributed under the terms of the Creative Commons-Attribution-Noncommercial-Share Alike License 4.0 International (<http://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly attributed, not used for commercial purposes, and, if transformed, the resulting work is redistributed under the same or similar license to this one.

DOI: 10.18438/ebliip29516

---

**Abstract**

**Objective** – To better define the skills, knowledge, and competencies necessary to data librarianship.

**Design** – Electronic survey.

**Setting** – Unknown number of research institutions in English-speaking countries with a focus on North America.

**Subjects** – Unknown number of information professionals who follow data-related interest group electronic mail lists or discussions on Twitter.

**Methods** – Author distributed an electronic survey via electronic mail lists and Twitter to information professionals, particularly those in biomedicine and the sciences, who self-determined that they spend a significant portion of their work providing data services. The survey asked respondents to rate the importance of various skills and expertise that had been selected from a review of the literature. In addition to other quantitative analysis, author performed cluster analysis on the final dataset to detect subgroups of similar respondents.

**Main Results** – 82 valid responses were received. Most respondents supported more

than one academic discipline and spent at least half of their time on data-related work.

Competencies in the “Personal Attributes” category (such as interpersonal, written, and presentation skills) were rated as most important, while those in the “Library Skills” category were rated as least important. A cluster analysis detected two groups that could best be described as subject specialists and data generalists. Subject specialists focus on a smaller number of disciplines and view a smaller number of tasks as important to their work compared to data generalists. In addition, data generalists are more likely to report spending most of their time on data-related work.

**Conclusion** – Data librarianship is a heterogeneous profession with many skillsets at play depending on the work environment, but the existence of two overarching subgroups – subject specialists and data generalists – deserves further study and may have implications for a number of stakeholders. Hiring institutions may consider the breadth of their user population’s needs before recruitment. Educational institutions as well as other on-the-job training opportunities may do well to focus more on “soft skills” as this is deemed more important by data librarians.

**Commentary** – In the past decade, there has been growth in the number of libraries offering data services, defined as services to researchers in relation to managing data. Examples of data services include data management guidance, data curation, and data visualization. Whereas Tenopir et al. (2012) found that only a small minority of U.S. and Canadian libraries offer any sort of services, a recent content analysis of library websites by Yoon and Shultz (2017) has revealed over 180 schools with services in place, though to varying degrees. Other scholarship has focused on the competencies required of librarians in data services roles, such as Xia and Wang’s (2014) analysis of social science data librarian job postings. The author’s survey contributes to the literature by asking self-designated data librarians in biomedical and scientific fields how these

competencies and skills are actually utilized in practice.

This summary relies on Boynton and Greenhalgh’s (2004) critical appraisal tool. Two aspects of the paper are worth highlighting here. First, the author performs cluster analysis on the categorical survey data to group respondents into the categories of generalist and specialist. This is an innovative and welcome analytical technique in LIS practitioner research. In addition, the data used in the analysis are openly available, well-documented, and reproducible. The questionnaire generally fares well against the critical appraisal tool. The items were developed in conjunction with a review of the literature and pilot tested. Skills were rated on importance using a five-point Likert scale ranging from “Not at all important” to “Absolutely essential” with an additional option for “Don’t know or N/A.” The biggest issue was the sampling method, which relied on a convenience sample from electronic mail lists and social media. This limits the generalizability of the results, although the author mentions this as a limitation and generally refrains from making wide-sweeping claims in the paper. (Differences between subject specialists and data generalists were tested with unpaired-samples t-tests and p-values were reported, but these should have been adjusted for multiple tests or else explicitly presented as exploratory.)

The study has unconventional implications for library science students and librarians looking to get into data librarianship. For students, it suggests they may be served more by developing soft skills and seeking discipline-specific skills rather than focusing on library-specific data management or curation courses. For librarians, on-the-job training and professional development opportunities, especially in a specific discipline if they wish to specialize, might prove more beneficial. Strong comfort with self-education may be a highly valuable skill to develop and certainly one to promote during the hiring process. The study also highlights potential barriers to data librarians that deserve further study. For instance, an open-ended survey comment

indicated that “our researchers have shown a strong bias towards working with ‘one of their own.’” It is unclear where this reluctance comes from, but it may be from a lack of awareness of the skills and competencies data librarians possess, suggesting that proactively demonstrating what they have done and can do to assist researchers would do much to reverse misconceptions.

## References

- Boynton, P. M., & Greenhalgh, T. (2004). Selecting, designing, and developing your questionnaire. *BMJ*, 328(7451), 1312–1315.  
<https://doi.org/10.1136/bmj.328.7451.1312>
- Tenopir, C., Birch, B., & Allard, S. (2012). Academic libraries and research data services: Current practices and plans for the future [White paper]. Retrieved October 26, 2018, from Association of College & Research Libraries:  
[http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir\\_Birch\\_Allard.pdf](http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf)

Xia, J., & Wang, M. (2014). Competencies and responsibilities of social science data librarians: An analysis of job descriptions. *College & Research Libraries*, 75(3), 362–388.  
<https://doi.org/10.5860/crl13-435>

Yoon, A., & Schultz, T. (2017). Research data management services in academic libraries in the US: A content analysis of libraries' websites. *College & Research Libraries*, 78(7), 920–933.  
<https://doi.org/10.5860/crl.78.7.920>