

Everyone's a data librarian now

Genifer Snipes

University of Oregon, Eugene, OR

Abstract: As data issues become increasingly important in education and everyday life, all librarians should consider data literacy an important skill to add to their professional repertoire.

Author Bio: Genifer Snipes is the Business and Economics Librarian at the University of Oregon, where she is interested in applied business research and teaching. Specifically, her work focuses on marketing and entrepreneurial research, as well as the growth of data-driven business librarianship.

Keywords: *databrarianship, big data, data literacy, information literacy, data librarianship*



This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

When I graduated from library school, I expected to enter a changing profession in which shrinking budgets, changing technologies, and evolving classroom pedagogies were converging to change how librarians did their jobs. I did not expect, however, that five years down the road I would be waving the data services flag. In that time, I have come to believe that supporting data-related activities is one of the most important roles librarians will play in coming years. As technology permeates virtually all aspects of academic, professional, and personal life, its natural product – data – is similarly growing in importance. The importance and pervasiveness of data in higher education, however, is easy to underestimate. This column will discuss the role of data in academic research and education, looking specifically at the implications for librarians in their roles as collection developers, educators, researchers, and subject specialists. The column concludes with ways for new librarians to add data skills to their professional repertoire.

What is data? In general terms, data is a collection of individual pieces of information that exists to provide insight or information about its subject. While the definition is simple enough, the reality of data is more complex. It's no exaggeration to say that today, data can be anything. We all know about the U.S. Census and scientific research outputs, but there's so much more. The length of your walking stride? That's data. Which storefronts shoppers look at the longest? Data. The changing use of the term "Dark Ages" in historical monographs? Also data. With the development of new technologies like the Internet of Things and optical character recognition, everything is quantifiable.

While STEM and social science disciplines have relied upon quantitative data for years, the spread of data gathering and analysis tools in the past decade has enabled other disciplines to incorporate data-driven methods into their work, often in ways not previously possible. Even historically qualitative fields like history and philosophy are applying quantitative techniques to the written word, allowing them to pursue projects like relationship mapping and moral decision-making analysis. Nor does the importance of data end at the academy's gates. Researchers are eager to translate this wealth of information into usable insights. Businesses, governments, and

other organizations are equally interested in putting those insights into practice. These organizations, however, face a massive shortfall of workers capable of analyzing and applying that data to solve problems, which colleges and universities have been slow to address (Manyika, et al., 2011). A 2017 report from the Business-Higher Education Forum found that 69 percent of employers prioritize hiring data literate job candidates, while only 23 percent of university leaders say they're producing data literate graduates (2017).

Unsurprisingly, the areas of librarianship in which data is making the most high-profile changes are in research and instruction libraries. These changes mean that English subject librarians are being called upon to teach students to scrape article repositories for textual analysis projects, business librarians are assessing, purchasing, and coordinating the storage of datasets for PhD projects, and art librarians could be assisting with NEA grant data management plans or helping artists obtain datasets for use in creating data-based creative works.

Although some areas of librarianship are experiencing the shift toward data-driven activities more visibly than others, no librarian can expect to avoid data issues completely. Technical services librarians are not only responsible for cataloging datasets, they are also analyzing usage data to guide collection development, tracking processing workflows, and quality checking metadata, all of which are data-related duties. Library administrators, in turn, must be able to synthesize the data their various departments produce, using it to make strategic decisions related to library budgets, project priorities, hiring, and marketing. Obviously, even at institutions that have been able to create dedicated Data Services departments or librarian positions, all librarians should be preparing for the day when data-related services become a regular part of their role—assuming that day hasn't already arrived.

Fortunately, even for those of us who didn't take data-related classes in library school, supporting data services doesn't require a degree in statistics. Because data literacy is such a significant issue for businesses and educational institutions, librarians have access to many

formal and informal learning options across a variety of platforms. Within the library world, groups including LITA and Library Juice Academy offer workshops and classes on topics including GitHub, altmetrics, GIS, and data management. An increasing number of books on data librarianship appear on our shelves, offering additional guidance and training for the interested librarian.

Looking beyond the library, tools for educating ourselves on data topics are everywhere. Virtually all online learning platforms offer courses on business and web analytics as introductions to data-driven decision making. These platforms include Lynda.com and Skillshare, which are often available through public libraries. Major MOOC providers like MIT OpenCourseWare, Coursera, and Udemy provide access to courses on basic data analysis, basic and advanced statistics, statistical software programs, and methods for presenting data analysis to non-specialist audiences. The educational options are endless, which means that no matter the need or skill level, resources are readily available to help expand our skills in a growing area of information services.

Concepts like Big Data, data-driven decision-making, and data literacy have become an increasingly prominent part of discussions about the future of business, education, and life in general, and this trend isn't likely to disappear anytime soon. As such, it is in the best interest of librarians, as information and educational professionals, to understand data issues, integrating them as seamlessly into our work as we have issues of information bias, copyright, and the impact of journal price inflation. Doing so will allow us to continue providing our institutions, communities, and patrons with the resources they will require when data-related information needs arise in their lives.

References

Business-Higher Education Forum. (2017). Investing in America's data science and analytics talent. Retrieved from <http://www.bhef.com/publications/investing-americas-data-science-and-analytics-talent>

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H. (2011, May). Big data: The next frontier for innovation, competition, and productivity. Retrieved from <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>