



Back to Basics: Considering Categories of Data Services Consults

Isaac Wink, University of Kentucky, Lexington, KY, USA, isaac.wink@uky.edu 

Abstract

Consultations are fundamental to data librarianship, serving as a vital means of one-on-one support for researchers. However, the topics and forms of support unique to data services consults are not always carefully considered. This commentary addresses five common services offered by data librarians—dataset reference, data management support, data analysis and software support, data curation, and data management (and sharing) plan writing—and considers strategies for successful patron support within the boundaries of a consultation.

Received: April 19, 2024 **Accepted:** July 7, 2024 **Published:** August 9, 2024

Keywords: data services, consultation, consult, reference

Citation: Wink, Isaac. 2024. "Back to Basics: Considering Categories of Data Services Consults." *Journal of eScience Librarianship* 13 (2): e931. <https://doi.org/10.7191/jeslib.931>.

The *Journal of eScience Librarianship* is a peer-reviewed open access journal. © 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

See <http://creativecommons.org/licenses/by/4.0>.

∞ OPEN ACCESS

Introduction

One-on-one support for researchers through consultations is an essential element of data librarianship in both theory and practice. The Medical Library Association Data Services Competency lists providing “training and consultation for data-related topics” as one of five essential skills for data services librarians (Federer et al. 2020), and in a 2018 survey of librarians conducting data-related work (n=82), 87% of respondents indicated that “one-on-one consultation or instruction” is either an absolutely essential or very important skill for data librarians (Federer 2018). These expectations are reflected in job ads for data services librarians, where providing consultation, support, or advice in research needs is one of the most frequently listed duties (Ramos Eclevia, La Torre Fredeluces, Lagrosas Eclevia, and Saguibo Maestro 2019).

Consults are the most basic level of research support that libraries offer to patrons; however, “basic” does not mean simple. The training that librarians receive in reference or other individualized support does not always address unique aspects of data services, which include data management planning, data analysis, and finding datasets for reuse. Nor does “basic” mean low value. A consult with a data services librarian may be the first time that a researcher frames concepts like data management as ongoing parts of the research process, and the strategies and resources that an effective librarian offers may be the starting point in significant improvement to a patron’s data competencies.

This commentary outlines some broad categories of data services consults that librarians may encounter and offers strategies for effective patron assistance in each case. The list is not meant to be exhaustive or prescriptive but rather to serve as a starting point for librarians reflecting on the structure of their consults. A rough categorization of topics can help librarians assess the level of support they can offer in a relatively brief consultation to make their work both impactful and sustainable.

Dataset Reference

In the case of finding data, the need for the librarian to ask the right questions to determine appropriate sources is critical, given that data reference interviews are, as Kristin Partlo has observed, “rife with problems of semantics and definitions (e.g., what exactly is meant by ‘raw data’), of process and of expectations” (2010). An effective librarian not only helps a patron identify a suitable dataset but also uses a reference consult as an instructional opportunity that helps the patron fit their data into their overall research process—for example, adjusting a research question to one that is answerable with available data.

Familiarizing patrons with the relevant landscape of data in their discipline can be another important outcome of this type of consult. Data services librarians can model the search process across a variety of different platforms and ask patrons to apply their own domain knowledge to narrow results. Depending on the discipline and type of data needed, the librarian might work with the researcher to identify organizations that collect and share relevant data, search for related scholarly publications that may include supplementary datasets, or explore generalist and domain-specific data repositories.

Finally, unlike with other scholarly outputs, the library is frequently not the only steward of datasets at an institution. Many university centers collect or facilitate access to specialized datasets that may interest patrons. At the author's institution, the Center for Clinical and Translational Science manages access to the National Institutes of Health's All of Us Research Hub, which provides health information for nearly a million participants. Knowing about these specialized resources enables librarians to connect patrons to data sources unavailable on the open web.

Broad Data Management Support

Patrons frequently come to data services librarians with a nebulous set of data management issues. They may have identified a series of problems in their relationship to their data—they lose track of their files, they struggle to make sense of analyses they produced in the past, they cannot understand or access materials produced by other members of their research team—but they may not yet have identified the particular practices that create those problems.

In a consult of this type, a data services librarian's essential value is the ability to break broad problems down into discrete and manageable ones. Through a series of questions, the librarian can help the researcher determine which elements of their workflow are leading to issues and then identify solutions. For instance, when a researcher expresses frustration with being able to understand data or analyses produced by other team members, the librarian may ask about the team's data documentation practices and make recommendations to improve current practices or propose new ones.

A data services librarian can take a strengths-based approach to data management consults by first identifying a researcher's existing practices and then proposing solutions that build upon them, making the solutions more likely to be implemented long-term. Kristin Briney's definition of data management as "the compilation of many small practices that make your data easier to find, easier to understand, less likely to be lost, and more likely to be usable during a project or ten years later" (2015) usefully reframes data management as a set of discrete activities rather than as a single overarching attribute.

Worksheets and template documents provide a useful starting point in helping researchers implement proposed data management practices. For example, the librarian can provide the researcher with a copy of a file naming convention worksheet (Briney 2023) or a README template file (Cornell Data Services 2024) during the consult and use it to drive a discussion about what the particulars of implementation may look like. Decisions are ultimately the responsibility of the researcher, but the librarian can make suggestions and answer questions.

Data Analysis and Software Support

The boundaries of data analysis or software consults can be difficult to negotiate in terms of both breadth and depth. Regarding breadth, individual librarians or teams of librarians cannot possibly have expertise in

every tool that researchers at their institution use or every method of analysis, meaning librarians must have a plan for consults requesting support that go beyond their knowledge. In promoting their consult services, they may choose to draw boundaries around the specific software or modes of analysis they support.

However, librarians may be able to offer meaningful support even to researchers working with software or methods unfamiliar to them. For instance, a data services librarian with statistics expertise can provide recommendations on tests to apply to data, even if they are not familiar with the specifics of implementing it in the researcher's chosen software environment. Oftentimes, the process of answering questions posed by a librarian can help a researcher view their data in new ways and achieve new insights.

Regarding depth, researchers may request a higher level of support than what librarians can reasonably provide, such as intensive data cleaning or writing code. When librarians want to provide more intensive support for a project, they may consider formalizing support by requesting co-authorship on a publication or inclusion on a grant proposal. These arrangements establish librarians as equal research partners in their institutions while also establishing limits that make librarians' work sustainable.

Data Curation

Data curation is an important step to take prior to sharing research data to ensure that it will be interpretable and usable by others in the future. It requires intimate knowledge of a dataset, limiting the scope of support a data services librarian can provide in a consult setting. However, the librarian can still often provide an important orientation to researchers who are new to data curation.

As with data management, librarians can take a strengths-based approach to data curation consults by asking questions to determine what curation activities the patron has already done. One option for framing those activities in a broader context is by using a tool such as the CURATE(D) Checklist from the Data Curation Network (DCN) to identify what additional steps a researcher should consider taking (Data Curation Network 2022). To provide discipline-specific support, the librarian can identify submission requirements for relevant data repositories to create an augmented checklist. Librarians can also review or direct researchers to the DCN's Data Curation Primers, which provide more specific information for particular data types, file formats, and key concepts.

Data Management (and Sharing) Plan Review

Libraries are increasingly providing one-on-one support for researchers writing data management (and sharing) plans, or DM(S)Ps, a required part of many grant applications for federally funded research. The need for DM(S)P support has grown with the rollout of the National Institutes of Health's Data Management and Sharing Policy and will grow further following the Office of Science and Technology Policy directing more funders to adopt similar requirements (National Institutes of Health 2023; White House Office of Science and Technology Policy 2022).

A DM(S)P is a compliance document, making individualized assistance in writing one a rather unique library service. Some researchers may view it as just another paperwork hurdle that must be overcome to receive funding. Their primary motivation in receiving library support is often a desire to produce a document that will pass muster with the funder, while a desire to improve their data management and sharing practices may be secondary or non-existent.

However, DM(S)P consults offer data services librarians the opportunity to reach new audiences of researchers who may not otherwise engage with library resources. Discussing data repository options, for example, provides an opening for describing the data curation practices and the services the library offers. Similarly, a researcher may express an interest in improving their data management and sharing workflows beyond the scope of a two-page DM(S)P, leading to further meetings with the librarian.

Data services librarians offering DM(S)P consults should maintain an awareness of other institutional stakeholders and be prepared to refer researchers to other offices rather than attempt to answer policy questions on which they are not well-versed. For instance, a librarian benefits from recognizing when a dataset may contain personally identifiable information but should refer researchers to their institutional review board or other appropriate body rather than attempt to provide answers beyond their expertise.

Conclusion

The particulars of their roles aside, data services librarians should keep two key facts in mind when providing consults. First, consults should make patrons active participants in the services they receive. Patrons may come to librarians with a “just do it for me” attitude to fixing code, cleaning data, or writing data management plans, as well as a sense of urgency that their issue be resolved as quickly as possible. As advocates of Slow Curation have noted, “someone else’s lack of planning is not our emergency” (Thielen, Marsolek, and Narlock 2023). A consult is, by nature, a relatively low level of engagement on a given project, and it is most effective when a patron leaves empowered to take the next necessary steps on their own. To keep their work sustainable, librarians should be prepared to say no to some service requests.

Second, maintaining connections to the institutional community of data stewards will help data services librarians know when a patron may benefit from being referred elsewhere. Data services librarians can cultivate impactful relationships with other campus units, including information technology, high-performance computing centers, research institutes and centers, and more. Patrons may not be aware of these other resources, and making connections to them reduces redundancy in services and allows for increased specialization.

Consults are a distinct tool for data services librarians to offer individualized support that is not always possible in other instructional settings, and they are most effective when librarians think critically about the skills they bring to them. The five categories of data services consults laid out here may be helpful to

librarians not only in strategizing for individual consults but also in communicating the value they provide to researchers.

Acknowledgements

The author would like to thank Colleen Barrett, Jennifer Hootman, and Stacey Greenwell for reading previous drafts of this commentary.

Competing Interests

The author declares that they have no competing interests.

References

- Briney, Kristin. 2015. *Data Management for Researchers: Organize, Maintain and Share Your Data for Research Success*. Research Skills Series. Exeter, UK: Pelagic Publishing.
- . 2023. *The Research Data Management Workbook*. Caltech Library. <https://doi.org/10.7907/z6czh-7zx60>.
- Cornell Data Services. "Guide to Writing 'Readme' Style Metadata." Accessed April 19, 2024. <https://data.research.cornell.edu/data-management/sharing/readme>.
- Data Curation Network. "The DCN CURATE(D) Steps." Accessed April 19, 2024. <https://datacurationnetwork.org/outputs/workflows>.
- Federer, Lisa. 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>.
- Federer, Lisa, Erin Diane Foster, Ann Glusker, Margaret Henderson, Kevin Read, and Shirley Zhao. 2020. "The Medical Library Association Data Services Competency: A Framework for Data Science and Open Science Skills Development: Journal of the Medical Library Association." *Journal of the Medical Library Association* 108 (2): 304–309. <https://doi.org/10.5195/jmla.2020.909>.
- Garrison, Betty, and Nina Exner. 2019. "Data Seeking Behavior of Economics Undergraduate Students: An Exploratory Study." *Reference & User Services Quarterly* 58 (2): 103–113. <https://doi.org/10.5860/rusq.58.2.6930>.
- Hoffman, Starr. 2015. "Data Reference and Instruction in Journalism and the Social Sciences." *DttP: Documents to the People* 43 (2): 14–17. <https://journals.ala.org/index.php/dttp/issue/view/issue/603/360>.
- Johnson, Andrew. 2023. "A Tiered Model for Data Management, Curation, and Sharing Support in Grant Proposals and Budgets." *Journal of eScience Librarianship* 12 (2): e702. <https://doi.org/10.7191/jeslib.702>.
- National Institutes of Health. "All of Us Research Hub." Accessed April 19, 2024. <https://www.researchallofus.org>.

- National Institutes of Health. 2020. "NOT-OD-21-013: Final NIH Policy for Data Management and Sharing." <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>.
- Partlo, Kristin. 2010. "The Pedagogical Data Reference Interview." *IASSIST Quarterly* 33 (4): 6–10. <https://doi.org/10.29173/iq884>.
- Ramos Eclevia, Marian, John Christopher La Torre Fredeluces, Carlos Jr Lagrosas Eclevia, and Roselle Saguibo Maestro. 2019. "What Makes a Data Librarian? An Analysis of Job Descriptions and Specifications for Data Librarian." *Qualitative and Quantitative Methods in Libraries* 8 (3): 273–290. <http://www.qqml.net/index.php/qqml/article/view/541>.
- Thielen, Joanna, Wanda Marsolek, and Mikala Narlock. 2023. "Conceptualizing Slow Curation." *Journal of eScience Librarianship* 12 (2): e740. <https://doi.org/10.7191/jeslib.740>.
- White House Office of Science and Technology Policy (OSTP). 2022. "Desirable Characteristics of Data Repositories for Federally Funded Research." Executive Office of the President of the United States. <https://doi.org/10.5479/10088/113528>.