Journal of eScience Librarianship
putting the pieces together: theory and practice

Commentary

Why You Need Soft and Non-Technical Skills for Successful Data Librarianship

Margaret Henderson

San Diego State University, San Diego, CA, USA

Abstract

There are many courses available to teach research data management to librarians and researchers. While these courses can help with technical skills, like programming or statistics, and practical knowledge of data life cycles or data sharing policies, there are "soft skills" and non-technical skills that are needed to successfully start and run data services. While there are many important characteristics of a good data librarian, reference skills, relationship building, collaboration, listening, and facilitation are some of the most important. Giving consideration to these skills will help any data librarian with their multifaceted job.

Correspondence: Margaret Henderson: margaret.henderson@sdsu.edu
Disclosures: The author reports no conflict of interest.

Introduction

> *Many day-to-day issues that my colleagues and I struggle with most are not discussed in literature. These issues include strategies for building relationships with busy researchers, demonstrating the value of the library perspective in this arena, advocating for improved research practices, and creating strategies for engaging with high-level administrators about these issues.*
>
> — *Heather Coates (2014, 52)*

There is more to providing research data services (RDS) than the technical skills necessary for data management. Reviews of job advertisements for data librarians show that there are many other skills listed. Eclevia et al. (2019) did a study of 104 job advertisements and found that interpersonal and communication skills, and the ability to work independently and collaboratively, occurred in job advertisements more often than knowledge of research data curation and management. Kahn and Du (2018) in a similar study of job advertisements, found the skills of research assistance and critical thinking/problem solving were in more job advertisements than knowledge of data sources or a statistical software package. Federer (2018) conducted a survey of librarians doing data related work, and found that communication skills, the ability to develop relationships with researchers, and other "soft skills", were especially relevant to their work.

"Soft skills", or interpersonal skills, and other non-technical skills, are very important when setting up research data management services, and continue to be important to the sustainability of services. Reference skills, relationship building, collaboration, listening, and facilitating access to de-centralized resources, are all important to the success of a service. Combining these skills with an understanding of the data life cycle, metadata, and some coding and statistics, will place librarians and libraries in a good position to support research at their institution. If you want to work with data, you will need to be an administrator and an entrepreneur, as well as a reference and instruction librarian, so data science knowledge is only a part of the job.

Reference

I have always advocated that reference skills are a librarian's superpower. Finding out what is really needed by users is a big part of what we do. For example, when I was working on a practicum with people in the university bioinformatics core, I had a chance to learn about searching the electronic health records (EHR). At the

time clinicians couldn't retrieve their own data, they had to request it through a programmer who worked on the database. He would do a search based on their request and send them the dataset, but it would be sent back because it didn't quite have what was needed. He would try again, and the dataset would be sent back with a few more comments, and so on. A good reference interview could have saved time and reduced the number of back and forth emails. In fact, when the university hospital was researching a new data warehouse for the EHR data, the request for proposals was going to include a data concierge to help people find the correct data.

Data scientist William Koehrsen (2018) confirms the importance of reference skills in a recent article about the non-technical skills needed in data science. His first step is to ask the right question or questions. For him, being able to ask the right question means he looks for the appropriate data for a project, rather than being overwhelmed by a mass of data.

Relationships

Our reference skills naturally lead into the ability to develop good relationships with different groups on our campus. We ask questions to learn more about what they are doing and what they need to do their work. And we work to support what they are doing, in research, teaching, or administration. I presented to the grant administrators from the schools and colleges at my institution, and then chatted with a few of them at other events on campus. Soon after, I had several faculty members contact me for help with data management plans, and they were referred by their grant administrator. By forging relationships, people feel comfortable recommending your expertise.

Because building relationships was such an integral part of what I did as director for RDS, I realized that creating relationship maps is an excellent way to show progress when assessing RDS. These maps can quantitatively express the strength of a relationship with different width lines or colors to show types of contact, meetings, emails, consultations, and how many of each. Figure 1 shows what could be the first year of relationships for a library data service, with outreach to sponsored programs and the provost in the form of emails to one or two people, but no further interactions. But institutional IT, the research office, and a researcher have darker lines with two arrows showing interactions with emails, and meetings with multiple people. With proper captioning, and maybe animation using a relationship mapping program like NodeXL, you could show changes in relationships over time, giving evidence of the growth and strength of your service.

Collaboration

Data librarians need to collaborate with people inside and outside of the library. Colleagues with subject expertise, knowledge of statistics, etc. could help with workshops. Researchers, schools and departments, and others with existing
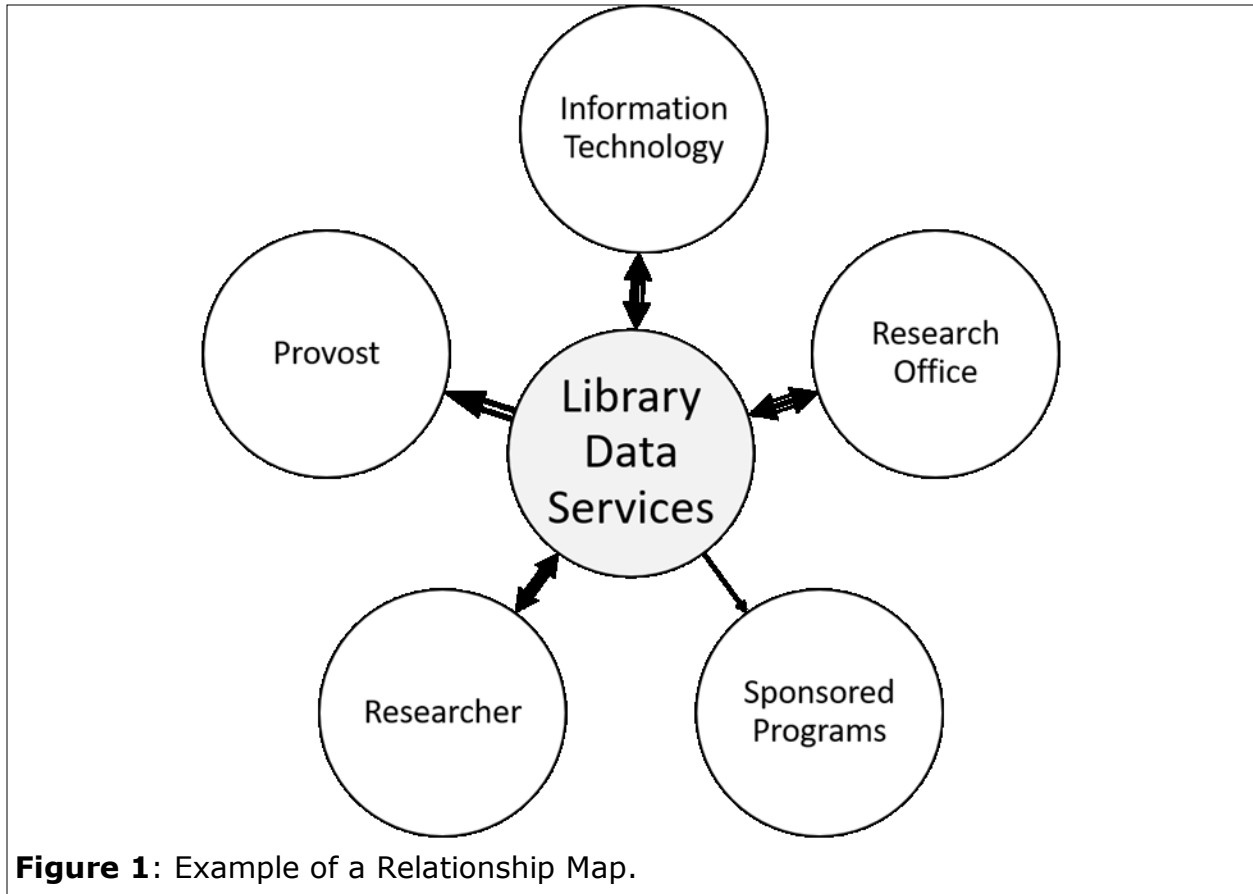
Figure 1: Example of a Relationship Map.

resources that are outside the library budget often have core services for everyone at the institution. I collaborated with the Statistics and Analysis service of the Department of Statistical Science and Operations Research to provide R training. They provided the trainer, and I organized the classes using a library classroom, handled the registrations, and provided food. In a collaboration with university Sponsored Programs I helped organize grant writing workshops, that included a session I taught on data management plans for grants. I also collaborated with data librarians at other institutions to hold data boot camps using real-time distance education resources to share expertise and teaching duties.

Before approaching a potential collaborator, be sure to think about what you are hoping to achieve. Developing personas for the people or groups you are interested in and using those personas to decide how to approach that person or group, is an excellent way to be sure you aren't at a loss for words when you meet people. Brianna Marshall and participants at the Midwest Data Librarians Symposium (2015) brainstormed a long list of potential collaborators or partners, and developed several personas and elevator speeches that can help you think about how to approach people and groups on campus. Betsy Humphreys (2018, 522-23) recommends that you have informed questions and ideas for joint action

ready when you first meet with a group, and to find something concrete and achievable to start a collaboration.

Listening

Good listening is the foundation for building relationships and fostering collaborations. In order to understand what a researcher needs, you have to listen for the whole message, check your understanding, and ask questions (Ross and Dewdney 1989, 33). And you need to avoid making assumptions. While working with a graduate student to develop a data entry form for an experiment, I listened to him go through the whole protocol once, so I understood what he was doing. Then we went through it again so I could ask questions to clarify the experimental steps, types of data being collected, and the measurements involved. A couple of times, I assumed things were being measured a certain way, but I made sure that when I entered the information in the form we were building, I reviewed my understanding with the student, and he was able to make sure I had the correct measurements.

Listening can be one of the hardest things to do, especially when you already have some data services in place. You might be interviewing a researcher or just chatting with a faculty member you meet in the hall. They might complain about what they can't get from the university or library, and how they need help with a DMP. While you want to correct them and tell them about the DMPTool (https://dmptool.org) is available, and IT has storage, and the university has a policy—this won't help. As Amanda Rinehart (2015) wrote, using active listening techniques can be very useful in these situations. You need to let them finish, show empathy and then try paraphrasing what they said so they know you understand. Then, you can suggest the solutions you have and offer your help. If you act defensive and cut them off, you will never learn all they need, and, without sympathy, they will remain frustrated. Remember that people do not really notice what the library offers until they actually need to do something, so keep up the marketing efforts and empathize with busy researchers and students.

Facilitation

Over my 30 plus years in libraries, one of the things I have regularly done at the reference desk is help people find something they need that is not a library resource. Helping students find tutoring services, or faculty find research equipment, or a visitor to campus find the proper building, are all part of facilitating access to services. In all my jobs, one of the first things I do is find out what the other departments in the library can help with, and then which departments around the institution can be of help to me, or those I am helping at the reference desk or in consultations. I also try to form relationships with the people working in the other areas, so I can refer people easily. Once, I worked with an Emergency Room research group that needed help on a large data collection project that covered multiple Excel spreadsheets. They did not think REDCap would work, although I was pretty sure it would. But I am not an expert,

so I put them in touch with REDCap administrators who were able to set things up in a much easier format than the original Excel. Even though I did not do much, I was pleased to be able to send people to the best services.

In an analysis of RDM surveys, Goben and Griffin (2019, 913) found that storage, sharing, and long-term access to data were the biggest worries of researchers, and few libraries are able to provide the hardware to offer these services. So, we need to develop relationships with institutional computing services to learn where to send researchers. As director of RDS, I met with the managers of various university and school IT services to find out what kinds of collaboration networks, as well as back up and long-term storage were available for researchers. While some services were only for associated school faculty, others were core services for everyone. By having a list of resources and contacts, I could tell researchers about available computing options.

Conclusions

These five skills are just the start of the non-technical skills data librarians need. In the report "Developing the Librarian Workforce for Data Science and Open Science", Federer et. al. (2020) include lifelong learning skills, like flexibility and adaptability, skills for developing programs and services, such as willingness to embrace failure, and other traditional library skills, like assessment and evaluation. All these skills are the foundation for the development of good research data services.

I'm not advocating that we ignore data science and technical skills. But I believe that if we give up the parts of our profession that differentiate us from data science, we devalue ourselves. We are bowing to the people who think librarianship isn't valuable because it is a female-intensive field and see computer and data science as more masculine so more valuable (Harris 1992, 16). Our combination of soft skills AND computing and data skills is what makes us more useful than a programmer or statistician who can't ask the right questions to figure out what a researcher needs. With our holistic view of the whole research process, we can be of more help than the grants administrator who just worries about policy compliance. We know how to work with all the groups involved with RDM and bring them together as a service that truly meets the needs of researchers.

References

Coates, Heather L. 2014. "Building Data Services From the Ground Up: Strategies and Resources." *Journal of eScience Librarianship* 3(1): e1063. http://dx.doi.org/10.7191/jeslib.2014.1063

Eclevia, Marian Ramos, Fredeluces, John C.L.T., Maestro, Roselle S. and Eclevia, Carlos L. Jr. 2019. "What Makes a Data Librarian? An Analysis of Job Descriptions and Specifications for Data Librarians." *Qualitative And Quantitative Methods In Libraries* 8(3): 273-290. http://qqml-journal.net/index.php/qqml/article/view/541

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, Sarah C. Clarke, and Maryam Zaringhalam. 2020. "Developing the Librarian Workforce for Data Science and Open Science." *OSF Preprints* January 16. https://doi.org/10.31219/osf.io/uycax

Goben, Abigail and Tina Griffin. 201). "In Aggregate: Trends, Needs, and Opportunities from Research Data Management Surveys." *College & Research Libraries* 80(7): 903-924. https://doi.org/10.5860/crl.80.7.903

Harris, Roma M. 1992. Librarianship: The Erosion of a Woman's Profession. Ablex: Norwood, NJ.

Humphreys, Betsy L. 2018. "How to earn a reputation as a great Partner." *Journal of the Medical Library Association* 106(4):521-528. https://doi.org/10.5195/jmla.2018.504

Khan, Hammad Rauf and Du, Yunfei. 2018. "What is a Data Librarian?: A Content Analysis of Job Advertisements for Data Librarians in the United States Academic Libraries." Paper presented at *IFLA WLIC 2018 – Transform Libraries, Transform Societies in Session 139 - Education and Training, Science and Technology, and Continuing Professional Development and Workplace Learning,* Kuala Lumpur, Malaysia. http://library.ifla.org/id/eprint/2255

Koehrsen, William. 2018. "How to Get the Right Data? Try Asking for It." *Medium: Towards Data Science,* July 26, 2018. https://towardsdatascience.com/how-to-get-the-right-data-why-not-ask-for-it-d26ced1bbd46

Marshall, Brianna, and MDLS15 participants. 2015. "Outline of Partnerships Session: Topic 3: Forming Partnerships on Campus." *Midwest Data Librarians Symposium,* University of Wisconsin, Milwaukee. UWM Digital Commons. https://dc.uwm.edu/mdls/2015/partnerships/8

Rinehart, Amanda K. 2015. "Getting Emotional About Data: The Soft Side of Data Management Services." *College & Research Libraries News* 76(8): 437-440. https://doi.org/10.5860/crln.76.8.9364

Ross, Catherine Sheldrick and Patricia Dewdney. 1989. *Communicating Professionally: A How-To-Do-It Manual for Library Applications*. New York:Neal-Schuman.