JENNIFER MUILENBURG

8

Data Services

Data services in libraries are not new, but they are related in many aspects to digital scholarship in libraries. Although the term "data services" tends to lean toward researchers working in data-intensive sciences, many of the practices that comprise data services mirror those necessary to digital scholarship research. To that end, the University of Washington's (UW) data services librarian and digital scholarship librarian have discussed the type of services they provide to campus users and have collaborated on campus programming and events. They also collaborate on planning future services. This chapter will outline the similarities and differences between data services and digital scholarship and will describe the data services assessment work, training program, and strategy for marketing services at the UW Libraries.

DEFINING DIGITAL SCHOLARSHIP AND DATA SERVICES

Before we can talk about digital scholarship and data services, we need to discuss some terminology and definitions. The concept of data services is confusing to many, even in the context of academic library services.[1] Similarly, as explained in the Association for Research Libraries' article "Digital Scholarship

From the Culture of Digital Scholarship in Academic Libraries, edited by Robin Chin Roemer and Verletta Kern (Chicago: American Library Association, 2019).

Support in ARL Member Libraries: An Overview," "Digital scholarship [DS] is not new and its panoply has evolved over time. DS is a shifting range of scholarly endeavor that can incorporate a number of definitions, methods, tools, and research outputs."[2] Even within this book, slightly different definitions of digital scholarship are used by different authors—despite the fact that all of us work at the same institution, if not on some of the same projects.

The term data services is also a common source of confusion, in part because "data" is a term that can mean multiple things. At UW, the Libraries has adapted a definition from the UK Data Archive: Data is "that which is collected, observed, or created, for purposes of analyzing to produce original research results. . . . Research data may be created in tabular, statistical, numeric, geospatial, image, multimedia or other formats."[3] In the book *Databrarianship,* the editors provide several examples of what data might mean in different disciplines before announcing a wide-ranging, catch-all working definition for the purposes of their text: "The data we are concerned with here are the product of taking that raw informational input and assembling it into a structure form for analysis. Data are a product of research as well as an input for research."[4]

There is also a lack of common, established definitions of several data-related terms, including "digital archiving" and "digital preservation," to name a few. For this reason, staff working in libraries may have different understandings of what it means to curate, archive, or share a digital item, unless specific institutional definitions are provided. This lack of common understanding extends to other disciplines are well. For example, a review of data management plans (DMPs) at UW in 2015 showed that terminology was an obstacle in implementing DMPs, especially around the terms "archive" and "sharing." For example, the review showed that researchers would use the word "sharing" when they meant anything from putting data on a shared drive in a lab, providing data upon emailed request, a verbal presentation at a conference, or depositing data in a disciplinary web-based repository with accompanying metadata.[5]

All of this is used to illustrate that when proceeding with descriptions of data or digital scholarship services at a particular institution, it will be in everyone's best interests to provide explicit definitions of various terms, along with relevant citations or links to specific software programs, metadata standards, or similar information so that all players have the same base information. In a nutshell: "the term digital scholarship itself is quite fluid and seems to offer many interpretations depending on a particular university's culture, institutional organization, and environment."[6] This means that the types of support services and education libraries can provide will also be in constant evolution. The case is similar with data services in libraries: although there are, generally speaking, a set of agreed-upon services that define "data services," the specifics of those services will vary not only over time, but at

each institution. Tenopir et al. listed the most common data services offered by libraries (items marked with an asterisk are services offered at UW), which include:[7]

- Consulting on data management plans*
- Consulting on data and metadata standards*
- Outreach with other data service providers on campus*
- Providing reference support in finding and citing data sets*
- Creating guides for finding data*
- Directly participating in research projects
- Discussing data services with others on campus*
- Training librarians and others on campus*
- Providing repository services, which includes:
  - Deaccessioning of data from a repository
  - Preparing data sets for deposit into a repository*
  - Creating or transforming metadata*
  - Identifying datasets that could be candidates for repositories

Working from the 2012 Associations of College and Research Libraries (ACRL) list of research data management activities in libraries, one can draw parallels to equivalent services for digital scholarship. In table 8.1, each of Tenopir's major data services activities are listed in the first column, with the second and third columns providing examples of what such services might look like for either research data services or digital scholarship services.

TABLE 8.1

Tenopir's Research Data Management activities and digital scholarship equivalents

Service	Research Data Management	Digital Scholarship
Consulting on data management plans	Library staff help researchers complete DMPs from various funding sources, including creating metadata and other documentation, finding appropriate repositories, and data curation. Example from NSF: https://www.nsf.gov/eng/general/dmp.jsp	Library staff help researchers complete DMPs from various funding sources, including creating metadata and other documentation, finding appropriate repositories, and digital object curation. Example from NEH: https://www.neh.gov/sites/default/files/2018-06/data_management_plans_2018.pdf
Consulting on data and metadata standards	Library staff help identify disciplinary metadata standards, if available, or general standards in their absence. www.dcc.ac.uk/resources/curation-reference-manual/chapters-production/scientific-metadata, www.dcc .ac.uk/resources/metadata-standards	Library staff help identify disciplinary metadata standards, if available, or general standards in their absence. http://dh101.humanities.ucla.edu/?page_id=35, www.dcc.ac.uk/resources/subject-areas/social-science-humanities

(cont.)

TABLE 8.1 (cont.)

Tenopir's Research Data Management activities and digital scholarship equivalents

Outreach with other data service providers on campus	Library staff works for information technology departments, scholarly communications and publishing experts, as well as experts in various departments and/or labs on campus. This keeps both sides aware of current research, any opportunities for new or improved services and tools, and helps create partnerships and collaborations for the entire campus.	Library staff works for information technology departments, scholarly communications and publishing experts, as well as experts in various departments and/or labs on campus. This keeps both sides aware of current research, any opportunities for new or improved services and tools, and helps create partnerships and collaborations for the entire campus.
Providing reference support in finding and citing data sets	This falls under the work of what is considered more traditional librarianship.	For digital scholarship, this is extension thereof—in both cases, oftentimes a researcher has created/gathered her own dataset or other digital research objects, but may need assistance supplementing their data, or may need help starting a project. Library staff can help locate data sources and/or help negotiate licensing terms.
Creating guides for finding data	Traditional librarianship	Traditional librarianship
Directly participating in research projects	Library staff can work alongside researchers as information specialists, research assistants, etc. Services can include metadata creation, data curation, research methods, acquiring research papers or datasets, etc.	Similar to data services—library staff can participate in many types of research projects, offering varying levels of expertise, from metadata creation at the beginning of a project, to curation at the end.
Discussing data services with others on campus	This is essential to creating and maintaining a robust research data community on campus. Networking, collaborations, wayfinding (pointing people to the right experts on campus)—all of these create opportunity for collaboration, consultation, and connection for campus researchers.	The same applies for digital scholarship, sometimes even more so since it is such an interdisciplinary methodology.
Training librarians and others on campus	Providing training to librarians on tools such as DMPTool, persistent identifiers, metadata creation, online research tools such as Open Science Framework, and possibly data analysis tools or methods.	Providing training on various software tools and how they might be appropriate for various projects.
Providing repository services	Library staff can either assist with deposit into local repositories, or help locate external repositories and provide assistance with remote deposit.	Although there are fewer repositories for digital humanities/scholarship work, the same consultation services would apply here to digital scholarship.

Adapted from Academic Libraries and Research Data Services: Current Practices and Plans for the Future. Prepared by Carol Tenopir, Ben Birch, and Suzie Allard. Chicago: Association of College and Research Libraries, 2012.

THE HISTORY OF DATA SERVICES AT UW

Staff have been working on data services issues in the UW Libraries since 2010, when the first UW data services librarian position was created to support campus use of research data. The position was created in response to growth in data-intensive research on campus. As with most campuses, at UW there are multiple centers or departments on campus that support researchers with data, including the Center for Social Science Computation and Research, the Center for Studies in Demography and Ecology, and the eScience Institute. Several departments have also historically been data-intensive, including the health sciences and social sciences, departments using geographic information systems (GIS) analysis and software, and physics/astronomy. As computer power and data availability have grown (along with the concept and analysis of "big data"), the Libraries have added a librarian dedicated to support and facilitate research data issues.

In addition to meeting with researchers from across campus and establishing the libraries as a "data concierge" for all campus users, one of the first big projects undertaken by the new data librarian was a campus-wide survey to assess the data needs of UW researchers. The survey and accompanying interviews were instrumental in developing UW's first set of data services, which included data management plan consultations, data acquisition, data-related referrals around campus, classroom teaching and workshops for library staff and campus researchers, and work with campus partners to create a community of practice around research data issues.

To further bolster the development of the data services program, a data services team was created, which included the geographic information systems (GIS) librarian, metadata librarian, staff from the Libraries' information technology and technical services, health sciences librarians, and the government documents librarian. The team participated in the activities listed above and helped disseminate information to their particular areas and from those areas to the data services team for discussion. Eventually, a second, part-time librarian was hired to support marketing and outreach for library data services across campus.

Over time, data-related work within the UW Libraries has evolved to include efforts and services that overlap strongly with digital scholarship. These include assessment work, staff training, and outreach and marketing services.

NEEDS ASSESSMENTS

Before deciding what services should be provided, libraries should perform an assessment to determine what digital scholarship work is currently underway, what needs the community has, and where any gaps might be. This type

of work should also include a frank assessment of library staffing and technological support, as well as identifying potential campus partners who can contribute to the effort.

The following are two examples of key assessments that have taken place within the data services team, or in collaboration with other UW Libraries units.

EXAMPLE 1

2012 Data Needs Assessment Survey[8]

In 2012, after several months of talking with library staff and campus departments, a more formal campus-wide survey was performed by the data services librarian to clarify the research data management needs of UW researchers. The key research questions for this assessment helped us to understand the level of awareness of data management plan requirements; determine the size and typical storage locations of researchers' datasets; and identify the attitudes of researchers toward data sharing, thus helping to shape our service priorities for UW researchers. Questions were reviewed by discipline to determine disciplinary need.

The survey targeted UW primary and co-primary investigators (PI and co-PI) who submitted grant proposals from 2010–2012 (see table 8.2). According to our records, the survey pool included over 3,100 individuals, of which 323 responded to the survey. When analyzing the data, we noted uneven response rates across academic disciplines, with a large number of respondents from the Health Sciences. This made it challenging to generalize results across disciplines.

To minimize the likelihood of confusion, at the start of the survey the data services team provided definitions for several data-related terms that were identified as ambiguous in pilot testing. These included:

TABLE 8.2

Respondents from 2012 UW Libraries' data needs assessment survey

	# Respondents	% of Respondents	% of Total Survey Population
Health	183	60.8	59.7
Sciences	82	27.2	28.5
Social Sciences	19	6.3	8.3
Humanities	5	1.7	.7
*Other**	12	4.0	2.8

*The category of "Other" included members affiliated w/ departments outside the four broader disciplinary areas (e.g. UW-IT, School of Business, Law School, Interdisciplinary Arts & Sciences, etc.).

Data: By data, we do not mean a synonym for information. We mean research data; that which is collected, observed, or created, for purposes of analyzing to produce original research results. Research data may be created in tabular, textual, statistical, numeric, geospatial, image, multimedia, or other formats.[9]

Data management: Data management pertains to the collection, cleaning, analysis, storage, sharing, disposal, and/or archiving of research data.[10]

Metadata: Metadata refers to the description of the content and context of data files. Examples include how, when, and by whom data was collected and how it is formatted.

Data cleaning: Data cleaning is a process used to detect and correct errors in a dataset. This can include the coding of missing data, correcting typing errors made in data entry, adjusting for column shift, and so forth.

In the main body survey, participants were asked to sort a number of services into low, low-medium, medium, medium-high, and high priorities. These levels were assigned values of 1 to 5, with a low value of 1 and a high of 5, and a mean priority level for each service was taken. According to results, the highest rated priorities were:

ensuring that data is secure—3.95

backing data up—3.95

short-term storage (5 years or less)—3.73

long-term storage—3.66

controlling/providing access to data—3.59

Other sections of the survey covered data storage needs and concepts around data sharing. For example, one section of the survey surfaced important trends in the duration of data storage needed by respondents. Only 9 percent indicated a need for storage for one to five years, while 27 percent were interested in "indefinite" storage. Additionally, 43 percent of the respondents were interested in storing their data for six to ten years, and 20 percent were seeking storage for longer than ten years but less than indefinitely. Separately, at least 68 percent of those responding wanted to be able to share their data over time (another 18 percent were unsure if they wanted shareable storage).

Respondents were also asked if they were willing to share data beyond the project team. A little more than half of respondents (55 percent) said that they were willing to do so. If health fields are excluded, this average rises to two-thirds (67 percent). Respondents in the health fields appeared less likely to share data with people outside of their research team; responses to the follow-up question asking for clarification indicated this was primarily due to concerns about the protection of anonymity and personal information of human subjects. Respondents in the sciences appeared to be most likely to

share their data; again, according to responses from the follow-up question, this may have been due in part to sharing requirements of various funding bodies. Researchers were most willing to share their final reports and least willing to share raw data.

Humanities respondents comprised only 1.7 percent of the total respondents to the survey. Nevertheless, by combining the information provided by both our report and a separate needs assessment by UW's digital scholarship librarian (see chapter 4), we were able to make connections between our units' target populations.

The data services team considered the survey and interview results in terms of what services the library was offering, what researchers said they'd like to have offered, and where there were gaps. In response, the team adjusted some of our outreach and created many more training opportunities for researchers of all levels (as well as librarians) on both tools and data management concepts. LibGuides were modified to reflect the types of information researchers were seeking, and we created a stronger relationship with our Office of Research as well as data-intensive departments on campus to make sure lines of communications were open in both directions. In response to the survey's findings that many researchers would share data more easily if there were a simple mechanism to do so, the library began investigating the possibility of building a UW data repository to accompany our institutional repository. Information on that project is detailed in chapter 6.

EXAMPLE 2

UW Libraries' Triennial Survey

The data services team also relies on the UW Libraries' Triennial Survey, a survey sent every three years to faculty and graduate students at all three campuses (Seattle, Bothell, and Tacoma), to all undergraduates at Bothell and Tacoma, and a sample of undergraduates at the Seattle campus. Results of the surveys are used for facilities or service improvements, library budget narratives, UW student technology fee proposals, and to demonstrate the library's impact on research, teaching, and learning.[11]

The most recent survey was completed in 2016. Respondents to the 2016 survey included 1,527 faculty (35 percent response rate) and 2,780 Seattle campus graduate and professional students (22 percent response rate).[12] The survey included several data-related questions.

Does your research and teaching involve any of the following digital activities/tools? Please check all that apply.

- ☐ My work does not involve any of these
- ☐ Text/data mining
- ☐ Data visualization (using tools such as Tableau)
- ☐ Web authoring or publishing (using tools such as Scalar or Omeka)

- ☐ Digital mapping/digital map making (using tools such as ArcGIS, Neatline, Google My Maps)
- ☐ Digital annotation (using tools such as hypothes.is and Lacuna)
- ☐ Other (please specify)

Responses indicated that 40 percent of both faculty and graduate student respondents said that their work uses one or more of the above activities and tools. Highest-ranked individual activities were data visualization and data/text mining, followed by digital mapping.

Another digital scholarship and data services question of interest that was included on the survey looked at services to support researchers' work:

Which of the following library services would be helpful to your academic work, research, and scholarship?

- ☐ Citation management software
- ☐ Consultation with a subject liaison librarian
- ☐ Help managing, archiving, and preserving research data
- ☐ Questions related to publishing issues
- ☐ Submitting theses and dissertations electronically

Surprisingly, 51 percent of graduate student respondents said that help on managing, archiving, and preserving research data would be helpful to them and 57 percent of faculty respondents said the same.

On the faculty side, survey respondents were asked:

Which of the following library services would be useful to your research and scholarly activities:

- ☐ Managing, archiving, and preserving research data
- ☐ Assessing the impact of your work or research
- ☐ Questions related to publishing issues
- ☐ Depositing your scholarship into ResearchWorks, UW's institutional repository
- ☐ Development of data management plans

44 percent of faculty said they would be interested in depositing their research into ResearchWorks, and 37 percent said that help developing data management plans would be useful (figure 8.1).

Results of the Triennial Survey are used to help library staff reassess current offerings, compare to what survey responses indicate is desirable, and to identify opportunities where there are service gaps. Data services staff used the 2016 survey results to develop an online research data management workshop offered via Canvas (described below), and to improve our educational materials, LibGuides, and workshops offered on tools such as DMPTool and Open Science Framework (OSF).

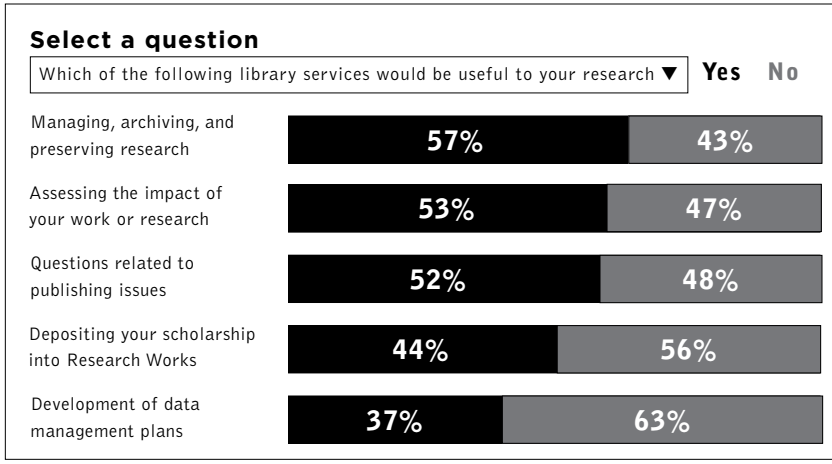

FIGURE 8.1
Faculty responses to "Which of the following library services would be helpful to your academic work, research and scholarship?"

KEEPING UP WITH NEEDS ASSESSMENTS

As with all technology, changes are rapid and constant. The solutions a library picks today to fit digital scholarship needs will change, whether that be in two years or ten. The tools researchers use now will likely change at the same pace; indeed, there are no doubt technological tools on the horizon that will allow types of research we have yet to imagine. To keep pace with these changes, libraries and librarians will need to make sure they stay on top of changes in the field. This can be done through partnerships with digital scholarship researchers and centers on campus, attending workshops, continuing education via software training and research methodologies, and by attending conferences. Professional organizations such as Research Data Access and Preservation and the International Association for Social Sciences Services and Technology are excellent places to find colleagues, continuing education, discussion groups, and conferences.[13] Keeping in touch with what other universities and data archives are doing in terms of facilities, programming, technology, and services will help library staff stay informed about changes in the field. Locally, designating a librarian or library staff member as a liaison to the various departments working on digital scholarship projects will help establish these relationships, as well as encourage their continuation.

Needs "re-assessments" will be essential to understanding an individual library's researchers and technological status. Although an extensive primary needs assessment is typical for a library new to digital scholarship, subsequent reassessments can be done on a smaller scale. For example, the library can

- interview a particular population of the researchers each year on a rotating schedule;
- create a community of practice (COP) that meets on a regular basis that includes library participation, which can be an important source of information; and
- make sure the COP knows the library is a resource and includes library staff in conversations on needs, projects, research, presentations, curriculum support, and so on to enhance and support this relationship.

This type of constant reassessment is essential when dealing with fast-paced changes in technology and methodology.

Once relationships are established that allow the library to understand the data and digital scholarship environment in which the University is working, the next step is keeping library staff and campus researchers aware and trained on services and tools available at the library.

COMMUNICATIONS, MARKETING, AND OUTREACH

Once you have decided on the kind of services that will be offered by your library, the next step is making sure the campus population is aware of those services. To get information and messages out, a communications and marketing plan can be an excellent tool to keep communications consistent, constant, and available via multiple channels.

The UW data services staff created the inaugural version of its own communications plan in 2014 (see appendix G). This plan included an outline for what types of communications would come from the unit, on what time line, on what channel (for example, Twitter, Facebook, Instagram, blog, etc.), and listed responsibilities by staff position. The intended audience for data-related communications was, ambitiously, anyone on campus—the idea being that parts of our services are relevant to undergraduates, research staff, faculty, graduate students, research assistants, and so on. It also included a calendar of when items are to be published, a crucial piece of any communications plan.

Locally, the data services staff also participates in groups for social media managers, both within the UW Libraries and the University as a whole, to hear about the trends, tools, and time lines they use. We follow other library social media accounts, as well as the accounts of data-intensive researchers or departments on campus, to learn about their events and to stay aware of various approaches to using social media when advertising services. Another, broader resource of interest to library staff is the annual Library Marketing and Communications Conference, which offers sessions and presentations on a variety of issues of interest to marketing and communications in a library setting.[14]

Many libraries already have certain communication channels in place: outreach to other library staff, particularly liaisons or subject librarians who work closely with departments; calendars of events; email newsletters; building signage; Twitter and other social media channels. However, sometimes these efforts are distributed among staff, and efforts may be duplicated, or individual items or channels may fall through the cracks. An organized communications plan is a good tool to prevent duplication and provide maximum impact for the library's message.

If your library has existing staff or guidelines for developing a plan, you'll want to start there. But even if there is an overall plan for your library, having an individual plan for your department will enhance your message and your presence on campus. Getting your own message out independently, or in addition to the library's overall message, will enhance patron's awareness of your services.

While creating the plan, make sure to consult with your library's social media and marketing staff (if any) to determine what channels are most effective. Twitter, for example, might be heavily used but perhaps only by a subset of your population. It may be that Instagram and SnapChat will have more uptake with your users. Connecting with your campus social media department or staff will help you find out what has the most impact and will help you keep up with trends, both by students as well as administrators.

You should determine the frequency that items will be posted to various forums (Twitter, Facebook, Instagram, blogs, etc.) and then create a calendar that anyone associated with marketing will adhere to. Put a point person in charge to remind writers when they're due to create content and to remind them when it's time to share information online.

As with all social media, be careful about operating in a vacuum: if you're not careful with your follows and followers, your consistent and useful messaging may only be reaching other library staff or campus users that are already aware of your services. Care must be taken to cultivate awareness of other groups, departments, and leaders on campus who are well-positioned to spread the library's message to their own constituencies. Having the library's message rebroadcast via retweets, email, newsletters, and so on by a respected professor, dean, or department head can have more impact than repeated messages from little-known accounts. Cultivating personal relationships will help improve the chances of this type of message rebroadcasting.

The communications plan is an excellent first step in creating a flow of information; however, personal relationships are what will give any marketing messages the weight and validity they need to impact user populations. One way to ensure that the library's message is distributed is to create and foster partnerships with various campus groups and departments. Subject and liaison librarians are a logical beginning, but outreach to other groups will work in the library's favor as well.

From the Culture of Digital Scholarship in Academic Libraries, edited by Robin Chin Roemer and Verletta Kern (Chicago: American Library Association, 2019).

At UW, the Libraries has a strong connection with both the eScience Institute, UW's home for data-intensive science (not coincidentally located in a former Libraries-owned space), and the Walter Chapin Simpson Center for the Humanities. These types of close partnerships, if bolstered by regular meetings and conversations, will enable a two-way chain of communication that will keep both parties aware of the other's activities. Librarians at UW work with individuals from both the institute and the center to collaborate on programming, consult on areas of expertise, and solicit feedback on activities and events in the library and the individual departments.

Essentially, the message is this: communications outreach must come from multiple outputs, multiple times, and you will still have to rally constantly to make sure your message is heard. Marketing and communications channels are being used not only to create awareness around issues relevant to data-intensive research but also to broadcast information about training and education offered by the library.

TRAINING

Libraries and their staffs are in excellent positions on campus from which to offer various types of trainings on the tools and services that support contemporary research. This is because libraries tend to be centrally located on campus, are discipline-agnostic, and their offerings are typically available to researchers of all levels. Libraries' trainings are also usually offered in various convenient formats: drop-in sessions within the building, collaborations with other campus departments, educational series, presentations, and online classes or workshops are only some examples of the education and training a library can provide.

CASE STUDY

Staff Training

The UW Libraries' data services team has taken the approach of first offering in-person data-related trainings to library staff, then turning to the campus community. These staff sessions are usually an hour long and deal with a single tool at a time, unless there are several tools that relate to one topic (e.g., UW has several tools that offer various persistent identifier capabilities). Each session typically provides a short introduction to the tool and what it is used for, followed by either hands-on time or a question-and-answer period (or both) at the end of the session. Workshop offerings have included training on tools such as DMPTool, Open Science Framework (OSF), ORCID, and DataCite, for example.[15] The idea behind these sessions is to provide introductory-level information to staff, who will then have enough information about a tool to

refer patrons to it when appropriate. Data services staff still handle reference-level questions about the tools, but the greater awareness about their capability allows more users to be directed to tools by word of mouth, instead of relying on Google searches or library reference guides.

The UW Libraries has also offered information sessions about products that are not supported by UW, but that could be acquired by individual interested parties. For example, data services staff helped electronic lab notebook company LabArchives offer a drop-in educational session to UW researchers, even though UW does not currently have a sitewide license for the software.[16] Having LabArchives demonstrate their software and be available for questions gave researchers an opportunity to determine whether or not this software would meet their needs. Although the UW Libraries isn't in a position to acquire the software on behalf of the campus, knowing about the software's abilities helps Libraries staff inform interested parties about free trials, grant-based allocations of the software, and helps the Libraries' staff advocate when necessary for other parties to acquire their own versions of the software.

Trainings can also be a worthwhile strategy for libraries looking to increase awareness of data-related issues beyond the uses of individual tools. Over the last few years, the UW Libraries has turned more of its attention toward developing this type of training, especially in combination with the Libraries' online learning efforts. One outcome has been the creation of a successful online multiday workshop on research data management skills, which we offer via the Canvas learning management system.

CASE STUDY

Research Data Management Skills Workshop

The Research Data Management Skills Canvas workshop was based on a 2013 UW workshop created from the New England Collaborative Data Management Curriculum (NECDMC).[17] The NECDMC was developed to teach research data management skills to researchers primarily in STEM disciplines. Each of the seven modules included a lecture, a slide presentation (sometimes with notes), activities to illustrate the learning concepts, extensive background information in the form of a Word document, and case studies.

In spring of 2013, librarians from UW's main campus and the health sciences library worked together to pilot all seven modules in a weekly, hour-long format. Staff worked to edit, enhance, and customize module content to reflect the culture, disciplinary realities, and technical tools available to UW's campus researchers. The resulting course was widely marketed, and had more than seventy-five enrollees, primarily from engineering, nursing, forestry, and biology. Unfortunately, the course suffered significant attrition over the seven weeks, with only seven of the seventy-eight enrolled researchers completing all modules. Although those who finished the class gave the content and instructors high ratings, the librarian instructors felt that the overall

impact of the course was minimal. Discussions followed among the data services staff about how the content could be reoriented and better delivered to more researchers in a more efficient way.

Luckily, a collaborative solution to this issue arose in the form of the UW Libraries' instructional design librarian, Robin Chin Roemer, who had already been piloting a new series of online Canvas courses to graduate students on how to acquire key research skills. Chin Roemer had attended several of the in-person NECDMC workshop sessions and suggested the content might do better if delivered in a multiday online format, similar to the asynchronous Canvas classes her team had been designing. The data services librarians agreed, and development of a new online workshop began.

Content for the new workshop was plentiful, as much of it could be drawn from previously customized NECDMC modules. That said, the most difficult task for the data services team was limiting the content to the most salient and helpful data management best practices, tips, and links to resources, both at UW and externally, that would provide the highest impact to attendees. Collaborating with the specialists within the UW Libraries' Instructional Design (LibID) unit was imperative to achieving this goal. Because they were not data experts, LibID staff were more easily able to focus on streamlining the content suggested by the data services team and formatting it for use in the online Canvas environment. The LibID team provided design assistance, utilized best practices in delivery and language, and helped the data services team members understand how best to "chunk up" the curriculum for maximum impact (see figures 8.2 through 8.4). Having this instructional design expertise at

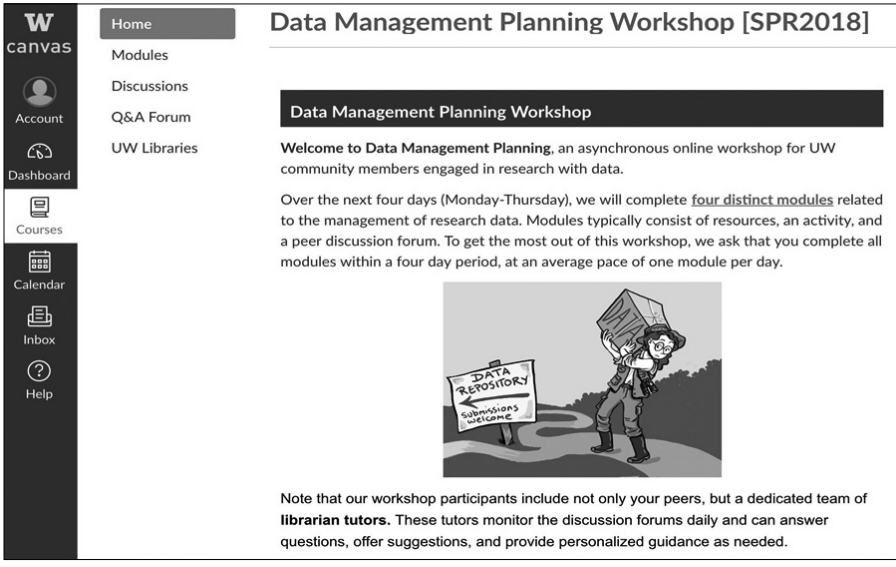

FIGURE 8.2
Homepage for Data Management Planning Workshop

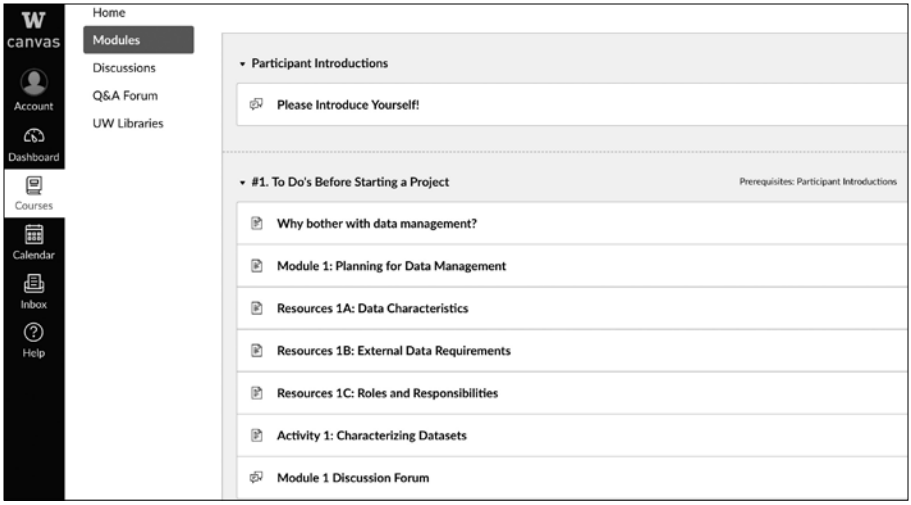

FIGURE 8.3
Example of Module 1's Content

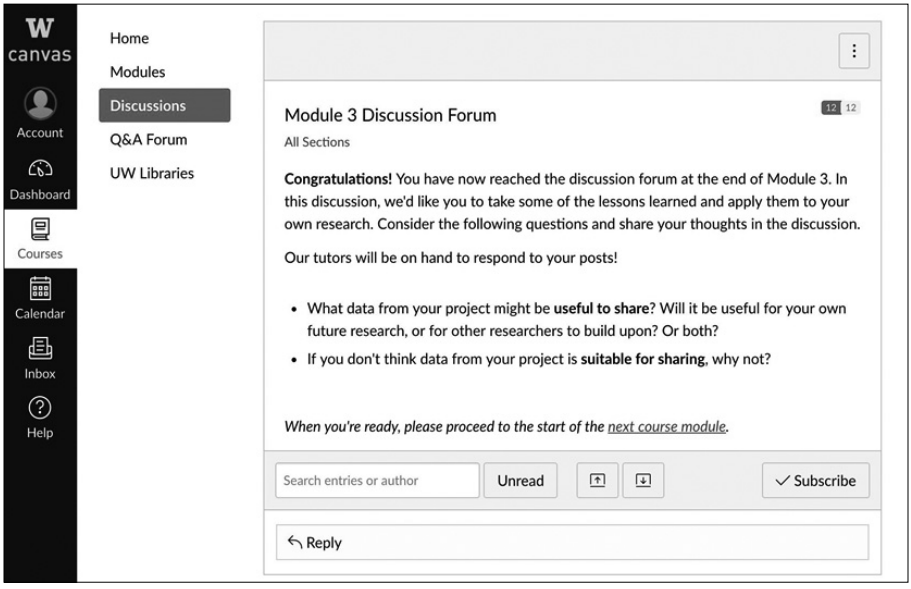

FIGURE 8.4
Module 3 Discussion Forum

hand made our revision work possible; without it, our course would have been ill-designed and overburdened with content.

The revamped Research Data Management workshop launched in August 2016. In consultation with LibID, we decided the class would be four days in duration, focus on asynchronous interactive learning between students and librarian "tutors," and would be limited in enrollment to forty students. Within days of its announcement, enrollment for this new non-credit class ran to sixty students, which forced us to create a significant waitlist. A 5:1 ratio of students to librarian tutors meant that every post was answered by a librarian in a timely fashion; in fact, because one of our tutors was working remotely from Germany, we had the benefit of the time difference to help make sure no query went too long without being answered. As with most online classes and massive open online courses (MOOCs), a minority of students did the majority of the participation, but that appears to be an issue with online, no-credit courses of all types.[18] Participants who completed the course rated the workshop highly and requested long-term access to the workshop materials in order to share information with peers.

The data services team followed up this class with a second offering in spring 2017, and then again in summer 2017. After looking at attendance numbers and seeing that a growing number of registrants were coming from health sciences, the data services librarian approached the librarians in the UW Health Sciences Library to ask if they'd be interested in creating their own workshop specific to their student populations. This quickly resulted in a new health sciences-focused version of the Research Data Management workshop, which was first offered concurrently with the sciences-focused workshop in spring 2018. For each instance of the workshop, maximum enrollment has been reached within a few hours or days of the initial marketing postings and waitlists are created each time. As of July 2018, a second iteration of the health sciences workshop was offered again during the summer term, concurrent with the sciences workshop. This concurrent two-workshop approach allows us to double the number of students and associated impact. It also allows librarian tutors for the two workshops to share information among themselves to better serve the students' needs.

The success of the health sciences-focused workshop has led to discussions among some UW librarians about offering other customized versions of the course to specific populations. These could easily include researchers who identify themselves with digital scholarship and digital humanities. Since a good portion of the workshop includes links to UW staff, IT support, and associated tools and resources—the portions of the workshop that don't require interpersonal connections to deliver the information—digital scholarship researchers would be well-served by a similar online workshop. Since many digital scholars are either unaware of or new to various relevant data management tools, a DS/DH class could provide up to 80+ researchers a year

with information relevant to their fields of study and to the UW in particular. A workshop like this could also be developed to help walk researchers through the creation of a data management plan, whether that plan be for the National Science Foundation or the National Endowment for the Humanities.

WHAT'S NEXT FOR DATA SERVICES AND DIGITAL SCHOLARSHIP

Many of the tools and services provided by a library's data services staff are things that help ensure the long-term preservation and access to research outputs. However, the infrastructure required for both digital scholarship curation and archiving, as well as data science curation and archiving, is not currently something the UW Libraries is able to provide. Although there are people on the UW campus with the technical expertise required to provide these types of services, in an academic environment with constantly shrinking budgets and the ever-present competition for funds, finding the funding to provide such a service on a continuing basis is an issue. Librarians are available to provide curation services, but that lands us right back in an issue of definition: what, exactly, are curation services? Will they include metadata services, data validation, future format migration, hosting, and backup, or some combination thereof? What budget will be used to pay for these services to ensure their longevity? Long-term storage costs money and requires staffing to maintain. Right now, these questions and their solutions remain up in the air.

There are two issues of paramount importance to data-intensive science as well as digital scholarship for which there are no clear-cut answers. As discussed in chapter 4, "the number one question received during Office Hours centers around where to host and store projects for free or at low cost." The sciences have an increasing need to both archive and provide access to interactive content and very large datasets and the associated software. Even if libraries are not able to provide these services for either data-intensive science or digital scholarship, they have expertise in long-term storage and access of many types of materials—for all disciplines. Being able to participate in discussions and develop services or referrals for these kinds of services will help ensure that best practices for archiving are being considered and will ensure the best long-term outcomes for preservation.

CONCLUSION

This chapter outlined the similarities and differences between data services and digital scholarship by describing the data services assessment work and identifying opportunities for collaboration through training and marketing

services at the UW Libraries. As noted above, there are many areas of library data services and digital scholarship that are similar, if not identical to, each other. Assessing its own library environment will help library staff develop high-demand services and will keep those services current. A staff knowledgeable about both data services and digital scholarship basics, trends, and technology will improve communication and community among library users. Using marketing strategies to get a library's message out will help ensure that library users know what the library can provide to assist them in their research. All of these efforts will foster a robust culture of digital scholarship and enhances researchers' experience—and success—creating and sharing their work.

Takeaways

- Data services and digital scholarship are both technology-driven fields with meanings that vary greatly from context to context. Establishing definitions and context local to a specific library or institution is essential to ensure that library staff and patrons understand what is being discussed.
- Assessment is important when introducing a service, but constant reassessment is necessary to deal with fast-paced changes in technology and methodology.
- Partnerships and personal relationships are essential, both within and external to the Libraries.
- Outreach must come from multiple outlets, be repeated multiple times, but you'll still have to work hard to ensure that the message is heard.

NOTES

1. Amanda J. Swygart-Hobaugh, "Data Services in Academic Libraries—What Strange Beast Is This?" *SLIS Student Research Journal* 6, no. 2 (2009): 1.
2. Association of Research Libraries, "Digital Scholarship Support in ARL Member Libraries: An Overview, https://perma.cc/ZVD2-LGM3.
3. Robin Rice, DISC-UK DataShare Project: Final Report, (JISC, 2009): 16, http://repository.jisc.ac.uk/336/.
4. Lynda Kellam and Kristi Thompson, *Databrarianship: The Academic Data Librarian in Theory and Practice* (Chicago: ALA Editions, 2009), 4-5.
5. Report internal to UW.
6. Heather McCullough, "Developing Digital Scholarship Services on a Shoestring: Facilities, Events, Tools, and Projects," *College and Research Libraries News* 75, no. 4 (2014): 187.
7. Carol Tenopir et al., *Academic Libraries and Research Data Services: Current Practices and Plans for the Future,* (Chicago: Association of College & Research Libraries, 2012).

8. Stephanie Wright et. al., *Fall 2012 Research Data Management Needs Assessment Results* (Seattle: University of Washington Libraries, 2013).
9. Adapted from Rice, DISC-UK DataShare Project, p. 16.
10. Adapted from University of North Carolina, Research Data Stewardship Report, https://sils.unc.edu/sites/default/files/general/research/UNC_Research_Data_Stewardship_Report.pdf, 2012, p. 11.
11. Funds from UW's student technology fee are used to pay for various technology resources for student use.
12. University of Washington Libraries Assessment Program, "Triennial Survey," www.lib.washington.edu/assessment/surveys/triennial.
13. More information on the International Association for Social Sciences Services and Technology may be found at www.iassistdata.org/. More information on Research Data Access and Preservation may be found at https://www.asist.org/rdap/.
14. More information on the Library Marketing and Communications Conference may be found at www.librarymarketingconference.org/.
15. DMPTool (https://dmptool.org/) is a tool that simplifies the creation of data management plans for many funding agencies. Open Science Framework (https://osf.io/) is a cloud-based collaborative research platform. ORCID (https://orcid.org/) is a provider of persistent identifiers for researchers. DataCite (https://www.datacite.org/) is a provider of digital object identifiers for research data.
16. LabArchives, (https://www.labarchives.com/), is an electronic lab notebook tool.
17. "New England Collaborative Data Management Curriculum," accessed October 3, 2018, https://library.umassmed.edu/resources/necdmc/index.
18. Amy Ahern, "The Flip Side of Abysmal MOOC Completion Rates? Discovering the Most Tenacious Learners," EdSurge, February 22, 2017, https://www.edsurge.com/news/2017-02-22-the-flip-side-of-abysmal-mooc-completion-rates-discovering-the-most-tenacious-learners.