Title: Situating Expertise in Practice: Domain-Based Data Management Training for Liaison Librarians

Jamie Wittenberg (corresponding)[a], Anna Sackmann[b], Rick Jaffe[c]

[a]Indiana University Libraries, Indiana University Bloomington. 1320 East Tenth Street, Bloomington, Indiana, 47405 United States. <jvwitten@iu.edu>

[b] Kresge Engineering Library, 110 Bechtel Engineering Center. University of California, Berkeley, 94720 United States. <asackmann@berkeley.edu>

[c] Research IT, Earl Warren Hall, 2195 Hearst Ave, Berkeley, CA 94720 United States. <rjaffe@berkeley.edu>

Abstract:

The research data management team at the University of California, Berkeley implemented a domain-based Librarian Training Program in order to upskill liaison librarians in research data management principles and create a community of practice among librarians providing research data support. The training program partnered with representatives from each subject division of the Library to integrate content from relevant disciplines. The training model emphasized scaffolding and concrete deliverables, teaching specific tools and concepts, and creating learning objects useful for instruction and outreach. Employing a situated, learning-based, pedagogical model, the program was more successful than previous attempts at library-wide research data management training at Berkeley. This analysis details the program management, curricular design, instruction, and outcomes that made the Library Training Program successful.

Introduction

Librarians in liaison and domain-centric roles often have opportunities to connect researchers with important information but are not always well-positioned to provide the consultation services necessary for emerging topics such as research data management (RDM). Given funders' increasing requirements for data management plans, data sharing, and reproducible research, librarians recognize a growing need to improve awareness and advocacy for RDM (Johnston et al., 2017; Latham, 2017; Carlson & Stowell-Bracke, 2013; Antell et al., 2014). At the University of California, Berkeley, an eighteen-month initiative to train subject librarians in research data management addressed the discrepancy between job requirements and librarian skill sets. The Librarian Training Team designed a domain-specific curriculum and outreach program to prepare librarians for the provision of research data management consultation and referral support. This training program grew from early Berkeley data management efforts to offer generalized training to early adopters selected from each domain. However, general research data management training was unpopular and insufficient to prepare early adopters to train their departments (Wittenberg & Elings, 2017). Domain-based research data management training, in conjunction with administrative buy-in, is more effective in engaging librarians, delivering relevant content, and creating a community of practice.

Literature Review

The Role of Librarians in Support of Research Data Management

In the library literature, a number of studies report on the importance of academic librarians providing data management support services and examine the methods in which librarians can best prepare to fill this new role. In March 2007, the NSF published a report suggesting that "university-based research libraries and research librarians are positioned to make significant contributions" in developing support for data curation, analysis, archiving, and the creation of digital libraries that index research outputs (Cyber Infrastructure Council, 2007). The conversations addressing librarians' role in data support continued in November 2008, during which attendees of the Research Data Management (RDM) Forum identified four key data roles for librarians called into the data support role: data manager, data creator, data librarian, and data scientist (Pryor & Donnelly, 2009). The data librarian's core skills include (among others) knowledge in data preservation, data appraisal and retention, and standards development.

Following the NSF report and the RDM forum, others examined how library organizations can and should adjust to this call for greater support in research data services. Delserone conducted a science assessment at the University of Minnesota (UM) and asked researchers in the sciences about the types of help they seek from the library. Their responses fit into three categories: data organization and manipulation; data storage, security, and sharing; and data stewardship. UM developed the Research Cyberinfrastructure Alliance (RCA) to further examine how the role of libraries and librarians could best support researchers in these three categories. The RCA recommended that the library provide support in multiple areas of data services, including: data stewardship, instruction, data management policy, and data repository certification (Delserone, 2008). Jaguszewski and Williams discovered a similar trajectory of library services when they interviewed administrators at five libraries of the Association of Research Libraries (ARL). Libraries are shifting their focus to "what users do (research, teaching, and learning) rather than on what librarians do (collections, reference, library instruction)" (Jaguszewski & Williams, 2013, p. 4). This shift in focus surfaced a number of areas that librarians needed to support, like data management and preservation plans. Librarians responded to these recommendations through a variety of training programs designed to upskill librarians in research data management topics.

Types of Trainings for Librarians

In the years following the Jaguszewski and Williams study, librarians have worked to find the proper mode of Research Data Management training through a combination of online, hybrid, and in-person trainings. Research Data MANTRA (Management Training), developed at the University of Edinburgh in 2011, was created as an online curriculum to teach researchers and librarians about research data management through use cases, scenarios, and best practices (Rice, 2014). This distance learning model was designed to support individual learners located remotely from the instructors and from one another. Initially, the creators of MANTRA were funded to create discipline specific materials. However, a needs assessment found that specific topics, such as confidential data, spanned multiple disciplines. Currently in its fourth iteration, the program is now linked to a certificate-granting Coursera MOOC designed for both librarians and researchers (Tibbo & Jones, 2015). MANTRA enables learners to work through data management training at a self-guided pace, and its open license has allowed others to build upon the training for their own local purposes. The training remains broad and, although highly accessible, does not provide the customized training incorporating local research culture that some librarians require.

Shortly after the development of MANTRA, a hybrid training model emerged that specifically focused on librarians (De Smaele et al., 2013). Librarians from the three Dutch Universities of Technology (3TU) developed Data Intelligence 4 Librarians, a mix of online and in-person intensive instruction. The model, which combined group meetings, online study, and homework assignments, reflected the collaborative environment of the modern scientific community. The course consisted of seven days of training (four days of face-to-face training) and covered 4 interdisciplinary modules: data management; technical skills; acquisition and advice; and actual topics (de Smaele et al., 2013). Participants in Data Intelligence 4 Librarians responded positively to homework and the resulting discussion; however, participants wanted real-world use cases contributed by librarians employed in an RDM role and researchers addressing how data are managed and how these behaviors might differentiate based on discipline. Similar responses were found when planning for and assessing other general training programs for librarians (Bresnahan & Johnson, 2013; Cox et al., 2014; Wittenberg & Elings, 2017).

In-person trainings have taken a number of forms. RDMRose, which originated in the UK, took a slightly different approach in their training for liaison librarians. The course content consisted of half-day sessions completed as self-directed learning exercises. This format enabled hands-on activities, such as researcher interviews. Participants found that the group discussions enabled them to see multiple perspectives of RDM; however, librarians desired greater inclusion of local institutional context, which may aid in translating the theoretical to concrete examples (Cox et al., 2014). Byatt et al., also held an intensive, in-person workshop around data management informed by survey results. This workshop for librarians focused on knowledge of research data management, making effective referrals, and end of lifecycle research data management (2013). The training combined a general introduction which was intentionally kept interdisciplinary in order to allow best practices to be transferred between disciplines. Future changes for Byatt, et al., include widening the scope of trainees to draw in a diverse set of expertise. Both of these in-person trainings highlight the difficult balance of providing discipline and institutional specific RDM knowledge from multiple areas with greater diversity of opinion.

Three additional in-person trainings have taken on a discipline-based and researcher-centric approach by integrating data workflows into the learning process. Bresnahan and Johnson conducted a needs assessment of librarians at University of Colorado Boulder supporting research data management and

found that librarians desired practical, hands-on training and expressed concern regarding the disciplinary differences required to work with researchers across the university (2013). As a result of this needs assessment, a day long workshop called *DataDay!* was developed and implemented for subject librarians at the university. The workshop included hands-on exercises and discussions through which participants worked with real datasets (Johnson & Bresnahan, 2015). Lyon created a 3-unit Research Data Services course at the University of Pittsburgh iSchool in which practitioners participated alongside graduate students. This exposed MLIS students to the practitioner perspective while providing practitioners with graduate level coursework and content. The course facilitated collaboration with faculty and researchers in four health and physical science laboratories to give participants a sense of daily research workflows (Lyon, 2016). Finally, librarians at the University of Pittsburgh adapted pieces of Lyon's work to develop a new research data management training for subject librarians that enabled a deep dive into subject areas while addressing research data competencies like data sources, metadata schemas, and data archives (Mattern et al., 2016). While this deep dive was successful in developing disciplinary based knowledge of RDM, the authors found a need for greater peer-feedback among librarians in order to develop a stronger community of practice.

Situated Learning and Communities of Practice

The literature shows that, when faced with the challenge of learning the skills and expertise needed to support a burgeoning new field, librarians have requested training that conveys real-world experience, prepares them to recognize disciplinary variation, and equips them with an understanding of both the broad perspectives and the local institutional context. However, the success of their efforts is equally dependent on the process by which they develop these new capabilities.

Lave and Wenger's construction of situated learning places value on learning through a process of participation with communities of practices (1991). Situated learning increases the effectiveness of learning new abilities and tools within a specific academic culture and discipline (Farrell & Badke, 2015). The concept of situated learning has been applied to information literacy as a way of knowing an information landscape within a specific context (Lloyd, 2007). The concept of communities of practice has evolved from Lave and Wenger's initial definition of 'a system of relationships between people, activities, and the world; developing with time, and in relation to other tangential and overlapping communities of practice' (1991) to a nuanced, deliberate group consisting of three elements: domain (a specific area of expertise that members share), community (a set of people who engage with one another), and practice (ways of dealing with problems typical of a domain) (De Cagna, 2001).

The RDM Librarian Training Team mixed general concepts with discipline-based research data workflows to create a domain-based model for librarian training. Each session was tailored to the priorities, needs and culture of a given division of the university's library. In this way, the librarians developed a community of practice with the sociocultural backing of their division leadership. The format of the division-based librarian training program allowed librarians to fully engage in the norms and workflows of the researchers they serve and of the librarians with whom they practice, contributing to the division's capacity for mastering new skills.

History of RDM Training at UC Berkeley

Berkeley established its research data management program in 2015 in response to a benchmarking report that highlighted research data services as an area in need of improvement. The University Library and the campus's central Research IT unit partnered to develop a discipline-agnostic program (though one that was aware of the differences across disciplines) that leveraged existing services in both organizations, along with services offered by other campus and University of California partners. The establishment of a research data management program required internal and external training to expand the campus pool of librarians and IT staff capable of providing RDM support, foster collaboration with groups such as the UC-wide California Digital Library, the Berkeley Institute for Data Science, and the Social Sciences-based D-Lab, and communicate service offerings to patrons.

Initial training efforts in 2015 included standalone workshops that combined introductory research data management content with descriptions of services. These presentations, open to all librarians and IT support staff, generated an enthusiastic response but were unsuccessful in spreading expertise or establishing an ongoing community of practice (Wittenberg & Elings, 2017). Even with the inclusion of the DataDay! curriculum, assessment data from the training indicated a need for nuanced, concrete training, that they could apply to their work with researchers. When asked "After this workshop, how clear are the goals of the RDM Program?" the median ranking for clarity was a 2.5 out of 5, with most participants reporting that the goals were moderately clear. Participants reported in a post-workshop survey that in the future they would like to see skills building for RDM, discussions of specific services, and information about communication with PIs. To better address the needs that arose from the first workshop, the RDM program adopted a cohort model and developed instructional content for librarians and IT staff selected for the diversity of their domain expertise and their potential to become early adopters of RDM practices. These sessions were moderately successful and the RDM cohort reported improved comfort with RDM concepts. Following the sessions, the eleven librarians who participated in the workshops requested that future trainings accurately reflect the workflow and everyday activities pertinent to their discipline (Wittenberg & Elings, 2017).

In 2016, the RDM Program developed the domain-based training model that underlies the program discussed here. This new model addressed research data management principles within the context of the liaison librarian's subject area, using real-world case studies relevant to everyday practices. It could be, and was, tailored to the five specific library divisions at UC Berkeley (Social Sciences, Life and Health Sciences, Engineering and Physical Sciences, Arts and Humanities, and the Bancroft Special Collections Library) and pulled together components of data science, data management, and theory to deliver customized, yet relevant and measurable content to all subject librarians. The goal of the domain-based training model was to train every liaison librarian according to their disciplinary expertise and practical needs.

Instructional Design and Training Development

Each library division was assigned a two-month training cycle, or sprint, during which the RDM Librarian Training Team was committed to a training program exclusively for the librarians in that division. The model training cycle consisted of 3-5 sessions, a total of four hours instructional time. Prior to each division's cycle, the team would suggest potential learning outcomes for the division in a planning document. This document would be sent to the division head along with a request for a meeting to discuss training. During this meeting, division heads were asked to identify specific learning outcomes and goals for the two-month training cycle and to assign a liaison librarian in their division to serve as a representative and collaborate in designing, organizing and presenting the division's curriculum.

Developing domain-based training sessions required buy-in from library division heads, liaisons, and instructors as well as curriculum development and assessment. Additionally, the Librarian Training initiative was one of many competing organizational priorities within the library and presented new workload and professional development demands on librarians. The RDM Librarian Training Team required strategies to smooth its reception and promote its adoption.

By meeting with division heads and utilizing a division representative, the team utilized a "top-down, bottom-up" method to gain buy-in. This planning and implementation method, which is applied to leadership and project planning in fields such as management and higher education, established roles and responsibilities and sought to deliver a service from the perspective of both management and librarians. In higher education, this principle draws on the model of distributed leadership, wherein leadership incorporates the activities, viewpoints, and an interdepence from multiple individuals (Spillane et al., 2004). Distributed leadership is closely related to the convergence model in higher education, which joins the top-down efforts of leaders with authority with bottom-up efforts (Kezar, 2012). While convergence may result in different types of institutional change depending on the context and implementation, the model greatly contributed to the success of the domain-based training. The design of the training closely follows key strategies for convergence, including: assessment of timing from the bottom-up; alignment with priorities of higher-up leadership from the library strategic plan (includes areas of growth in data support) (California, 2017); and the utilization of a translator (division representative) who framed the presentation of curriculum and created a channel of communication to leaders at the top (Kezar, 2012).

The training team was systematic in training development and established a ten-step process for instructional design for each workshop cycle:

1. Identify and determine division learning outcomes with division head(s)
2. Plan content delivery dates and session locations with division representative and guidance from the division head(s)
3. Develop communication plan for the training with the division representative
4. Contact guest speakers (where applicable)
5. Designate responsibility for curriculum and develop assessment questions
6. Rehearse and review (return to step 5)
7. Deliver workshop
8. Review the workshop and analyze feedback with division representative
9. Begin planning for the next workshop (repeat 5-8)
10. Review training cycle with division head(s)

The first meeting with the division representative included determining logistics and creating learning outcomes for each training session. Discussion centered on the focus, format and timing of the sessions. Together, the team and the division representative crafted a survey of division librarians to gauge areas of interest, preferred format and available dates. The division representative was given responsibility for communications to other librarians in the division in order to increase the likelihood of response. The librarians' responses shaped the overall design of the cycle. Subsequent weekly meetings with the representative were used for curriculum design, revision and rehearsal.

Delivering Relevant Content

In order to build a baseline capacity library-wide, the training team developed three learning outcomes that spanned all division trainings:

1. Librarians and library staff will be able to recognize data management questions and confidently respond.
2. Librarians and library staff will actively build and participate in a research data management community of practice.
3. Librarians and library staff will understand the data management needs of researchers in their domain and incorporate that awareness into reference and instruction work.

These three common learning outcomes served as a starting point for the division-specific curriculum and provided a framework on which to build activities and more specific outcomes depending on skill level, need, and interest. Working with the division representative on the individual session learning outcomes, curriculum, and design was a crucial component in providing content that was appropriate for the librarians' daily workflow.

The division representative helped the training team determine if the level of content appropriately matched the services provided by libraries in a given division. For example, the Arts and Humanities and the Bancroft divisions required introductory materials in order to generate conversation around the definition and general practices of data within their domain. The Life and Health Sciences and Engineering and Physical Sciences requested a deep-dive into research data management topics and tools in their given domain.

The Social Sciences division participated in the first training cycle. The division comprised 14 librarians, whose knowledge in RDM ranged from beginner to expert. Because of this wide range of expertise, the team set general workshop outcomes, including: librarians becoming familiar with the most frequently asked questions regarding research data management within their domain; being able to incorporate data management questions into the reference interview (when appropriate); and learning how to make appropriate referrals. The training cycle took place over the course of two months and consisted of four, one-hour sessions: two workshops and two topic driven discussion sessions. During the social science training cycle, the training team identified where librarians' work with researchers intersects with data management; led a discussion on ICPSR as a model repository for the social sciences; created online learning objects; and finished the cycle with a discussion on data literacy.

The second training cycle combined the Life and Health Sciences Division and the Engineering and Physical Sciences division. Both divisions are smaller in size and frequently collaborate on instruction and collection development, making this an ideal combination. Based on feedback from the Social Science librarians, the trainers repeated the 'entry points into RDM' activity with the science librarians. Whereas the Social Science librarians had found the exercise helpful and used the entry point examples to generate conversation, the science librarians found the entry points to be unrealistic due to the way in which scientists generally communicate with librarians. Aside from the entry points activity, a new training format with new content was developed for this cycle. Instead of spreading the training out over two months, the librarians requested a "bootcamp" style training and schedule three, 90-minute sessions during a two-week duration. Librarians in the combined sciences divisions requested workshops on specific topics and tools frequently utilized by researchers within their domains, such as versioning with git and GitHub and network analysis and visualization. By focusing on these two concepts and tools as they fall within the research data lifecycle, the librarians developed a sense of researcher workflows

and research methods that they could apply to their consultation services. The team developed tools-focused learning outcomes that included: librarians understanding the practice of versioning and the process behind using git and GitHub as a version control system. The training cycle ended with a discussion on reproducibility in the sciences in order to generate conversation surrounding the different types of reproducibility, gray areas for researchers, and how to best adhere to funder and publisher data sharing requirements despite these challenges.

The third and final training cycle targeted the librarians in the Arts and Humanities division and archivists at the Bancroft Library, Berkeley's special collections. This training presented a different challenge in translating RDM methodologies to make them relevant to librarians and archivists who do not necessarily think of their work as data-centric. The trainers developed a curriculum of practical sessions underlined by RDM methodologies that focused on real workflows from researchers and librarians on campus. This cycle consisted of five workshops, and although the target audience consisted of librarians in the arts and humanities and campus archivists, all librarians and researchers in the digital humanities were invited to attend the series. By taking an approach that explored components in the research data lifecycle through humanist use cases, librarians and researchers could map abstract concepts to recognizable problems and tools. After beginning with an introductory session to establish language and the research data lifecycle, trainers presented sessions on active data storage tools (Box and Drive); Adobe Bridge and Shared Shelf for digital asset management; the DMPTool for writing NEH grants, and the Humanities Commons data repository. The learning outcomes for this cycle included understanding RDM concepts, recognize RDM concepts in their daily work, and gain confidence in their ability to provide effective reference around RDM-related topics.

The training cycle for each division included one common element: a role-playing activity to identify "entry points" into research data management. These are typical reference interactions in which a student or researcher may ask about a piece of software or literature review, and how a librarian can use that as an opportunity to generate conversation around data management. The purpose of the exercise was to demonstrate the context for RDM in the everyday work of the librarians. It was also an effort to bring all the participants to a common starting point, defining the phrase "research data management" through concrete examples. In practice, the role-playing activity -- a workshop leader and one of the attending librarians were asked to read from a script – served as ice-breakers and promoted discussion within the group.

Nurturing Communities of Practice and Leveraging Expertise More Broadly

Beyond instilling familiarity among librarians with the day-to-day work in their division, trainings sought to build comfort and shared expertise across the wider group of librarians and other research support staff. The Social Sciences librarians, who work at different libraries located across the Berkeley campus, requested that two of their four sessions remain closed from the rest of the library community in order to utilize that time to build a stronger community of practice within the division. However, they wanted half of their sessions to be an open invitation to other campus librarians in order to learn more about data management practices in other domains. The librarians in the Sciences division requested tool-based sessions offered only to members of the division. However, two librarians from Lawrence Berkeley Lab were welcomed into their sessions in hopes of fostering new relationships. The Arts, Humanities and Bancroft training series, under the guidance of the Literatures and Digital Humanities Librarian representing the division, was opened to all who wanted to attend with the purpose of generating conversation and building a community of practice around data and digital humanities

outside of the library. One session that coincided with "Love Your Data Week" was advertised as a campus-wide event.

Assessing Librarian Engagement

Assessment played a crucial role in the implementation of the training. The training team assessed several components in each cycle, including: attendance; specific learning outcomes, self-reported lessons learned; topics that librarians were still unsure of; overall satisfaction of individual sessions; and overall satisfaction of the full domain cycle. Additionally, the team asked how the program could be improved to inform future librarian training efforts at UC-Berkeley. For each session, participants registered their attendance and completed assessment forms. In some cases, pretest and posttest assessment was performed. After all sessions were held, the training team analyzed feedback and reviewed the training cycle during a follow up meeting with division heads. This follow up meeting included an overview of participant feedback, session participation, and suggestions for continuing RDM education.

After completing each training cycle, the training team examined recorded attendance in order to measure the effectiveness of the distributed leadership model. The average attendance for each cycle's group workshops was well over 50% of the division's librarians. The combined Life and Physical Sciences cycle had an average attendance of 93%; the Arts, Humanities, and the Bancroft average attendance was 73%; and the average attendance of the Social Sciences was 69%. Total number of librarians invited to participate was generated from the division rosters. For workshops and cycles that were opened up to the greater academic community for events like "Love Your Data Week," we noted a greater number of participants than that of the division roster (Fig. 1.).

Cycle Cohort	Total number of librarians invited to participate	Average attendance (number of individuals)	Range	Average attendance (percent of those invited)
Arts, Humanities, and the Bancroft	33	24	8-35	73%
Life and Physical Sciences	14	13	11-14	93%
Social Sciences	16	11	4-15	69%

Fig. 1: Attendance based on sign-in sheets for cycle cohorts

Due to the varying needs of the individual cycles, the team assessed different components for each workshop and cohort. While this approach informed changes and improvements for future workshops and cohorts, it failed to provide meaningful assessment data that could be used for benchmarking purposes. Respondents were consistently asked about overall satisfaction regarding individual workshops and the cycle cohort in which they participated on a five-point Likert Scale ranging from unsatisfied to very satisfied. The training team focused heavily on free text assessment questions to generate improvement and change as new workshops and cycles were planned. While asking about overall satisfaction regarding individual sessions and the entire cycle provided a brief snapshot to the team, overall satisfaction scales typically skew more positive than self-reported satisfaction and if the

question were to be framed in negative terms (Peterson, 2017). Because three of the five workshops delivered to the Arts, Humanities, and Bancroft cycle cohort were advertised beyond the divisions, the "overall satisfaction" question for the entire cycle was not asked (Fig. 2).

Cycle Cohort	Average Level of Satisfaction per session (scale of 1-5)	Overall Level of Satisfaction (scale of 1-5)
Arts, Humanities, and the Bancroft	4.33	N/A
Life and Physical Sciences	4.32	4.38
Social Sciences	4.56	4.75

Fig. 2 Participant satisfaction per session and satisfaction overall

Compared to assessment data from the RDM Program's first workshop – a general data management workshop held prior to the development of the domain-specific model - participants were more satisfied overall with domain-specific instruction, rating the first workshop and average of 4.1 out of 5. Domain-specific session satisfaction ranged from 4.38 - 4.75 out of 5.

Following the introduction sessions in the Social Science and Arts, Humanities, and Bancroft training, the team asked attendees to record something new they learned. Coded responses included the following: personal use of RDM; referrals and RDM support on campus; data and RDM as concepts; specific tools; and fielding RDM reference questions. The heads for Arts & Humanities, the Bancroft, and the Social Sciences each indicated a wide range of skill level for their divisions, which indicated a need for session to introduce concepts like the research data lifecycle and the variety of ways to define the term "data" based on discipline. For example, the librarians who work with data on a daily or weekly basis (i.e. Data Librarian and the Literatures and Digital Humanities Librarian) have higher levels of domain knowledge with data and data management. However, other librarians' skills may range from sufficient to beginner. Because of this wide range, the training team developed introduction sections for their training cycles. (Fig. 3).

Coded Response	Arts, Humanities, and the Bancroft Cohort (21 collected responses)	Social Sciences Cohort (13 collected responses)	Total Responses (34)	% of total responses
Personal use of RDM	1	1	2	5.80%
Referrals and RDM support on campus	7	3	10	29.40%
Data and RDM as concepts	8	5	13	38.20%
Specific tools	1	2	3	8.80%
Fielding RDM reference questions	4	2	6	17.60%

Fig. 3 Categories of something new attendees learned.

The same question was not asked in the Life and Physical Sciences cohort because the division heads requested that the training not include an introductory session but focus instead on tools and researcher workflows around concepts like reproducibility and versioning.

Because each training cycle consisted of highly customized sessions, the general learning outcomes developed for the entire training program were difficult to assess. While the team did not develop a definitive confidence marker during the assessment process, free-text feedback may have provided some indication of an increase in confidence following a session. Librarians across all cycles who attended tools-based sessions reported that not only would they be able to explain the given tool (i.e. Box, Drive, GitHub, Adobe Bridge, or Shared Shelf) to a faculty member, researcher, or student, but they realized how the tool could be incorporated into their own workflow. When asked to report something new they learned, the team received comments regarding a librarian's own ability to think of reference questions from the perspective of data management.

Next Steps

Librarians in the Social Sciences and Life and Physical Sciences cohorts who commented on improvements for the training as a whole provided excellent feedback that was later echoed in follow-up conversations with division heads. During the training cohort cycles, the team had not formulated a plan for opportunities to continue education in research data management. Several attendees and all of the division heads commented that they would like to see the training program continue in some form at regular intervals to enable librarians to continue learning and exploring. Other librarians commented on the need for more information resources and hands-on opportunities for learning. One librarian suggested that the training could be more effective by hearing directly from researchers to develop a better sense of their workflows and approaches to data management.

Academic liaison librarians offer a domain expertise in data identification, selection, organization, preservation, and access which should be leveraged by the research data management consultants in order to provide a comprehensive service (Johnston et al., 2017). As best practice, Research Data Management consultants at UC Berkeley now regularly communicate with liaison librarians to keep them abreast of consultation and training requests. When RDM receives a request for a consultation or group training, they alert the library liaison who provides support for the researcher's department. This current level of communication provides the liaison with an opportunity for direct involvement in the consultation or training, or with an opportunity to observe, build their practice, and continue to foster the relationship with researchers in their department. Pulling liaisons into the RDM program consulting network is critical for the program's ability to develop relationships with departments.

Liaisons have the knowledge and ability to identify connections with the research lifecycle and data, which compliments the RDM consultants whose expertise lies in IT. For example, five librarians from the social sciences division have advised on consultations in their departments regarding secure data workflow within a research group, digital asset management, and data management planning. The Demography and Sociology Librarian and the Government Documents Librarian assisted in the development and delivery of a data management training for the Demography Department. The Chemical Information Librarian frequently co-consults on data management plans with researchers in the engineering and physical sciences, which provides her with an opportunity to weave in her expertise on metadata schemas and the use of electronic lab notebooks to improve data management.

In September 2017, Berkeley hired a Research Data Management Program Manager, who has been tasked with strategically scaling the RDM Program. The Program Manager and UC Berkeley's Data Librarian developed the Data Initiatives Expertise Group, which will consist of librarians from a variety of disciplines who are currently interested in developing greater expertise around research data. The group will focus on building a community of practice to keep librarians engaged with campus data initiatives and will provide ongoing educational data programming to support librarians' professional development. Berkeley Library division heads have expressed a clear desire that on-going training is needed for liaison librarians in the field of data management. Building on the domain-based training model is a goal for the RDM program as it continues to provide consultation and training services.

Conclusion

The Librarian Training Program proved to be highly successful in the UC Berkeley Library environment. Librarians in all domains discussed their ability to incorporate RDM concepts into their workflow and utilize tools learned in the sessions. Participants, on average, were more satisfied with domain-based RDM training than they were with general RDM training. Division heads indicated an overall high level of satisfaction and a desire for continued training. Research Data Management has since identified liaison librarians from the training sessions who are willing and interested in taking part in data management consultations and trainings taking place with researchers in the departments they support. As the Research Data Management Program at Berkeley continues to grow, additional librarians will be identified and trained to assist with data support.

A variation of this training model has been adopted at the Indiana University Libraries to train liaison librarians on research data services and the implementation of an open access policy. In Indiana University's adaption of the model, the training team has emphasized an outreach agenda in lieu of an instructional agenda. Rather than convening for presentations and workshops to learn tools, concepts, and create learning objects, librarians collaborate with the Scholarly Communication department to develop subject-specific outreach material to educate their constituents about open access and data management services in the Libraries. Indiana University's adaptation of UC Berkeley's training program demonstrates the flexibility of the model as a tool to meet other disciplinary training goals in the library.

The program's initial and ongoing success is dependent upon intensive group work that establishes and strengthens situated learning in a community of practice. The program is effective in fulfilling short-term, concrete needs for disciplinary learning objects, but more importantly, it provides a framework for scaffolding data management pedagogy and facilitating new relationships across groups of stakeholders at the university. Tying new communities of practice to operational components of the research data management service (through ongoing training, workshops, and collaborative consultations) was critical to cementing the role of liaison librarians in research data support at UC Berkeley.

References

Antell, K., Foote, J., & Turner, J. (2014). Dealing with data: Science librarians' participation in data

management at Association of Research Libraries institutions. *College & Research Libraries*, *75*(4). http://doi.org/https://doi.org/10.5860/crl.75.4.557

Bresnahan, M., & Johnson, A. (2013). Assessing scholarly communication and research data training needs. *Reference Services Review*, *41*(3), 413–433. Retrieved from http://www.emeraldinsight.com/doi/pdf/10.1108/RSR-01-2013-0003

Byatt, D., Luca, F. De, Gibbs, H., Patrick, M., Rumsey, S., & White, W. (2013). *Supporting researchers with their research data management: Professional service training requirements—a DataPool Project Report*. Retrieved from http://eprints.soton.ac.uk/352107/

California, University of (2017). University of California, Berkeley Library strategic plan. Retrieved April 12, 2018, from http://stories.lib.berkeley.edu/strategicplan/

Carlson, J., & Stowell-Bracke, M. (2013). Data management and sharing from the perspective of graduate students: An examination of the culture and practice at the Water Quality Field Station. *Portal: Libraries and the Academy*, *13*(4), 343–361. http://doi.org/10.1353/pla.2013.0034

Cox, A., Verbaan, E., & Sen, B. (2012). Upskilling liaison librarians for research data management. *Ariadne*, (70). Retrieved from http://www.ariadne.ac.uk/issue70/cox-et-al

Cox, A., Verbaan, E., & Sen, B. (2014). A spider, an octopus, or an animal just coming into existence? Designing a curriculum for librarians to support research data management. *Journal of eScience Librarianship*, *3*(1). http://doi.org/10.7191/jeslib.2014.1055

Cyber Infrastructure Council. (2007). *Cyberinfrastructure Vision for 21st Century Discovery*. Retrieved from http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf

De Cagna, J. (2001). Tending the garden of knowledge: a look at communities of practice with Etienne Wenger. *Information Outlook*, *5*(7), 6–12.

de Smaele, M., Verbakel, E., & Potters, N. (2013). Data intelligence training for library staff. *International Journal of Digital Curation*, *8*(1), 218–228. http://doi.org/10.2218/ijdc.v8i1.255

Delserone, L. M. (2008). At the watershed: Preparing for research data management and stewardship at the University of Minnesota Libraries. *Library Trends*, *57*(2), 202–210. http://doi.org/10.1353/lib.0.0032

Farrell, R., & Badke, W. (2015). Situating Information Literacy in the Disciplines : A Practical and Systematic Approach for Academic Librarians. *Reference Services Review*, *43*(2), 319–340.

Jaguszewski, J. M., & Williams, K. (2013). New Roles for New Times: Transforming Liaison Roles in Research Libraries, (August), 1–17. Retrieved from http://www.arl.org/storage/documents/publications/nrnt-liaison-roles-revised.pdf

Johnson, A., & Bresnahan, M. (2015). DataDay!: Designing and assessing a research data workshop for subject librarians. *Journal of Librarianship and Scholarly Communication*. Retrieved from http://jlsc.ubiquitypress.com/articles/10.7710/2162-3309.1229/galley/89/download/

Johnston, L. R., Carlson, J. R., Hswe, P., Hudson-Vitale, C., Imker, H., Kozlowski, W., … Stewart, C. (2017). Data curation network: how do we compare? A snapshot of six academic library institutions' data repository and curation services. *Journal of eScience Librarianship*, *6*(1). http://doi.org/10.7191/jeslib.2017.1102

Johnston, L. R., Olendorf, R., Stewart, C., Carlson, J., Hudson-vitale, C., Imker, H., & Kozlowski, W. (2017). *Data curation network: A cross-institutional staffing model for curating research*. Retrieved from http://hdl.handle.net/11299/188654

Kezar, A. (2012). Bottom-up/top-down leadership: Contradiction or hidden phenomenon. *The Journal of Higher Education*, *83*(5), 725–760. http://doi.org/10.1353/jhe.2012.0030

Latham, B. (2017). Research data management: Defining roles, prioritizing services, and enumerating challenges. *The Journal of Academic Librarianship*, *43*(3), 263-265. http://doi.org/10.1016/j.acalib.2017.04.004

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York:

Cambridge University Press.

Lloyd, A. (2007). Learning to put out the red stuff: Becoming information literate through ciscursive practice. *The Library Quarterly*, *77*(2), 181–198. http://doi.org/10.1086/517844

Lyon, L. (2016). Librarians in the lab: Toward radically re-engineering data curation services at the research coalface. *New Review of Academic Librarianship*, *22*(4), 391–409. http://doi.org/10.1080/13614533.2016.1159969

Mattern, E., Brenner, A., & Lyon, L. (2016). Learning by teaching about RDM: An active learning model for internal library education. *International Journal of Digital Curation*, *11*(2), 27–38. http://doi.org/10.2218/ijdc.v11i2.414

Peterson, R. A. (2017). Asking specific questions. In *Constructing Effective Questionnaires* (pp. 83–100). Sage.

Pryor, G., & Donnelly, M. (2009). Skilling up to do data: Whose role, whose responsibility, whose career? *International Journal of Digital Curation*, *4*(2), 158–170. http://doi.org/10.2218/ijdc.v4i2.105

Rice, R. (2014). Research data MANTRA: A labour of love. *Journal of eScience Librarianship*, *3*(1), 4. http://doi.org/10.7191/jeslib.2014.1056

Spillane, J., Halverson, R., Diamond, J. (2004). Towards a theory of leadership practice: A distributed perspective. *Journal of Curriculum Studies*, *36*(1), 3–34.

Tibbo, H., & Jones, S. (2015). Research data management and sharing. Retrieved September 18, 2017, from https://www.coursera.org/learn/data-management

Wittenberg, J., & Elings, M. (2017). Building a research data management service at the University of California, Berkeley: A tale of collaboration. *IFLA Journal*, *43*(1), 89–97. http://doi.org/10.1177/0340035216686982