



Team-based data management instruction at small liberal arts colleges

International Federation of
Library Associations and Institutions
2017, Vol. 43(1) 105–118
© The Author(s) 2016
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0340035216678239
journals.sagepub.com/home/iff



Ryan Clement
Middlebury College, USA

Amy Blau
Whitman College, USA

Parvaneh Abbaspour
Lewis & Clark College, USA

Eli Gandour-Rood
University of Puget Sound, USA

Abstract

This paper describes a collaborative approach taken by librarians at five small, regional liberal arts colleges to developing/enhancing research data management services on their campuses. The five colleges collectively belong to a consortium known as the Northwest Five Consortium. Over 10 months, librarians from the five schools collaborated to plan a data management and curation workshop with the goals of developing relationships with researchers working with data, developing their own research data management skills and services, and building a model for future training and outreach around institutional research data management services. This workshop brought together research teams including faculty, students, and librarians, and incorporated active learning modules as well as in-depth pre-workshop discussion. This article will discuss the context and background for this workshop, the model itself, and the outcomes and possibilities for future developments.

Keywords

Communities of practice, data services, information literacy and instruction, preservation and conservation, research data management

Submitted: 17 May 2016; Accepted: 19 September 2016.

Introduction

Over the past decade academic libraries at institutions of all sizes have been developing services to support research data management and curation (Choudhury, 2008; Goldstein and Oekler, 2011). Spurred by mandates passed down from federal funding agencies that are creating new imperatives for data management and sharing among the research community, academic libraries have seized this opportunity to address this evolving information need on campus (Heidorn, 2011; Hswe and Holt, 2011; Walton, 2010). Extending their training in organizing and managing information, as well as digital preservation and records

management, librarians are disposed both by their skill-set and their roles in scholarly communications support to address these needs (Brandt, 2007: 365). Moreover, the provision of research data services is an opportunity for libraries to demonstrate their relevance to the campus community at a time when conceptions of scholarly communication are evolving. On campuses where libraries compete with other

Corresponding author:

Ryan Clement, Middlebury College, 110 Storrs Ave, Middlebury, VT Vermont 05753, USA.
Email: rclement@middlebury.edu

departments for resources, the provision of support for research data management (RDM) can be a fresh approach for libraries to build their value on campus in both the scholarly communications and instructional domains.

Common challenges in developing research data management services

Training and time investment. There are some common challenges encountered by librarians working to establish RDM services. Although data management is a ready match for librarians' skill-sets, they may lack specific training in these areas or be unfamiliar with the domain-specific knowledge needed to apply more general curation skills effectively. Furthermore, given that the vast majority of research data is produced digitally, managing this information requires a perpetual investment of time to stay up to date with new technologies. RDM is also a relatively new field experiencing rapid development across a variety of disciplines, requiring further time to stay abreast of domain-specific best practices.

Building campus relationships. Beyond requirements for professional development and training, the establishment of RDM services and workflows requires developing stakeholder relationships across the institution. Providing effective RDM frequently requires that the library develop new relationships and build existing partnerships across campus with units such as Information Technology (IT) and Sponsored Research. As familiar and accepted as conversations about RDM have become in the academic library community, other units may be entirely unfamiliar with the library's new (and often self-defined) mandate to support data management and may, for a variety of reasons, be reticent to engage with the library to address these needs. This initial caution may extend to the very stakeholders these services aim to support – researchers.

Outreach to faculty partners. Beyond these initial questions about new services and roles, librarians seeking to establish RDM services may encounter other challenges in working with faculty researchers. Among these is the deceptively simple task of informing researchers that these new services exist, and retraining faculty to reach out to the library when RDM needs arise. This retraining process involves two further challenges – convincing the researchers that the library is capable of providing these services, and defending the novel concepts of data management plans in particular and data sharing in general.

Researchers may fiercely defend disciplinary traditions around sharing research, and see these new mandates and requests for their output as an encroachment, even as these requirements have often been crafted with the utmost deference to establish disciplinary practices.

Background on the Northwest Five Consortium institutions.

This paper describes a collaborative approach taken by librarians at five small, regional liberal arts colleges to developing/enhancing research data management services on their campuses. These colleges are Lewis & Clark College, Reed College, Whitman College, Willamette University, and University of Puget Sound. The five colleges collectively belong to a consortium known as the Northwest Five Consortium (NW5C), which has the mission of promoting cross-institutional collaboration and is supported by a grant from the Andrew W. Mellon Foundation. Over 10 months, librarians from the five schools collaborated to plan a data management and curation workshop with the goals of developing relationships with researchers working with data, developing their own RDM skills and services, and building a model for future training and outreach around institutional RDM services. The workshop was supported with a mini-grant from the NW5C Fund for Collaborative Inquiry.¹ This article will discuss the context and background for this workshop, the model itself, and the outcomes and possibilities for future developments.

All five institutions are small, private liberal arts colleges, with student enrollments ranging between 1400 and 2600 full-time equivalent (FTE). They share several metrics by which the Carnegie Classification of Institutions of Higher Education (Indiana University Center for Postsecondary Research, n.d.) classifies colleges and universities in the United States. They are four-year Baccalaureate Colleges with an Arts & Sciences Focus – the majority of degrees awarded at each college are non-professional undergraduate degrees (Bachelor of Arts or Bachelor of Sciences in the arts, humanities, social sciences, or natural sciences, rather than in disciplines such as business, nursing, or engineering). All five institutions are primarily residential (at least half of the undergraduate students live on campus and at least 80% attend full-time) and all rank among the most selective (80th to 100th percentile) of baccalaureate institutions. The concentration on undergraduate education in a residential setting, the relatively small number of students and the generally low student to faculty ratio mean that there is a strong expectation that students will know their professors and fellow

students well, and that the personal attention that students receive from faculty and staff will translate into “high-impact” educational experiences for students (Kuh et al., 2008). The faculty expectations at selective liberal arts colleges generally follow a “teacher-scholar” model; while excellent teaching is a fundamental expectation, that teaching is rooted in disciplinary scholarship. Faculty at these institutions therefore have active research agendas, which often rely upon research assistance from their undergraduate students.

While many of the challenges libraries face in establishing RDM services are common across a wide variety of institution types, some of these challenges are magnified in unique ways at small liberal arts colleges. The intimate, student-centered environment that is characteristic of liberal arts colleges poses additional challenges to developing RDM services that must be addressed by librarians in these contexts. Faculty tend to be engaged in a particularly time-intensive style of teaching and mentoring, leaving limited time to spend on lab management. In this context, data management can fall by the wayside. Research assistants are often undergraduate students with little exposure to, and instruction in, managing research data. Furthermore, there tends to be a high turnover rate among these research assistants; many will spend only 2–3 years working in a lab before graduation, sometimes with interruptions for time off or study abroad.

These immediate, situational challenges are often compounded by the larger institution-level privileging of curricular needs over research in the allocation of resources. By virtue of being in a smaller institution there is often less infrastructure (both human and technological) in place to support research and fewer resources that can be devoted to these needs than in large research-intensive universities. This holds true in the library where data management responsibilities are often added to a traditional liaison position in addition to other instructional, technical, and functional work roles.

At the same time, the opportunities for data management and data information literacy at liberal arts colleges and primarily undergraduate institutions (PUIs) are also significant. The close-knit environment lends itself to collaboration between librarians and faculty, and data management services can act to reinforce or grow liaison relationships, especially with departments less reliant on the library for more traditional services. The combination of liaison responsibilities with RDM services allows for the flexibility to be able to consult with faculty and students *in situ*, in a laboratory setting, if that works best for the research group. There are many possibilities

for developing innovative approaches to data management and curation services and for providing data information literacy training for undergraduate students.

Literature review

Library data services

While most of the initiators of library data services have been large research universities (Antell et al., 2014; Cox and Pinfield, 2014; Heidorn, 2011; Soehner et al., 2010; Tenopir et al., 2012, 2013, 2014), there is a growing body of literature about data services at Master’s universities and PUIs (Goldstein and Oelker, 2011; Scaramozzino et al., 2012; Shorish, 2012; Stamatoplos et al., 2016; Touns and Hughes, 2013), and about how outreach efforts related to data literacy and data management may be implemented with undergraduate students at a range of institution types (Ball and Medeiros, 2012; MacMillan, 2010; Mooney et al., 2014; Piorun et al., 2012; Qin and D’Ignazio, 2010; Reisner et al., 2014; Shorish, 2015; Stephenson and Caravello, 2007; Strasser and Hampton, 2012; Zilinski et al., 2014). Librarians at PUIs have argued that while resources may be limited at smaller institutions, the same imperatives to act as data stewards and to transmit best practices apply (Shorish, 2012). The key factor for smaller institutions is the adaptation of emerging best practices for their environments:

Large research institutions may have more resources and staff, and their need for data curation may be greater. But we at smaller institutions are poised to learn from their pioneering work, borrow accordingly, and tailor data support services to the local needs of our patrons. (Touns and Hughes, 2013: 232)

Library data support services for undergraduate students draw upon sets of competencies, or literacies, related to data production or use. Calzada Prado and Marzal (2013: 126) define *data literacy* as “the component of information literacy that enables individuals to access, interpret, critically assess, manage, handle and ethically use data”. Qin and D’Ignazio (2010: 2) define *science data literacy* as a praxis-based skillset with emphasis on “functional ability in data collection, processing, management, evaluation, and use”. Carlson et al. (2011: 634) distinguish *data information literacy* from data literacy, statistical literacy, and information literacy with respect to the production of information in addition to its consumption, bringing the various literacies together as it “merges the concepts of researcher-as-producer and researcher-as-consumer of data products”. They situate data

information literacy competencies within the ACRL Information Literacy Competency Standards in order to propose the essential components of a data information literacy program administered by librarians (Carlson et al., 2011: 652). Considering data information literacy in light of the ACRL *Framework for Information Literacy for Higher Education*, Shorish (2015: 100) argues that data information literacy “should be treated as any of the other literacy competencies and incorporated into the workflow of outreach librarians”.

Many examples of librarian-led data information literacy instruction for undergraduates are in classroom settings. However, it can be difficult to find space in a tight undergraduate science curriculum for a stand-alone data literacy or data management course (Qin and D’Ignazio, 2010: Lessons Learned 3). This is one reason to look to the undergraduate research experience as a place to integrate data literacy or data management skills into undergraduate research experiences; another reason is the ability to immediately relate skills learned to their authentic contexts. In their needs assessment of data information literacy for faculty and students, Carlson et al. (2011: 648) find that while there were common areas of need, faculty needs were related to data they created themselves, while the students surveyed were using data from external sources as part of a course. However, Stamatoplos (2009: 240) points out that the differences between information needs for students and faculty are reduced when students are undertaking authentic research projects.

Undergraduate research

The Council on Undergraduate Research defines undergraduate research as “an inquiry or investigation conducted by an undergraduate student that makes an original intellectual or creative contribution to the discipline” (Rowlett et al., 2012: 2). Undergraduate research differs from most coursework because of its open investigative nature, usually managed through the mentorship of a faculty researcher (Stamatoplos, 2009: 237). Involvement with undergraduate research has been identified as a “high-impact” practice which contributes to student engagement and success (Kuh et al., 2008: 14). While the Boyer Report called for making inquiry-based learning the standard at research universities in order to improve the quality of undergraduate experience at such institutions (Boyer Commission on Educating Undergraduates in the Research University, 1998: 15), the emphasis on active, collaborative learning and faculty mentorship has been standard at liberal

arts colleges, where undergraduate students rather than graduate students work as research assistants for faculty research groups (Shorish, 2015: 101).

By taking part in undergraduate research experiences (UREs), students not only help to create new knowledge, but also learn about the norms of scholarly work in their discipline. Thus, undergraduate research experiences allow students to be integrated into communities of practice (Hunter et al., 2007: 37). However, formal instruction in data information literacy is not necessarily the norm in these communities of practice. Hunter et al. (2007: 46) identified eight categories of gains for students who participated in science UREs, including intellectual gains of “thinking and working like a scientist”, learning professional norms of “becoming a scientist”, and acquiring skills in various information literacy-related categories such as “work organization”, “computer”, and “information retrieval”. Hunter et al. do not consider data management competencies explicitly; they might be included in any of the above-mentioned categories, although the praxis-based elements probably are best categorized with skills.

The undergraduate research experiences described by Hunter et al. do not include instruction or mentorship from librarians. Stamatoplos (2009: 239) calls for formal involvement of librarians with undergraduate research programs, as students conducting independent research have even greater need for advanced information and library skills. Hensley (2015: 722) has also shown ways in which library information literacy (IL) initiatives can intersect productively with UREs, as they share emphases on developing “critical thinking and problem-solving skills”. Respondents to her survey provided IL instruction for undergraduate researchers in a variety of formats, including one-on-one with students or faculty/mentors, workshops with lab groups, and involvement as team-teacher or as instructor of record in a credit-bearing course (Hensley, 2015: 730). The examples of IL topics for undergraduate research, however, focused primarily on database searching and citation management; explicitly data-related aspects of IL accounted for just over 2% of the topics taught (numeric and spatial data 1%, data visualization .7%, developing a data management plan .5%) (Hensley, 2015: 735).

An instance of librarians working directly with undergraduate research groups is the social-work pilot project undertaken by Mooney et al. They describe their work embedding librarians with undergraduate research teams over the course of two semester-long research projects in social work, with biweekly meetings (Mooney et al., 2014: 374). Among the advantages to this approach are just-in-time lessons that fit

the development and workflows of the projects. Mooney et al. identify the potential for partnership between librarians and research teams at a place of knowledge production. The involvement of librarians with research groups helps to integrate a range of information literacy skills into the undergraduate research experience, something that Mooney et al. find lacking overall:

It is our assessment that data management and the broader scope of data information literacy, indeed even basic library research skills, are not widely perceived as explicit goals of participation in undergraduate research despite an overarching goal for the advancement of real world research experience. (Mooney et al., 2014: 371)

Because the approach taken to embed librarians in these projects will not scale, Mooney et al. (2014: 383) foresee offering librarian-run workshops for faculty and undergraduate researchers, complemented by librarian attendance at research team meetings.

Librarian training

The growth in interest in data services in libraries also has significant impact on librarian training in data services. While library and information schools are adding data services and data curation tracks for new graduates (Creamer et al., 2012; Keralis, 2012), additional training of librarians, especially in departmental liaison or institutional repository roles, may be needed for them to evaluate and develop services in their libraries (Bresnahan and Johnson, 2013), especially if data curation roles were not part of their original career plans (Pryor and Donnelly, 2009: 164). Tenopir et al. (2013: 72) find that while over 75% of librarians who support research data services as an integral part of their job responsibilities feel prepared to do so, less than 50% of those who occasionally support research data services feel prepared.

Both online and in-person training opportunities in RDM exist for librarians, with a broad range of instructional models, fee structures, and time commitments (ANDS, n.d.; EDINA and Data Library, University of Edinburgh, n.d.; Guy, 2013; Research Data Netherlands, 2016; Rice, 2014; Verbakel et al., 2013). A few models of librarian training that explicitly incorporate face-to-face training or interaction with researchers are of interest for the incorporation of elements important to the workshop model discussed in this article. The Data Information Literacy Project (datainfo.org) presents a model in which librarians familiarize themselves with types of research in specific disciplinary areas, interview researchers using the Data Curation Profile, and then work with faculty

to develop data information literacy programs for graduate or undergraduate students (Carlson et al., 2015). The immersive Informatics training program at the Library at the University of Melbourne, Australia, embedded library and IT support staff in a research context with data producers for a ten-module series of topics presented over 16 weeks (Shadbolt et al., 2014). In order to provide required training for a library representative for Statistics Canada data at participating institutions across Canada, the Data Liberation Initiative (DLI) instituted a common curricular model built on regional workshops. The in-person regional trainings created a “sense of community” among participants and trainers that continued beyond the workshop proper (Watkins et al., 2004: 19).

Northwest Five Consortium workshop model in theory

The planning process for the workshop took place over 10 months and was conducted primarily via videoconferencing meetings by the five librarian co-facilitators. The group was convened by the principal investigators in September of 2014 to respond to an open invitation from the NW5C for grant applications. Funding was available to support thematic workshops: cross-disciplinary initiatives to enhance teaching and learning at all member institutions. The grant application was submitted in October, and notification of funding approval was received in December. The librarian co-facilitators met more frequently beginning in January 2015 to plan the workshop. In these meetings the planners formulated a workshop model to help address the common challenges faced by librarians developing data management services at the NW5C member colleges. The model proposed bringing together teams from each school for a one-and-a-half-day workshop led by an outside facilitator. Each detail of the model from team composition to the breakdown of the curriculum was designed to maximize the impact of the workshop on the larger endeavor of establishing RDM services at each of the institutions.

Team model

The team model proposed bringing together a team from each school composed of a faculty researcher, one or two student researchers working with that faculty member, a librarian, and an educational technologist or IT support staff. Each team united those involved in addressing the data management challenges in a particular lab. This team model also served to model the type of librarian-researcher collaboration

the planners sought to develop on their campuses. To this end, once a team was identified from each institution, the librarian would work with the team to undertake a preliminary assessment of current RDM practices.

Facilitation

Due to the relative inexperience of the majority of the planning team with RDM services, an outside facilitator who was skilled at leading a diverse audience through an introductory data management curriculum was hired. Selecting an outside facilitator brought two advantages: first, it allowed the librarians to fully engage as members of their teams, working side-by-side through each activity; second, it provided an opportunity for the librarians to receive the RDM training themselves, learning both the content as well as the approach to teaching RDM.

Curriculum

The workshop curriculum was crafted by the planners, in collaboration with the facilitator, with the aim of developing participant awareness and understanding of data management issues. The curriculum for the first day was adapted from the New England Collaborative Data Management Curriculum (Lamar Soutter Library, University of Massachusetts Medical School, n.d.) as well as the DataONE modules (DataONE, 2015) and involved lecture and discussion sections interspersed with group activities to create an active, engaging experience. The curriculum was calibrated based on the participating teams' interests or stated RDM challenges (this question was posed in the recruitment process). The following half-day of the workshop was set aside as work time for teams to apply the lessons learned the previous day while the librarians gathered separately to talk about their work to establish RDM services at their respective institutions and further develop their community of practice.

Workshop model in practice

The workshop model evolved over the course of the planning process. Recruitment challenges affected the makeup of the teams and their disciplinary backgrounds, changes were made to the pre-workshop preparation process to accommodate the needs of participants, and the curriculum itself was modified prior to the delivery of the workshop to provide more participant-centered outcomes. The foundational elements of the model remained: a flexible, segmented curriculum centered around active-learning modules;

a team of researchers from each institution; participation by faculty researchers who generate data in their teaching lab groups; and involvement of undergraduate research assistants.

The planned makeup of the institutional teams, each consisting of a faculty principal investigator, one or two student researchers, a librarian, and an educational technologist or IT support person, was a deliberate attempt to bring together stakeholders with diverse perspectives and complementary skillsets. In fact, none of the teams ended up in the planned configuration of one faculty, one student, and one IT support, although all retained the librarian and faculty participants. All of the librarians reported approaching multiple faculty members at their institutions before each was successful in identifying a faculty member who fit all of the necessary criteria to attend: an interest in RDM, an active research program with undergraduate students involved in generating and working with data, and, crucially, availability on the date of the workshop. Although the original intent had been to recruit faculty from both social science and science disciplines, the planners found it challenging to recruit from the social sciences, since several of the librarians themselves were primarily liaisons to natural science departments. The final participating teams consisted of two chemists, one developmental biologist, and two environmental scientists. The fifth team, with a sociologist, started but did not complete the first day of the workshop.

In addition to the narrowing of disciplinary scope, there was not a consistent configuration of team participants. One institution sent two faculty researchers in environmental science but lacked an undergraduate student researcher, while the chemists, coming from separate institutions, each brought two undergraduate students. None of the teams brought a dedicated technology specialist. Another team brought in place of an IT support person a second faculty participant who, while not a member of the primary research group, was the primary technician for the instruments which were used by the members of the research group in their data collection. This proved to be a helpful team configuration, as the second faculty participant, while not primarily a teaching faculty member, had substantial experience both with the learning needs of undergraduate students, and also with the nature of the data sets being generated by undergraduates within their home department. During workshop activities requiring teams to make plans for future data management workflows within their lab, the second faculty participant enabled the group to delineate very specific procedures based on her expertise and familiarity with the instruments most commonly used by the lab.

Further modifications to the model were made during the pre-workshop preparatory period, when the librarians engaged with their faculty members and completed a modified Data Curation Profile (DCP) interview (Carlson, 2010). Several faculty members expressed some resistance to the time-intensive nature of completing an entire DCP. During the course of the oral interviews, the librarians were able to take stock of each faculty participant's comfort level and experience with data management principles, as well as their immediate needs. In general, faculty had little to no experience explicitly engaging with RDM as a skillset, and their priorities were focused on learning how to increase efficiency in their lab settings. These were unsurprising concerns for faculty at teaching-focused institutions, but did cause the librarians to suggest prioritizing the creation of practical documents such as lab protocols as an outcome of the workshop.

While incorporating these changes, the facilitator did retain as a central component of the curriculum a series of engaging activities to allow the participants to begin to immediately begin putting into practice skills related to the best practices being covered. This emphasis on active learning was a key element of the workshop's success, and was mentioned by the majority of participants in the feedback collected during the post-workshop assessment.

Outcomes and assessment

Overall outcomes

Overall, the workshop was successful in meeting several high-level goals of the planners. The team model developed new learning opportunities for students while also increasing faculty and staff communication in a collaborative cross-institutional environment. Within the workshop itself, participants increased their awareness and understanding of data management topics. The NW5C workshop piloted a model for librarian-researcher collaboration across campuses, and the incorporation of the individual DCP interviews allowed librarians, together with researchers, to create a preliminary assessment of RDM practices. The process of applying for the grant and planning the workshop itself allowed librarians to better understand the RDM landscape at their own institutions and at peer institutions. These findings are based on three assessments carried out before, immediately after, and eight months after the workshop.

A pre-workshop survey was distributed to all participants to assess base levels of understanding of and familiarity with the material to be covered. A post-workshop survey was distributed immediately



Figure 1. Results from pre- and post-workshop surveys of faculty and student workshop participants on describing best practices in research data management.

following the workshop. Separate assessment instruments were used for the faculty/student research teams and the librarian participants. A follow-up survey was distributed to faculty/student research teams ($n = 12$, due to unavailability of one team) eight months after the workshop to assess their continued awareness of RDM and curation best practices and their longer-term evaluations of the workshop's effectiveness. Questions from all instruments are available in Appendix A. Total registered attendees of the workshop included 14 faculty/student researchers and six librarians; pre-workshop completion was 79% for the faculty/student research teams ($n = 11$) and 100% for the librarians ($n = 6$) while post-workshop completion was 57% for the faculty/student teams ($n = 8$) and 67% for the librarians ($n = 4$). The lower post-workshop completion rate was primarily due to a faculty/student team from one institution leaving the workshop. The follow-up survey had a response rate comparable to the two earlier assessments, with an overall response rate of 83% from students and faculty ($n = 10$). While the sample sizes are too small to yield statistically significant inferences, the trends in both the numerical data and the qualitative responses support the conclusion that the workshop successfully met its objectives. All respondents who completed the post-workshop survey agreed with the statement that the workshop "was an effective way to learn about Research Data Management".

The quantitative ratings from students, faculty, and librarians show that in all cases, participants' confidence in their knowledge of RDM practices increased as a result of the workshop (see Figures 1 and 2). The answers to qualitative questions provide more insight into what participants found most useful. Both students and faculty noted specific skills (e.g. file naming conventions, file types, readme files) as well as general concepts (e.g. data curation, metadata) as



Figure 2. Results from pre- and post-workshop surveys of faculty and student workshop participants on finding and applying data management principles appropriate for their research.

important takeaways from the workshop. After the workshop, participants expanded their definitions of “data management” to emphasize the importance of planning, an aspect that had been highlighted throughout the workshop presentations, and was absent from pre-workshop answers.

Faculty outcomes

All responses to the post-workshop survey from faculty members referenced the importance of working in teams and fostering collaborations. Faculty supported extending the workshop’s mission to other groups on their home campuses. One faculty member wrote, “Working in teams was KEY. We need a follow up either at our own institution or another NW5C workshop!” Following the workshop, faculty members defining the “research data life cycle” were more likely to provide more detailed explanations, specifically emphasizing data storage and preservation issues. Several faculty members commented that the preliminary DCP work was helpful in preparing for the workshop and for further RDM planning. Multiple faculty commented on their intent to establish protocols and training procedures for their labs which were previously non-existent. In general, faculty feedback focused on appreciation for newly learned skills which could be immediately implemented in their labs for data backups, efficient data storage, and consistency in recording and formatting data.

Following the workshop, faculty members informally reported to librarians that their research groups had either instituted or begun to develop new file-naming plans and research data workflows to facilitate more consistent documentation and backup. The answers to qualitative questions in the follow-up survey indicate that the researchers have pursued these

preliminary plans. One group has “a protocol for file naming and creating/storing metadata on every set of experimental data we produce”. Another has “used the file naming conventions routinely and [has] converted old, ‘pre-workshop’ files to the convention”. Research groups report that they still face challenges in areas such as storage, backing up data and dealing with older files, not to mention the consistent implementation of file-naming and metadata conventions. In one case, a faculty member noted, “I believe some of these were addressed in the workshop, but the ‘doing’ is much harder” and suggested further communication and support among research groups and librarians to encourage both brainstorming and follow-through.

The quantitative data indicate that after eight months, faculty (and students) felt at least as well prepared to describe best practices and apply discipline-appropriate data management to their research as they had immediately after the workshop, and their answers to the open-ended question of how they would describe RDM were consistent with that self-assessment. One faculty answer in particular summed up key components of data management:

The data we collect in our lab is the foundation of our work and as such is extremely valuable so we must take good care of it. This means we must keep our data useful (describe it thoroughly with detailed and consistent metadata), keep our data organized (use consistent file naming conventions and store it in the correct folders), and keep our data safe (back it up consistently and provide accurate links in our notebooks).

Student outcomes

Student evaluations immediately following the workshop included the recognition of the value of the skills and concepts covered. A student wrote:

I found it very helpful to be introduced to the concepts behind data management . . . Being aware of the advantages of data management, I can progress in my career and find various applications for [the concepts]. This seems far more advantageous than simply being told by a PI that the lab follows XYZ protocols, which I may or may not continue to apply in the future.

One of the student researchers voiced a desire that future workshops provide more space for interaction and discussion with student workshop participants from other institutions. While not all teams were able to bring students along, there was a sense that the teams who did so found it valuable. Students were able to disseminate the lessons they learned on

returning to their campuses both informally, in talking with other student research assistants in their labs, and, in at least one case, through a formal presentation to the research teams in their department.

Librarian outcomes

In their post-workshop assessments, all librarian participants reported a significant increase in comfort levels with the subject matter and in feeling prepared to do outreach to faculty or colleagues. All librarians mentioned networking or collaboration with NW5C peers when asked about the most important outcome of the workshop. Librarians also raised some issues that were not within the explicit scope of the workshop, such as building a shared data repository, and posed questions about how data management plans might be applied on an individual level to undergraduate research.

The lessons learned for the librarian organizers of the workshop encompassed planning issues on many levels. The librarians had varying levels of training in RDM before the workshop, and some would have appreciated additional “train the trainer” instruction before the workshop itself. Cross-campus collaborations were possible in large part due to technological affordances of shared online documents and Google Hangouts. Having a standing online meeting was very important for planning purposes. Post-workshop use of some of these digital connections could have been improved; having a listserv or other digital meeting space set up before the end of the workshop would have facilitated its ongoing impact.

Conclusion

The NW5C Data Curation Workshop model is a successful, and sustainable, model for developing RDM services and building communities of practice for small liberal arts colleges. While the costs for the first iteration of this workshop model were grant-funded, the planners believe that costs for future iterations of the workshop would be greatly reduced and potentially feasible without grant funding. The primary reduction in costs comes from the fact that the planning librarians gained valuable RDM skills themselves, as well as being exposed to how the facilitator actually taught the material, thus providing valuable professional development training while simultaneously providing outreach services to faculty constituents. Aside from this facilitation cost, the most expensive part of the implementation was reimbursement of travel costs for participants – a situation which was exacerbated by the remote nature of at least one participating institution. Taking further

advantage of digital tools and distance-learning technologies for long-distance collaboration could lower these costs even further.

Faculty outcomes in the area of RDM skills were positive, but the most important long-term outcomes for faculty participants were their development as RDM evangelists for their campuses. In seeking to create a campus culture that values solid RDM practices, librarians need faculty partners who will speak to their importance. Another important outcome was the opportunity to extend the classic liberal arts teacher-scholar model to the domain of RDM skill development. While undergraduate student research assistants are more transitory than graduate research assistants, students attend liberal arts institutions because of the belief in education through close contact with faculty both inside and outside of the classroom. RDM training at the undergraduate level can also help to prepare upcoming graduate students to recognize and promote the importance of data management. This workshop demonstrated the feasibility of faculty, undergraduates, and librarians learning and developing their RDM skills in a collaborative environment.

The development of a community of practice was one of the most powerful outcomes of this experience for the planning librarians. The planning librarians have stayed in regular contact with each other concerning RDM developments on their campuses, and have continued working together on conference presentations as well as articles about their model. These activities have continued despite one of the planners moving to a new position at an institution on the opposite coast. An often unspoken challenge of starting an RDM program at a liberal arts college library is that, in contrast to larger institutions, the librarian tasked with developing these services is typically working alone. With many of the models for these services coming from larger institutions, the new data librarian may have few colleagues to turn to when wondering what will or will not work on their campus, and how to scale these services appropriately. Though just beginning, the development of a community for discussion, planning, and sharing of best practices particular to a smaller institution is invaluable.

A second iteration of the model, with reduced costs, was planned for the summer of 2016, and there is interest in replicating this instructional model for liberal arts institutions outside of the Pacific Northwest. The modules and tools used to develop this workshop are all openly available, and the planners are enthusiastic about sharing their experiences, including the successes and challenges. Whether they choose to proceed independently, or as part of the

growing community of liberal arts colleges developing RDM services, institutions that use a research-group-based approach to teaching RDM will not only improve practices on their own campus, but also help to strengthen relationships and change campus cultures.

Acknowledgements

The authors gratefully acknowledge those who dedicated their time and effort to this project: Mini-grant Co-PIs: Mark Dahl, Dena Hutto; Other Workshop Planners: Gary Klein, Michael Spalti; Workshop Facilitator: Amanda Whitmire, PhD; and all workshop participants.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note

1. The co-principal investigators on this mini-grant were Dena Hutto, Norman F. Carrigg College Librarian, Reed College and Mark Dahl, Director of the Aubrey R. Watzak Library, Lewis & Clark College.

References

- ANDS (n.d.) *23 (Research Data) Things*. Australian National Data Service. Available at: <http://ands.org.au/partners-and-communities/23-research-data-things> (accessed 18 April 2016).
- Antell K, Foote JB, Turner J, et al. (2014) Dealing with data: Science librarians' participation in data management at Association of Research Libraries institutions. *College & Research Libraries* 75(4): 557–574.
- Ball R and Medeiros N (2012) Teaching integrity in empirical research: A protocol for documenting data management and analysis. *Journal of Economic Education* 43(2): 182–189.
- Boyer Commission on Educating Undergraduates in the Research University (1998) *Reinventing Undergraduate Education: A Blueprint for America's Research Universities*. Stony Brook, NY: SUNY Stony Brook; Carnegie Foundation for the Advancement of Teaching.
- Brandt DS (2007) Librarians as partners in e-research: Purdue University Libraries promote collaboration. *College & Research Libraries News* 68(6): 365–396.
- Bresnahan MM and Johnson AM (2013) Assessing scholarly communication and research data training needs. *Reference Services Review* 41(3): 413–433.
- Calzada Prado J and Marzal MÁ (2013) Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri: International Journal of Libraries & Information Services* 63(2): 123–134.
- Carlson J (2010) *Data Curation Profiles Toolkit*. Available at: <http://docs.lib.purdue.edu/dcptoolkit/> (accessed 21 November 2016).
- Carlson J, Fosmire M, Miller CC, et al. (2011) Determining data information literacy needs: A study of students and research faculty. *portal: Libraries and the Academy* 11(2): 629–657.
- Carlson J, Sapp Nelson M, Bracke M, et al. (2015) *The Data Information Literacy Toolkit*. Purdue University. Available at: <http://docs.lib.purdue.edu/dilsymposium/2013/interviewinstruments/1/> (accessed 30 March 2016).
- Choudhury GS (2008) Case study in data curation at Johns Hopkins University. *Library Trends* 57(2): 211–220.
- Cox AM and Pinfield S (2014) Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science* 46(4): 299–316.
- Creamer AT, Morales ME, Kafel D, et al. (2012) A sample of research data curation and management courses. *Journal of eScience Librarianship* 1(2): 4.
- DataONE (2015) *DataONE Education Modules*. Available at: <https://www.dataone.org/education-modules> (accessed 30 March 2016).
- EDINA and Data Library, University of Edinburgh (n.d.) *DIY Research Data MANTRA Training Kit for Librarians*. Available at: <http://datalib.edina.ac.uk/mantra/libtraining.html> (accessed 20 March 2016).
- Goldstein S and Oelker S (2011) Planning for data curation in the small liberal arts college environment. *Sci-Tech News* 65(3). Available at: <http://jdc.jefferson.edu/scitechnews/vol65/iss3/4> (accessed 20 March 2016).
- Guy M (2013) *RDM Training for Librarians*. DCC RDM services case studies, Digital Curation Centre. Available at: <http://www.dcc.ac.uk/sites/default/files/documents/publications/case-studies/Training-for-librarians-case-study.pdf> (accessed 20 March 2016).
- Heidorn PB (2011) The emerging role of libraries in data curation and e-science. *Journal of Library Administration* 51(7/8): 662–672.
- Hensley MK (2015) A survey of instructional support for undergraduate research programs. *portal: Libraries and the Academy* 15(4): 719–762.
- Hunter A-B, Laursen SL and Seymour E (2007) Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. *Science Education* 91(1): 36–74.
- Hwse P and Holt A (2011) Joining in the enterprise of response in the wake of the NSF data management planning requirement. *Research Library Issues* 274: 11–17.
- Indiana University Center for Postsecondary Research (n.d.) Definitions. In: *The Carnegie Classification of Institutions of Higher Education*. 2015 edn. Available at: http://carnegieclassifications.iu.edu/classification_descriptions/basic.php.
- Keralis SDC (2012) Data curation education: A snapshot. In: *The Problem of Data*. CLIR Publication No. 154. Washington, DC: Council on Library and Information

- Resources, pp. 32–43. Available at: <https://www.clir.org/pubs/reports/pub154/pub154.pdf> (accessed 6 April 2016).
- Kuh GD, Schneider CG and Association of American Colleges and Universities (2008) *High-impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter*. Report, Washington, DC: Association of American Colleges and Universities.
- Lamar Soutter Library, University of Massachusetts Medical School (ed.) (n.d.) *New England Collaborative Data Management Curriculum*. Available at: <http://library.umassmed.edu/necdmc/index> (accessed 30 March 2016).
- MacMillan D (2010) Sequencing genetics information: Integrating data into information literacy for undergraduate biology students. *Issues in Science & Technology Librarianship*. Available at: <http://www.istl.org/10-spring/refereed3.html> (accessed 11 May 2015).
- Mooney H, Collie WA, Nicholson S, et al. (2014) Collaborative approaches to undergraduate research training: Information literacy and data management. *Advances in Social Work* 15(2): 368–389.
- Piorun M, Kafel D, Leger-Hornby T, et al. (2012) Teaching research data management: An undergraduate/graduate curriculum. *Journal of eScience Librarianship* 1(1). Available at: <http://escholarship.umassmed.edu/jeslib/vol1/iss1/8> (accessed 11 May 2015).
- Pryor G and Donnelly M (2009) Skilling up to do data: Whose role, whose responsibility, whose career? *International Journal of Digital Curation* 4(2): 158–170.
- Qin J and D'Ignazio J (2010) Lessons learned from a two-year experience in science data literacy education. In: *International Association of Scientific and Technological University Libraries, 31st annual conference*. Available at: <http://docs.lib.purdue.edu/iatul2010/conf/day2/5> (accessed 30 March 2016).
- Reisner BA, Vaughan KTL and Shorish YL (2014) Making data management accessible in the undergraduate chemistry curriculum. *Journal of Chemical Education* 91(11): 1943–1946.
- Research Data Netherlands (2016) *Essentials 4 Data Support*. Available at: <http://datasupport.researchdata.nl/en/> (accessed 6 April 2016).
- Rice R (2014) Research data MANTRA: A labour of love. *Journal of eScience Librarianship* 3(1). Available at: <http://escholarship.umassmed.edu/jeslib/vol3/iss1/4> (accessed 17 April 2016).
- Rowlett RS, Blockus L and Larson S (2012) Characteristics of excellence in undergraduate research (COEUR). In: Hensley N (ed.) *Characteristics of Excellence in Undergraduate Research (COEUR)*. Washington, DC: Council on Undergraduate Research, pp. 2–19. Available at: http://www.cur.org/assets/1/23/COEUR_final.pdf (accessed 31 March 2016).
- Scaramozzino JM, Ramírez ML and McGaughey KJ (2012) A study of faculty data curation behaviors and attitudes at a teaching-centered university. *College & Research Libraries* 73(4): 349–365.
- Shadbolt A, Konstantelos L, Lyon L, et al. (2014) Delivering innovative RDM training: The immersive informatics pilot programme. *International Journal of Digital Curation* 9(1): 313–323.
- Shorish Y (2012) Data curation is for everyone! The case for Master's and Baccalaureate institutional engagement with data curation. *Journal of Web Librarianship* 6(4): 263–273.
- Shorish Y (2015) Data information literacy and undergraduates: A critical competency. *College & Undergraduate Libraries* 22(1): 97–106.
- Soehner C, Steeves C and Ward J (2010) *E-Science and Data Support Services: A Study of ARL Member Institutions. Report*. Washington, DC: Association of Research Libraries. Available at: <http://www.arl.org/storage/documents/publications/escience-report-2010.pdf> (accessed 31 March 2016).
- Stamatoplos A (2009) The role of academic libraries in mentored undergraduate research: A model of engagement in the academic community. *College & Research Libraries* 70(3): 235–249.
- Stamatoplos A, Neville T and Henry D (2016) Analyzing the data management environment in a Master's-level institution. *Journal of Academic Librarianship* 42(2): 154–160.
- Stephenson E and Caravello PS (2007) Incorporating data literacy into undergraduate information literacy programs in the social sciences: A pilot project. *Reference Services Review* 35(4): 525–540.
- Strasser CA and Hampton SE (2012) The fractured lab notebook: Undergraduates and ecological data management training in the United States. *Ecosphere* 3(12): 1–18.
- Tenopir C, Birch B and Allard S (2012) *Academic Libraries and Research Data Services: Current Practices and Plans for the Future. Report*. Association of College & Research Libraries, June. Available at: http://www.ala.org/acrl/sites/ala.org/acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf (accessed 31 March 2016).
- Tenopir C, Sandusky RJ, Allard S, et al. (2013) Academic librarians and research data services: Preparation and attitudes. *IFLA Journal* 39(1): 70–78.
- Tenopir C, Sandusky RJ, Allard S, et al. (2014) Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research* 36(2): 84–90.
- Toups M and Hughes M (2013) When data curation isn't: A redefinition for liberal arts universities. *Journal of Library Administration* 53(4): 223–233.
- Verbakel E, Noordegraaf M, de Smaele M, et al. (2013) Data-intelligence training for library staff. In: *Framing the digital curation curriculum, International conference*, Florence, Italy, 6–7 May 2013, CEUR Workshop Proceedings, paper 13. Available at: <http://ceur-ws.org/Vol-1016/paper13.pdf> (accessed 6 April 2016).
- Walton G (2010) Data curation and the academic library. *New Review of Academic Librarianship* 16(1): 1–3.

Watkins W, Hamilton E, Boyko E, et al. (2004) Creating a national peer-to-peer training program for data librarians in Canada. *IASSIST Quarterly* 28(2/3). Available at: http://www.iassistdata.org/sites/default/files/iqvol282_3watkins.pdf (accessed 17 April 2016).

Zilinski LD, Nelson MS and Van Epps AS (2014) Developing professional skills in STEM students: Data information literacy. *Issues in Science and Technology Librarianship* 77. Available at: <http://dx.doi.org/10.5062/F42V2D2Z> (accessed 30 March 2016).

Author biographies

Ryan Clement has been the Data Services Librarian at Middlebury College, in Middlebury, VT, since 2015. Prior to this, he was the Data Services Librarian at Reed College in Portland, OR. He holds an MSc in Information from the University of Michigan School of Information, as well as an MA in Philosophy from Temple University. He currently serves on the American Psychological Association Library Advisory Board. He has published work in *Internet Reference Services Quarterly* and the recent book from the Association of College and Research Libraries (ACRL), *Databrarianship: The Academic Data Librarian in Theory and Practice*. He has presented on data services and instruction at liberal arts colleges at the ACRL Annual Conference, the Digital Library Federation Forum, the Research Data Access & Preservation Summit, and the Oregon Library Association Annual Conference. His research interests include data services planning and outreach, data/information literacy, library assessment, and user experience/service design.

Amy Blau has been Instructional and Data Services Librarian at Whitman College in Walla Walla, Washington since 2014. She holds an MSc in Information from the University of Michigan School of Information and a PhD in Comparative Literature from the University of Illinois, Urbana-Champaign. Her professional and research interests include data services planning and outreach, information literacy, digital preservation, and the digital humanities.

Eli Gandour-Rood is the Science Liaison Librarian at the University of Puget Sound, a private liberal arts college in Tacoma, WA. He received his MLIS from the University of Washington Information School in 2013. Currently, he serves as the Vice-President/President-Elect of the Washington State Chapter of the Association of College and Research Libraries.

Parvaneh Abbaspour is the Science & Data Services Librarian at Lewis & Clark College in Portland, Oregon. She received her MLIS from McGill University in 2013.

Appendix A: Assessment questions

Pre-workshop survey, librarians

1. Prior to signing up for this workshop, how familiar were you with the data researchers at your institution were working on? (1 = Very Unfamiliar, 5 = Very Familiar)
2. How prepared do you feel to talk to faculty about their research data? (1 = Very Unprepared, 5 = Very Prepared)
3. How prepared do you feel to describe best practices in research data management to faculty or students? (1 = Very Unprepared, 5 = Very Prepared)
4. Please briefly describe or list what actions constitute effective data management.
5. How prepared do you feel to explain to other librarians what research data is? (1 = Very Unprepared, 5 = Very Prepared)
6. How do you define research data when talking to other librarians?
7. How prepared do you feel to support researchers in finding and applying discipline-appropriate data management approaches/principles to their research? (1 = Very Unprepared, 5 = Very Prepared)
8. How would you rate the current importance of good data management at your institution? (1 = Very Unimportant, 5 = Very Important)
9. How prepared do you feel to explain the research data lifecycle at your institution? (1 = Very Unprepared, 5 = Very Prepared)
10. Please briefly describe the research data lifecycle at your institution. What areas does your library currently support?
11. What areas of data management are you prepared to support for your researchers? What other departments on campus might you reach out to to assist in data management?
12. What is your institution? (coded as alpha character)

Pre-workshop survey, faculty/students/staff

1. How prepared do you feel to describe best practices in research data management? (1 = Very Unprepared, 5 = Very Prepared)
2. Please briefly describe or list what actions constitute effective data management.
3. How prepared do you feel to describe the research data lifecycle? (1 = Very Unprepared, 5 = Very Prepared)
4. Please briefly describe the research data lifecycle.

5. How prepared do you feel to find and apply discipline-appropriate data management approaches/principles to your research project? (1 = Very Unprepared, 5 = Very Prepared)
6. What do you most hope to learn from this workshop?
7. What motivated you to attend this workshop this summer?
8. What is your institution? (alpha-coded)
9. What is your role?
14. Would you recommend a workshop on this topic to your colleagues? Why or why not?
15. Do you have other suggestions for how NW5C colleges can (individually or collaboratively) develop and improve their support of research data management?
16. What is your institution? (coded as alpha character)

Post-workshop survey, librarians

1. Prior to signing up for this workshop, how familiar were you with the data researchers at your institution? (1 = Very Unfamiliar, 5 = Very Familiar)
2. How prepared do you feel to explain to other librarians what research data is? (1 = Very Unprepared, 5 = Very Prepared)
3. How do you define research data when talking to other librarians?
4. How prepared do you feel to talk to faculty about their research data? (1 = Very Unprepared, 5 = Very Prepared)
5. How prepared do you feel to support researchers in finding and applying discipline-appropriate data management approaches/principles to their research? (1 = Very Unprepared, 5 = Very Prepared)
6. How prepared do you feel to explain the research data lifecycle at your institution? (1 = Very Unprepared, 5 = Very Prepared)
7. Please briefly describe the research data lifecycle at your institution. What areas does your library currently support?
8. What areas of data management are you prepared to support for your researchers? What other departments on campus might you reach out to to assist in data management?
9. How would you rate the current importance of good data management at your institution? (1 = Very Unimportant, 5 = Very Important)
10. Please briefly describe or list what actions constitute effective data management.
11. What stands out as the most important skills or pieces of information you gained from this workshop?
12. What information or activity was missing that you would have liked to cover?
13. Do you feel that this workshop structure was an effective way to learn about this topic? Why or why not?

Post-workshop survey, faculty/students/staff

1. How prepared do you feel to describe best practices in research data management? (1 = Very Unprepared, 5 = Very Prepared)
2. Please briefly describe or list what actions constitute effective data management.
3. How prepared do you feel to describe the research data lifecycle? (1 = Very Unprepared, 5 = Very Prepared)
4. Please briefly describe the research data lifecycle.
5. How prepared do you feel to find and apply discipline-appropriate data management approaches/principles to your research project? (1 = Very Unprepared, 5 = Very Prepared)
6. What stands out as the most important skills or pieces of information you gained from this workshop?
7. What information or activity was missing that you would have liked to cover?
8. Do you feel that this workshop structure was an effective way to learn about this topic? Why or why not?
9. What, if anything, from this workshop would you most like to see shared with peers and colleagues at your institution?
10. Do you have other suggestions for how NW5C colleges can (individually or collaboratively) develop and improve their support of research data management?
11. Other comments
12. What is your institution? (alpha coded)
13. What is your role?

Follow-up survey, faculty/students/staff

1. What is your institution? (alpha coded)
2. What is your status?
3. How prepared do you feel to find and apply discipline-appropriate data management approaches/principles to your research project?
4. How prepared do you feel to describe best practices in research data management?
5. How would you describe effective research data management to a new research assistant?

who joined your laboratory/research group this year? (2-3 sentences)

6. Describe how you used something you learned in the workshop in the last 8 months. (1-3 sentences)
7. Describe any issues or challenges you have encountered in the past 8 months with respect to the storage, organization, or sharing of your research data that you wish had been covered in the workshop.
8. How do you think research data management skills/practices/concepts might be incorporated into regular classes or labs in the curriculum?