JESLIB 2014; 3(1): 86-99
doi:10.7191/jeslib.2014.1064

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

Initiating Data Management Instruction to Graduate Students at the University of Houston Using the New England Collaborative Data Management Curriculum

Christie Peters and Porcia Vaughn

University of Houston, Houston, TX, USA

Abstract

The need for graduate-level instruction on data management best practices across disciplines is a theme that has emerged from two campus-wide data management needs assessments that have been conducted at the University of Houston (UH) Libraries since 2010. Graduate students are assigned numerous data management responsibilities over the course of their academic careers, but rarely receive formal training in this area. To address this need, the UH Libraries offered a workshop entitled Research Data Management 101 in April, 2014, and all graduate and professional students on campus were invited to attend. The New England Collaborative Data Management Curriculum (NECDMC) served as the basis for the workshop, and two general sessions were planned. A research group in the College of Natural Sciences & Mathematics requested a special session after advertisements for the workshop were distributed. 105 individuals registered for the event, 65 signed into the workshop, and 63 completed the end-of-workshop assessment. The results from this assessment, general lessons learned, and plans for future sessions will be discussed.

Introduction

The need for graduate instruction on data management best practices across disciplines on the UH campus is a theme that has emerged from two campus-wide data management needs assessments conducted at the UH Libraries since 2010. Faculty in science and engineering fields who were awarded large NSF or NIH grants in fiscal year 2010 were invited to participate in the first assessment, which explored general data management practices of principal investigators working on federally funded research just prior to the role out of the NSF data management plan (DMP) mandate in January 2011 (Peters and Dryden 2011). In 2013, the Libraries conducted a second interdisciplinary assessment modeled on Purdue's Data Curation Profile Toolkit and not dependent upon funding agency (http://datacurationprofiles.org/). Thirty researchers across 7 colleges (College of Liberal Arts & Social Sciences (CLASS), Honor's College, Architecture, Engineering, Natural Sciences & Mathematics (NSM), Pharmacy, and Technology) and 20 departments were interviewed for one or both of these two studies, which reveal that graduate students are rarely taught all of the competencies that are necessary to properly manage research data even though they are expected to assume many data management responsibilities over the course of their academic career. When this type of instruction is in place, it tends to be specific to a particular

Correspondence to Christie Peters: cpeters@uh.edu
Keywords: data management training, library instruction, NECDMC, New England Collaborative Data Management Curriculum, graduate students, assessment, instruction

area of research and focused on limited student responsibilities. Interviews with faculty at other institutions indicate that many feel they lack the experience or knowledge necessary to teach students data-information literacy competencies (Carlson et al. 2013). Given the current pervasiveness of data-driven research, this limited and ad hoc way of approaching data management instruction is a disservice to both the student and research communities.

Data services for students and faculty in the social sciences have existed in research libraries for decades, but it was the rise of computational research in the sciences and engineering -- and the data deluge that followed -- that led to the development of research data management services, defined here as the storage, curation, preservation, and provision for continuing access to digital research data (Hey and Trefethen 2003, Lewis 2010). Computational research in the social sciences has developed more slowly, although it is beginning to make progress, due in no small part to access and privacy restrictions that are inherent in social science research and the infrastructure requirements of distributed monitoring, permission seeking, and encryption (Lazer et al. 2009). Digital scholarship is still emergent in the humanities, but the increasing availability of various materials in digital format and the use of a variety of data analytics are enabling humanists to interrogate sources in new ways (Borgman 2009). The American Council of Learned Societies recognizes the need in the humanities and social sciences for infrastructure similar to the cyberinfrastructure utilized in the sciences, but one developed more specifically for the research needs of scholars in those fields (American Council of Learned Societies 2006). When data is defined simply as the output of any systematic investigation that results in the production of new knowledge, it is clear that scientists, social scientists, and humanists all 'do data' and will benefit from the development of research data management services (Pryor 2012).

The dangers inherent in conducting research without understanding what proper data management entails are many. Mismanagement of data over the lifecycle of a project can result in questions of research accuracy, reliability, integrity, and security. Access becomes an issue if data is not properly described, which then becomes a compliance issue. Only a concerted effort to educate current and future researchers to adopt better practices will alter the inconsistent data management practices that plague research across disciplines (Association of Research Libraries 2006). If these efforts are not undertaken or if they fail, the continued development of e-Research, defined here as "the use of digital tools and data for the distributed and collaborative production of knowledge," will be hindered by a lack of infrastructure, standardized processes, and personnel trained in the management and curation of research data (Carlson et al. 2011, Meyer and Schroeder 2009).

The scenario of graduate students who are insufficiently trained in data management best practices is not unique to the University of Houston. There are currently no widely accepted instructional standards for data management, and there appears to be no concerted effort across institutions to educate graduate students about data management best practices before allowing them to embark upon their graduate research. Libraries are well situated to help address this problem, although the traditional model of structuring and staffing research libraries around disciplines might complicate the development of data-related instructional services that are necessarily interdisciplinary in nature (Association of Research Libraries 2007). Anna Gold suggests ways that that librarians can position themselves as partners in research by playing a more "upstream" role in data science, but she refers specifically to direct involvement in the creation of data curation prototypes and support for the use of documentation, practices, or standards that will assure the longevity of the data downstream (Gold 2007). Providing

instruction to future researchers about data management best practices is arguably just as important an upstream role in data science, even if it is one step removed from actual collaboration.

Library-led data management instruction, which focuses on best practices across the entire data lifecycle, has much to offer e-Research and the campus research community. Liaison librarians who are very knowledgeable about the research needs of the faculty and graduate students they serve are well situated to put data management best practices into a disciplinary context that researchers understand by combining the comprehensive data management expertise that researchers often lack with the domain-specific knowledge that drives their research, both of which are necessary for the data curation required for e-Research (Gabridge 2010, Tenopir, Birch, and Allard 2012, Jahnke, Asher, and Keralis 2012, Garritano and Carlson 2009). The resulting instruction contributes to a more data-literate research community and prepares researchers to engage in the sound data curation practices that e-Research entails, while simultaneously educating the campus community about the data management and curation expertise that exists within the library. On a research university campus where the pressure to secure research funding from agencies with increasingly stringent data management requirements is at an all-time high and funding at an all-time low, the importance of having a data literate research community cannot be overstated.

The library also stands to gain from the development of data-related instructional services. A 2010 Association of College and Research Libraries report on the value of academic libraries states that academic libraries should align themselves with the mission of their institution (Oakleaf 2010). The UH mission statement includes goals to become a nationally competitive public research university and to create an environment in which student success can be ensured (http://www.uh.edu/about/mission/goals/). To align with these goals, the UH Libraries' 2013-2016 Strategic Directions includes the directive 'target specific user groups with customized services and niche collections' (University of Houston Libraries 2013). Recommended strategies for achieving this goal include expanding library services to graduate students and enhancing faculty research support. Data management instruction benefits graduate students by providing them with the information that they need to effectively manage the research data associated with their theses and dissertations, and it helps faculty increase their research efficiency and the strength of their grant proposals, which in turn contributes to the national competitiveness of the university as a whole. Library administrators can leverage this significant contribution to the university mission to argue the benefits of the research library to campus administrators and to advocate for campus collaborations with other units that offer related services, such as the Office of Sponsored Research and campus IT. Establishing collaborations around research data management has been challenging for many libraries, but such collaborations are essential for the development of truly comprehensive data management services on the research university campus (Verbaan and Cox 2014).

A number of instructional models were considered when the UH Libraries decided to offer a data management workshop for graduate students. In 2010, the University of Minnesota Libraries began offering workshops specifically aimed at the creation of NSF data management plans (Johnston, Lafferty, and Petsan 2012). While this approach has obvious relevance for students who plan on undertaking grant funded research, we felt that this type of workshop would be too limited in scope and might alienate students working on research that is not funded by NSF. Librarians at Purdue University, the University of Minnesota, and the University of Oregon collaborated on the Data Information Literacy (DIL) project,

which aims to develop educational interventions to meet identified data-related educational needs of graduate students in disparate disciplines (Carlson et al. 2013). This will undoubtedly revolutionize embedded and targeted data management instruction, but it is not the best solution when developing stand-alone workshops aimed at a diverse, interdisciplinary group of students. We know there is a need for data management instruction at the University of Houston, but we do not know the extent of need among our faculty and students. We felt it important to find a curriculum that we can modify to fit a diverse targeted audience and assess for the development of future data management services and instruction.

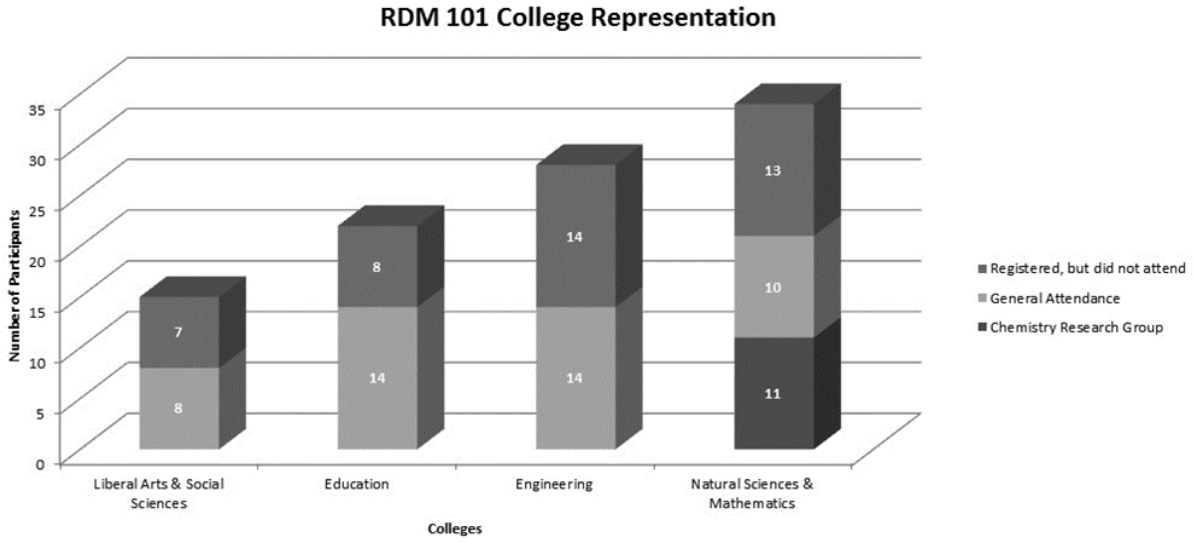
The Lamar Soutter Library at the University of Massachusetts Medical School and collaborators developed the New England Collaborative Data Management Curriculum (NECDMC) as an instructional tool to teach data management best practices to undergraduates, graduate students, and researchers in the health sciences, sciences, and engineering disciplines (http://library.umassmed.edu/necdmc/index). While students across disciplines at the University of Houston were invited to attend RDM 101, the instructors (both science librarians) believed that the majority of participants would come from STEM fields. The curriculum's focus on the data lifecycle, its scalability, and the ease with which it can be modified were among the reasons that the NECDMC was chosen over other curricula as the basis for this workshop.

Methods

The NECDMC curriculum is comprised of seven modules that can be used individually or in conjunction with one another, including: 1) overview of research data management; 2) types, formats, and stages of data; 3) contextual details needed to make data meaningful; 4) data storage, backup, and security; 5) legal and ethical considerations for research data; 6) data sharing and reuse policies; and 7) archiving and preservation. The lesson plan for RDM 101 included a one-hour lecture based on module 1 of the NECDMC and a hands-on activity using the NECDMC research case *Combining data from 10 years of research for retrospective studies on the effects of exercise and diet on the risk of diabetes*. For reasons that will be discussed below, we replaced this research case in the second RDM 101 session with the mini-case *Identifying Data Types and Stages of Data* that is located with the materials for module 2, and we dropped the activity altogether in the third session. We chose not use the 53-slide Powerpoint that accompanies module 1 because we thought non-science participants might find the heavily science-oriented and text-based slides off-putting and using so many slides is not conducive to discussion. We supplemented the curriculum with information from other modules and external sources when deemed necessary. For example, we used the YouTube video *Data Sharing and Management Snafu in 3 Short Acts* -- which was developed by librarians at the NYU Health Sciences Library -- to set the stage for the workshop, and it was very well received (http://youtu.be/N2zK3sAtr-4).

The stated objectives of module 1 include: 1) recognize what research data is and what data management entails; 2) recognize why managing data is important for your research career; 3) identify common data management issues; 4) learn best practices and resources for managing these issues; and 5) learn about how the library can help you identify data management resources, tools, and best practices. In an effort to keep the objectives manageable for a 1.5 hour workshop and suitable for a general audience, they were narrowed down to 1) recognize what research data is and what data management entails; 2) describe current issues within data management; and 3) identify resources, tools, and services related to data management, all in order to develop and apply data management best practices to one's own research.

Figure 1: 86% of RDM 101 registrants and 88% of participants came from four colleges



Participants registered for the workshop session by using a web form linked to the library website, and they signed into the workshop using a Survey Monkey form that was embedded in the Data Management Research Guide (LibGuide). Both forms asked for participant name, email address, college, and department, with the sign-in form additionally asking for advisor name and if the student's advisor recommended or required that they attend the workshop. Participants responded to a 17 question assessment administered using Survey Monkey at the conclusion of the workshop (Appendix). This assessment was based largely, but not exclusively, upon the assessment that accompanies NECDMC module 1. It gauged participant satisfaction with the workshop, the nature of data-related workshops and services that students would like to see in the future, and the likelihood of participation in future data management workshops. We used Survey Monkey because it has statistical and collaborative features that accommodate the mixed-method survey approach used in the assessment, which included qualitative and quantitative data that was analyzed through counts and frequencies.

A number of methods were used to market RDM 101. An electronic flyer for the event was distributed to colleges and departments by liasions, uploaded to the library's digital signage, pushed twice to the graduate and professional student listserv by the University's newly established Graduate School, and linked to the rotating image gallery on the library website's homepage with a link to the registration page. Personal invitations were also sent to all researchers who participated in one of the campus-wide data management needs assessments mentioned above inviting them to encourage members of their research group to attend one of the workshops.

Results

Demographics. The number of students (and faculty) who registered for RDM 101 surpassed our expectations. A total of 105 individuals registered for one of the two general sessions, and a Chemistry faculty member requested a dedicated session for 10 members of his research group. The most effective marketing strategy was having the Graduate School push workshop flyers to the graduate and professional student listserv. The vast majority of registrations occurred within 24 hours of each Graduate School

Table 1: RDM 101 participation by college and department

COLLEGES	DEPARTMENTS	PARTICIPANTS
Business	Marketing	1
Education	Curriculum & Instruction Counseling Psychology Educational Psychology	2 2 10
Engineering	Chemical & Biomolecular Civil & Environmental Electrical & Computer Mechanical Petroleum	1 7 4 1 1
Hotel & Restaurant Management	N/A	1
Law Center	N/A	1
Liberal Arts & Social Sciences	English Health & Human Performance Political Science Psychology	1 4 1 2
Natural Sciences & Mathematics	Biology & Biochemistry Chemistry Computer Science Earth & Atmospheric Sciences	4 11 (Research Group) 1 5
Optometry	N/A	1
Pharmacy	N/A	1
Technology	Mechanical Engineering Tech. Network Engineering Communication	1 1
Other	Baylor College of Medicine	1

push. A total of 65 individuals signed into one of the three sessions, 30 (46%) of whom claimed that they were asked to attend by their advisor. Of these, 16 (25% of the total) were the advisees of one of two researchers who had been interviewed for one or both of the campus-wide data management needs assessments. A number of others were asked to attend by faculty at the recommendation of a subject liaison.

While RDM 101 was marketed to graduate students across disciplines, 90 of the 105 registrants (86%) and 57 of the 65 participants (88%) came from just four of the twelve academic colleges on the UH campus (Figure 1). Of the participants, 12% came from the College of Liberal Arts & Social Sciences (CLASS), 22% from the College of Education, 22% from Cullen College of Engineering, and 32% from the College of Natural Sciences & Mathematics (NSM).

A close examination of the departmental data reveals that 68% of RDM 101 participants are in science or engineering-related disci-

Table 2: Average Likert ratings for Q2-Q9

Question	Session 1	Session 2	Session 3	Average Rating
Q2	3.79	4.35	4.40	**4.1**
Q3	4.24	4.43	4.60	**4.37**
Q4	4.24	4.14	4.30	**4.21**
Q5	3.55	3.74	4.50	**3.77**
Q6	4.03	4.39	4.50	**4.24**
Q7	3.17	3.65	4.10	**3.5**
Q8	3.62	3.96	4.20	**3.84**
Q9	3.43	3.85	4.67	**3.7**

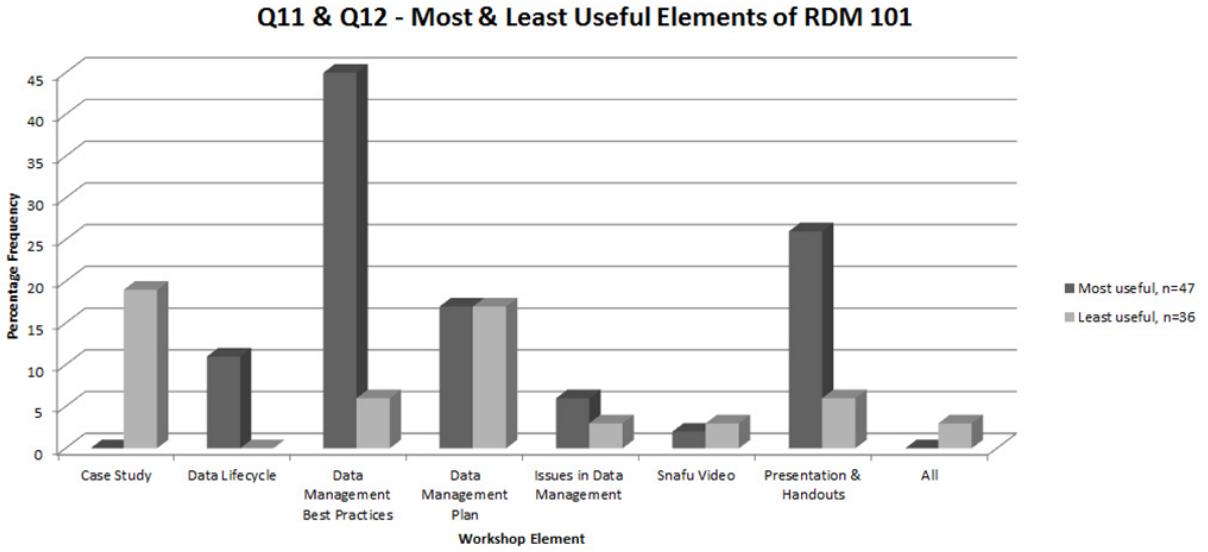
plines and 31% in social science-related disciplines (Table 1). There was only one participant, a graduate student in the Department of English, who is in the humanities.

Assessment. The RDM 101 assessment gauged participant satisfaction with the workshop, the nature of data-related workshops and services that students would like to see in the future, and the likelihood of participation in future data management workshops. We allotted 15 minutes at the end of the workshop for the assessment, which effectively took half of the time we allotted for a hands-on activity, but we decided to move forward with both the activity and the assessment in spite of the time crunch because we felt that both were important. In the end, due to the influence that the assessment will have on the development of future workshops and other data-related services, it became our number one priority and the activity was eliminated from the final workshop entirely.

Q2-Q9 asked participants to rank various aspects of the RDM 101 workshop using a Likert scale that ranged from one to five with (1) indicating *not at all well*/ *not at all* and (5) *very well*/*very likely*. For the purpose of analysis, we determined that an average rating of four or above indicates that the respondent is confident in their ability to explain the data management concept addressed in the question, while an average rating under four indicates that the respondent lacks that confidence. Based on these criteria, the overall average rating for four questions (Q5, Q7-Q9) indicates data management concepts covered in the workshop that participants were not confident they could explain at the workshop's conclusion (Table 2).

Q5 asked participants to indicate how well the workshop familiarized them with the data management plan (DMP) requirements used to characterize a plan for the lifecycle of research data. While the average rating for this question was 3.77, 66% of the respondents replied with scores greater than or equal to four. Similarly, when participants were asked if workshop goals met their expectations in Q7, 52% of respondents selected a 4 or higher on our rating scale, a fact that is overshadowed by the average rating of 3.5. These discrepancies could be indicative of differences in prior knowledge about

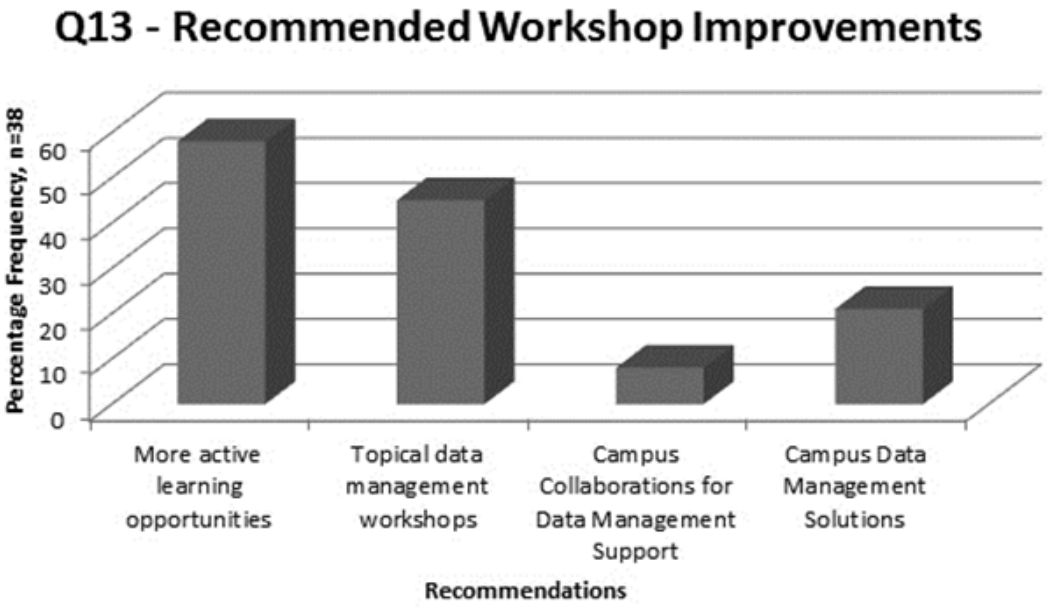
Figure 2: Workshop elements that participants labeled as most and least useful.



the topic across disciplines. If that is the case, it seems to indicate that students with very little knowledge about research data management, i.e. the students we are hoping to impact the most, did not learn enough about the topic during the workshop. Q8 asked participants to rate how useful the presentation portion of the workshop was in regard to their learning needs of research data management concepts. As with the results for Q5 and Q7, the average rating of the presentation was 3.81, but 67% of respondents selected a four or higher on the Likert scale. The results for Q5, Q7, and Q8 indicate a certain level of confidence with the content addressed, but instruction clearly needs to be revisited in these areas.

Q9 asked participants to rank the case study *Combining Data from 10 Years of Research for Retrospective Studies on the Effects of Exercise and Diet on the Risk of Diabetes*. This question remained on the assessment for all three sessions, even though we switched to the mini-case *Identifying Data Types and Stages of Data* in the second session of the workshop and used no activity at all in the session for the research group. We planned to simply discount this question given the change of plans, but were intrigued that the average rating across all sessions was 3.7, higher than one might expect given that it only applies to the first session. When Q9 average ratings are examined for each session, the results are even more interesting. The lowest rating for this question (3.43) occurs in the first session. Unlike Q5, Q7, and Q8, each of which had a significant number of ratings over 4, in spite of an overall average rating less than 4, only 38% of the respondents from this session rated the case study with a 4 or 5. This reflects a level of dissatisfaction with the case study that we did not see in the previous questions. The average rating for Q9 increased in the second session (3.85) even though a different case study was used. One possible explanation for this is that respondents rated the case study that was used, even though it was not the case study specified in the question. If that is the case, the second case study fared better than the first, but still fell short of the 4.0 threshold. It is more difficult to explain why the case study is ranked highest in the last session for the research group (4.67) with 50% of the respondents rating the case study with a 5. Likert ratings in this session were higher across the board, so the

Figure 3: Participant Recommendations for RDM 101 Improvements



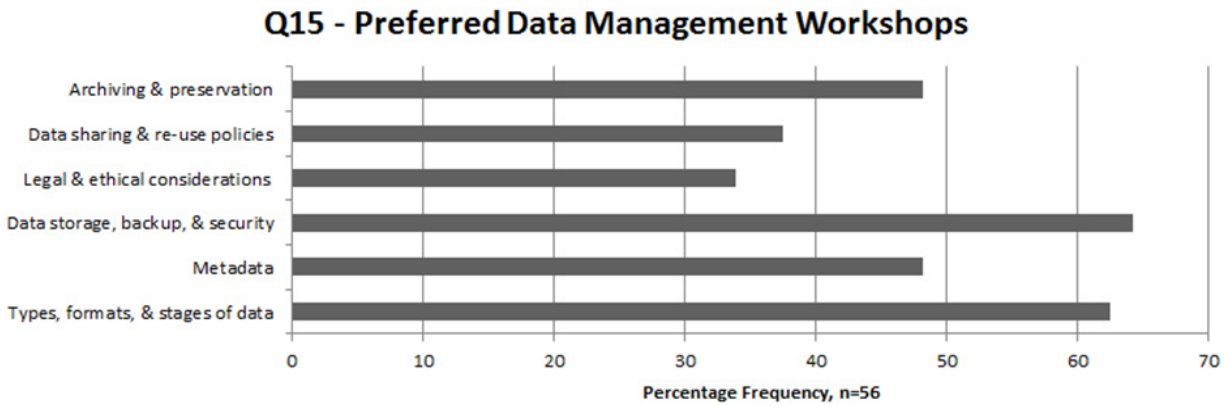
students may have simply been answering positively to everything without giving the questions much thought. If so, this speaks to the benefit of providing targeted data management instruction to small research groups, rather than to large, diverse groups of students.

Q10 inquired about satisfaction with the length of the workshop and how much time participants would be willing to commit to similar workshops. Three quarters of the respondents said that the workshop was *Just about right*, but 49% of those respondents subsequently commented that they would prefer to spend an hour or less of their time in similar workshops. Given the difficulty that we had conveying all of the information we prepared for RDM 101 in an hour and a half, we need to consider the apparent unwillingness of graduate students to attend a workshop that exceeds this length as we develop future workshops.

The assessment included a number of open-ended questions that address individual perceptions about RDM 101. Q11 and Q12 asked participants to point out the elements of the workshop that they found most and least useful (Figure 2).

The following workshop elements were used to code responses: (1) the Snafu video; (2) the data life cycle; (3) data management best practices; (4) issues in data management; (5) general workshop presentation and handouts; (6) data management plans, including the DMP Tool; (7) case study activity; and (8) all. The "all" category reflects responses that mentioned every element individually or responded "all of it" or "everything." Comments that were not relevant to the question were not coded or included in the analysis. Data management best practices (45%) and the general workshop presentation and handouts (26%) were considered the most useful elements of the workshop, while the case study (19%) and information on data management plans (17%) were considered to be the least useful. Interestingly, the same number of participants rated information about data management plans the most useful and the least useful aspects of the workshop demonstrat-

Figure 4: Participant preferences for topical data management workshops



ing that a one size fits all approach to data management workshops should not be the only solution on any campus.

Q13 asked participants to recommend improvements that they feel would help them better understand the various research data management concepts (Figure 3).

The most prominent categories include more active learning opportunities (58%), topical data management workshops (45%), and information on campus collaborations and data management solution options (29%). Participants suggested that they be allowed to work with their own research data or another actual data set to provide a real-world application of the concepts being taught. The request for more active learning activities indicates that the learning by listening pedagogy is not the best approach when teaching data management best practices. It is important for students to be able to apply what they are learning while they are learning these concepts in this workshop.

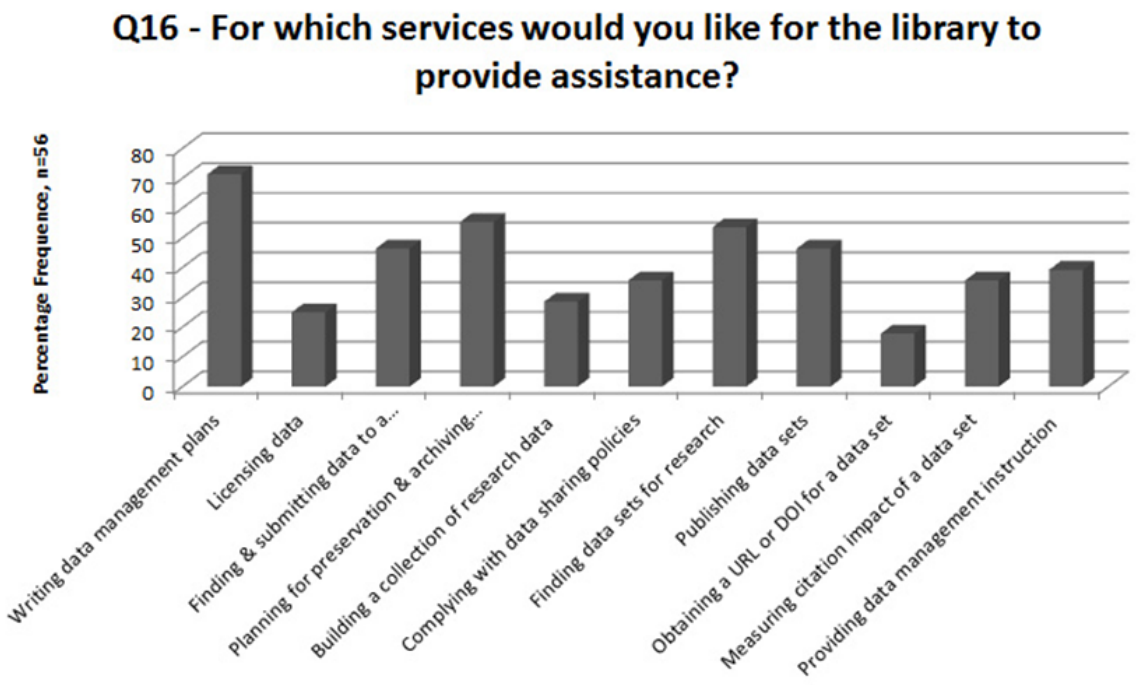
A number of responses to Q13 indicate that the workshop was too general to be useful and that a topical workshop would better meet the students' expectations. In anticipation of this need, Q15 asks participants to select data management-related workshops that they would be interested in attending (Figure 4).

The most highly sought after workshop is data storage, backup, and security (68%), an observation that was reinforced by a high number of questions about the storage, backup, and security solutions that are available both on and off campus. A workshop on types, formats, and stages of data (63%) follows closely behind, and there is interest in archiving & preservation (48%) and metadata (48%), but less interest in data sharing & re-use policies (38%) and legal and ethical considerations (34%). The last two recommendations for the workshop include more active collaborations between the library and other units on campus that provide data management support and services (8%) and more information on campus-wide data management solutions (21%).

Q16 requested that participants select data-related services that they are interested in the library or some other unit on campus providing (Figure 5).

Support for writing data management plans, which includes interest in DMPTool, is the service of most interest to respondents (72%). This is followed by planning for preservation and archiving (55%), and assistance finding data sets for research (54%). Finding and submitting data to a repository and publishing data sets are of equal interest to participants (46%). Assistance obtaining a URL or DOI for a data set is the service of

Figure 5: Participant preferences for data-related services



least interest to participants (18%). The services that are not highly sought after, such as licensing data, were not covered in great detail in the workshop. This could reflect a need to cover these issues in greater detail in future workshops.

It would be interesting to examine the types of workshops and services that students and faculty in particular fields of study are interested in. Unfortunately, we collected demographic information separately from the assessment in an effort to elicit more truthful responses from participants. Future assessments will include at least some non-identifying demographic questions.

Discussion

Close examination of the assessment data and our own perceptions of the workshop reveal that while we did not cover all of the material found within module 1 of the NEC-DMC, we tried to cover too much information in a single workshop. We chose the research case *Combining Data from 10 Years of Research for Retrospective Studies on the Effects of Exercise and Diet on the Risk of Diabetes* for our hands-on activity (even though it is a science-based case study) because we felt that the lessons conveyed by this particular example would resonate with students across disciplines. We were wrong. The complex nature of the case study, the short duration of the workshop, and our determination to assess the workshop at its conclusion made it impossible to successfully work through this activity in the time allotted. We selected the mini-case *Identifying Data Types and Stages of Data* as the activity for the second session of the workshop because it deals with a topic that participants in the first workshop were particularly interested in, and it is less complicated than the first case study. This activity worked slightly better than the original, but we still did not have enough time for participants to complete and reflect on the activity with guided discussion. For this reason, we decided to eliminate the activity altogether in

the last session of the workshop. Moving forward, we need to be more thoughtful about the amount of content we cover in each workshop and the instructional methods we select to accomplish learning outcomes.

We knew it would be a challenge to provide an overview of an unwieldy topic like research data management in a brief one-shot instruction session, but we were not willing to commit to a longer workshop or to a workshop series without first determining the need and interest for this type of instruction. The high level of participation in the first RDM 101 workshop and the responses that we obtained through the assessment demonstrate that students and faculty do, in fact, recognize the need for graduate-level instruction on data management best practices. Finding pedagogical approaches that facilitate active learning and allow participants to apply the data management competencies that we teach is an area where we will continue to improve.

We were not surprised that well over 50% of our participants came from the sciences and engineering, but we were surprised by the high level of participation from students and faculty in the social sciences, especially from the College of Education. The College of Liberal Arts and Social Sciences (CLASS) is the largest academic college at the University of Houston, housing 16 individual departments, yet it only ranked as one of the four most represented colleges because of participation from researchers within the science-oriented Department of Health and Human Performance (Figure 1). It is interesting that there appears to be a high level of interest for data-related services in the College of Education, but very little in data-driven CLASS departments like Political Science and Psychology. This suggests that we need to reach out to researchers in these departments to see if their data needs are being met elsewhere or if they simply are not aware of the services we are offering. We were not surprised that we had low participation from students in the humanities, but engagement in digital humanities is on the rise at UH. We will work with liaison librarians who support faculty and students in these departments to identify relevant data management needs and services.

As a result of the success (and failures) of RDM 101, a data instruction team has been assembled within the library and development of a series of workshops for spring 2015 is underway. This team is comprised of individuals who have expertise with general data management best practices and/or various aspects of the data lifecycle, i.e. metadata, archiving and preservation, data backup, and storage. In hindsight, we should have worked more closely with the Library's instruction team when developing RDM 101 to ensure sound pedagogical practice, so the data instruction team will meet with individuals from this unit to talk about 1) the mission of the group; 2) possible learning objectives for the workshops that we envision for spring, which currently includes a new and improved version of RDM 101, how to write a DMP, and describing, preserving and storing research data; and 3) assessment of these workshops. The team will also work on supplemental data management resources that address specific data management topics that can be distributed as the need arises.

Based on our first experience with RDM 101, we feel that the NECDMC requires significant modification to be suitable for a workshop aimed at a general audience that includes students outside of the sciences and engineering. It was more difficult than anticipated to modify the science-based content of NECDMC module 1, which contains too much content to cover in a single workshop, and the majority of the research cases involve science and engineering-based scenarios. It is unlikely that we will ever use the NECDMC as an out-of-the-box solution, but it is certainly a fantastic resource. Evidence from the RDM 101 assessment indicates that students are interested in activities that

are relevant to their research. Fortunately, a faculty member in the Department of Educational Psychology is interested in collaborating on the development of a research case study specifically aimed at researchers in her field. This will give us the opportunity to test this theory, and it is an exciting opportunity for the library to develop targeted data-management instruction for students in a department where there is demonstrated interest and need. It remains to be seen how we can develop an activity that will allow participants to use their own data or another actual data set, especially when working with students across disciplines, but the idea warrants consideration.

Perhaps the greatest achievement of the NECDMC is its attempt to standardize instruction around the unwieldy topic of research data management. It is our hope that pilots of the NECDMC like the one carried out by the UH Libraries will encourage the Lamar Soutter Libraries and their collaborators to make modifications to the curriculum that will make it more applicable to researchers across disciplines.

Electronic Content

Appendix
An online supplement to this article can be found at http://escholarship.umassmed.edu/jeslib/vol3/iss1/ under “Appendix”.

References

American Council of Learned Societies. 2006. “Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences.” http://www.acls.org/uploadedFiles/Publications/Programs/Our_Cultural_Commonwealth.pdf

Association of Research Libraries. 2006. To Stand the Test of Time: Long-term Stewardship of Digital Data Sets in Science and Engineering. In *A Report to the National Science Foundation from the ARL Workshop on New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe*. Arlington, VA.

Association of Research Libraries. 2007. Agenda for Developing e-Science in Research Libraries: Final Report and Recommendations to the Scholarly Communications Steering Committee, the Public Policies Affecting Research Libraries Steering Committee, and the Research, Teaching, and Learning Steering Committee.

Borgman, Christine L. "The Digital Future is Now: A Call to Action for the Humanities." *Digital Humanities Quarterly* 3, no. 4 (2009).

Carlson, Jacob, Michael Fosmire, C. C. Miller, and Megan Sapp Nelson. "Determining data information literacy needs: A study of students and research faculty." *portal: Libraries and the Academy* 11, no. 2 (2011):629-657, http://dx.doi.org/10.1353/pla.2011.0022

Carlson, Jake, Lisa Johnston, Brian Westra, and Mason Nichols. "Developing an approach for data management education: A report from the data information literacy project." *The International Journal of Digital Curation* 8, no. 1 (2013):204-217, http://dx.doi.org/10.2218/ijdc.v8i1.254

Gabridge, Tracy A. 2010. Covering the Last Mile: Developing e-Science Data Services with Liaison Librarians. Paper read at International Association of Scientific and Technological University Libraries, 31st Annual Conference, at Purdue.

Garritano, Jeremy R., and Jake R. Carlson. "A subject librarian's guide to collaborating on e-Science projects." *Issues in Science and Technology Librarianship* 57 (2009), http://dx.doi.org/10.5062/F42B8VZ3

Gold, Anna. "Cyberinfrastructure, Data, and Libraries, Part 2." *D-Lib Magazine* 13, no. 9/10 (2007).

Hey, Tony, and Anne Trefethen. "The Data Deluge: An e-Science Perspective." In *Grid Computing: Making the Global Infrastructure a Reality*, edited by F. Berman, G. Fox and T. Hey, 809-824. Chichester: Wiley, 2003.

Jahnke, Lori, Andrew Asher, and Spencer D. C. Keralis. 2012. "The Problem of Data." *Council on Library and Information Resources (CLIR) Report* no. pub. #154.

Johnston, Lisa, Meghan Lafferty, and Beth Petsan. "Training researchers on data management, a scalable, cross-disciplinary approach." *Journal of eScience Librarianship* 1, no. 2 (2012):79-87, http://dx.doi.org/10.7191/jeslib.2012.1012

Lazer, David, Alex (Sandy) Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, and Marshall Van Alstyne. "Life in the network: the coming age of computational social science." *Science* 323, no. 5915 (2009):721-723, http://dx.doi.org/10.1126/science.1167742

Lewis, Martin. "Libraries and the Management of Research Data." In *Envisioning Future Academic Library Services Initiatives, Ideas and Challenges*, edited by Sue McKnight, 145-168. London: Facet Publishing, 2010.

Meyer, Eric T., and Ralph Schroeder. "Untangling the web of e-Research: Towards a sociology of online knowledge." *Journal of Infometrics* 3, no. 3 (2009):246-260, http://dx.doi.org/10.1016/j.joi.2009.03.006

Oakleaf, Megan. 2010. Value of academic libraries: a comprehensive research review and report. Association of College and Research Libraries.

Peters, Christie, and Anita Riley Dryden. "Assessing the academic library's role in campus-wide research data management: A first step at the University of Houston." *Science & Technology Libraries* 30, no. 4 (2011):387-403, http://dx.doi.org/10.1080/0194262X.2011.626340

Pryor, Graham. "Why Manage Research Data?" In *Managing Research Data*, edited by Graham Pryor. UK: Facet Publishing, 2012.

Tenopir, Carol, Ben Birch, and Suzie Allard. 2012. Academic Libraries and Research Data Services: Current Practices and Plans for the Future. edited by ACRL.
University of Houston Libraries. 2013. Strategic Directions 2013-2016.

Verbaan, E., and A. M. Cox. "Occupational Sub-Cultures, Jurisdictional Struggle and Third Space: Theorising Professional Responses to Research Data Management." *The Journal of Academic Librarianship* 40, no. 3-4 (2014):211-219, http://dx.doi.org/10.1016/j.acalib.2014.02.008

Disclosure: The authors report no conflicts of interest.

ISSN 2161-3974