Building the New England Collaborative Data Management Curriculum

Donna Kafel
University of Massachusetts Medical School

Et al.

Let us know how access to this document benefits you.

Repository Citation

Kafel D, Creamer AT, Martin ER. Building the New England Collaborative Data Management Curriculum. *Journal of eScience Librarianship* 2014;3(1): e1066. https://doi.org/10.7191/jeslib.2014.1066. Retrieved from https://escholarship.umassmed.edu/jeslib/vol3/iss1/7

UMass Chan
MEDICAL SCHOOL

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

Building the New England Collaborative Data Management Curriculum

Donna Kafel,[1] Andrew T. Creamer,[2] Elaine R. Martin[1]

[1] University of Massachusetts Medical School, Worcester, MA, USA
[2] Brown University, Providence, RI, USA

Abstract

The need for a curriculum designed for librarians to use for teaching STEM research data management skills to their constituents from diverse STEM disciplines has been widely identified (Qin and D'Ignazio 2010). From 2012-2014, a collaborative group of New England librarians, led by a project team from the University of Massachusetts Medical School, developed lecture notes, presentation slides, assignments, readings, and case studies for teaching research data management. The New England Collabora-tive Data Management Curriculum (NECDMC) is unique in its flexibility; providing subject agnostic instructional materials in a modular format for teaching common data management best practices along with a suite of teaching cases illustrating data management in disciplinary contexts. This article is a follow-up to the "Teaching Research Data Management: An Undergraduate/Graduate Curriculum (Piorun et al. 2012) that was published in the Journal of eScience Librarianship.

Introduction

In the 2010 OCLC report "A Slice of Research Life: Information Support for Research in the United States", Susan Kroll and Rick Forsman concluded that "research findings and products are scattered across institutions, older documents and data are becoming inaccessible, and it is uncertain how today's digital content will be carried forward as information technology evolves. Most universities have not addressed this concern and researchers are unable to do so individually." (Kroll and Forsman 2010). Since this report was published, research libraries have developed new services that facilitate the accessibility, management, and long-term preservation of their institution's locally produced research data. One such service libraries are increasingly providing is instruction on research data management best practices for researchers and students who produce and collect research data at their institutions. Many librarians delegated to teach research data management are new to the concept of data information literacy and are investing significant professional development time learning research data management competencies and skills that science data literacy encompasses. Many of these novice data management instructors voice frustration that they lack the time and expertise to develop comprehensive data management teaching materials that are relevant to the STEM disciplines of their research communities. Librarians who have taught research data management to heterogeneous classes have commented on the challenge of comprehensively addressing discipline-specific data management issues

Correspondence to Donna Kafel: donna.kafel@umassmed.edu

and practices. (Qin and D'Ignazio 2010; Whitmire 2013).

Adding to these challenges is the reality that many STEM librarians have non-STEM educational backgrounds, as Creamer et al. found in a 2009 assessment of the educational backgrounds of New England health sciences and STEM librarians (Creamer 2011). Librarians lacking a science background and who are unfamiliar with scientific research environments (e.g. research personnel, project design, workflows, instrumentation, protocols) are concerned about their competence for delivering contextual research data management instruction to their researchers. The literature asserts that research libraries need reliable information on researchers' data needs and data curation practices to develop services that support outreach and teaching, and that these services need to be developed with the combined efforts of subject liaisons. (Scaramozzino et al. 2012; Gabridge 2009). These findings highlight the need for comprehensive educational tools that librarians can use for teaching data management to researchers in diverse STEM fields. Recognizing this need, three initiatives, DataONE, the Data Information Literacy Project, and the New England e-Science Collaboration each developed research data management curricula. The DataONE Education Modules are geared for environmental sciences and the University of Minnesota's Data Management Course (developed as part of the Data Information Literacy project) is intended for engineering students. The New England Collaborative Data Management Curriculum (NECDMC) stands apart for its subject agnostic approach. Authored by a coalition of STEM, metadata, and repository librarians, NECDMC addresses the need for data management educational materials that librarians can use to teach research data management to diverse STEM disciplines. Composed of seven instructional modules that address key components of the National Science Foundation's data management plan recommendations, the modules include lecture notes, readings, activities, STEM teaching cases and sample data management plans, NECDMC is a flexible tool that librarians can easily customize to meet the research data management learning needs of their particular audiences.

The Frameworks Project

The development of NECDMC was a direct outcome of an earlier collaboration between the University of Massachusetts Medical School (UMMS) and Worcester Polytechnic Institute (WPI). Faculty, researchers, graduate students and librarians from both institutions partnered on an IMLS planning grant to develop frameworks for an openly accessible data management curriculum for undergraduate and graduate students and researchers from a broad range of science, health sciences, and technology fields (Piorun et. al 2012). Collaborators stressed the need to develop teaching materials that were relevant to multiple STEM disciplines. Noting that data, research culture, and workflows varied immensely from one discipline to another, faculty from both institutions suggested that the data management curriculum be case-based so that students in various disciplines could understand data management concepts in familiar contexts. Heeding this suggestion, the project leaders interviewed three faculty members about their projects and developed the initial cases and design template, which served as a model for future case development.

In addition to the cases, the frameworks, including lesson plans for a seven module online curriculum were also developed. The frameworks would serve as the outline for later content development. Following the conclusion of this initial planning grant, the UMMS team of librarians was awarded funding by the National Network of Libraries of Medicine New England Region (NN/LM, NER) to fully develop lectures for the seven instructional modules and teaching cases. The UMMS librarians led a collaboration of New England librarians who authored the

instructional modules and additional teaching cases.

Launching the Curriculum

Launching the curriculum was a multi-phase collaborative endeavor. First, the project team secured additional funding from the National Network of Libraries of Medicine, New England Region (NNLM/NER). Next, the development team was expanded to include additional partners with expertise in metadata, institutional repositories, data management planning, etc. to author the modules. In their discussions with potential partners, the team outlined what the responsibilities of project partners would be: co-authoring curriculum modules; disseminating the curriculum to faculty at their institutions; and delivering the content in classes at their institutions.

Responses to the project team's invitation to partner on content development and to implement the curriculum at local settings revealed a great deal of interest from the corresponding libraries. Several potential partners had been planning to teach research data management to their staff, students, and researchers, and recognized the advantages of participating in a group effort to develop the curriculum. They noted that working with others to develop content that could easily be adapted for their local use would ultimately save them time and effort, and would also leverage contributors' expertise in specific areas of research data management. Libraries from the following institutions accepted the team's initial invitation to partner with the project team on the curriculum: University of Massachusetts Amherst, Tufts, Northeastern, and the Marine Biological Laboratory and Woods Hole Oceanographic Institution (MBLWHOI). Table 1 notes the seven modules and their corresponding partner authors. In addition to taking responsibility for fully developing the content for their modules, some partners contributed additional research cases and examples of data management plans.

From beginning to end, the partners took about nine months to outline their modules, develop the content drafts and revisions, teaching slides, etc. Following this phase, the project coordinators sent the entire curriculum to outside experts (librarians practicing in the areas that the modules covered) for peer review. Upon completion of the review process, the partners submitted revisions and met several times to finalize content. The New England Collaborative Data Management Curriculum (NECDMC) (http://library.umassmed.edu/necdmc/index) was published in November 2013. A Creative Commons Attribution-NonCommercial Share Alike license allows users of the curriculum to customize the full curriculum or parts of it, and modify it to the needs of their students. Along with the lesson plans and modules, the NECDMC site includes research cases, sample data management plans, student activities, instructions on using the curriculum, and a "Join the Collaboration" page that invites librarians to become pilot partners and implement NECDMC at their institutions.

Collaborative Curriculum Development Challenges

As to be expected, developing content via collaboration poses challenges as well as exciting opportunities for richer content development. An initial challenge was locating a suitable partner to develop module 7 in the initial development stage. All originating partners lacked expertise in archiving and preservation. The curriculum was originally launched without module 7 fully developed; however, over the summer of 2013 University of Connecticut and Harvard Countway joined the collaboration and completed module 7. A second challenge noted by partners was the potential for overlap as partners initially developed their modules independently of each other. Some content may be more relevant to one module than another or duplicated in several modules. In the cases where overlap occurred, partners decided on a case-by-case basis during the final review stage how each instance should be ad-

Table 1: NECDMC Module Co-authors

Module	Author(s)
Module 1: Overview of Research Data Management	Regina Raboin, Tufts University Andrew Creamer and Donna Kafel, University of Massachusetts Medical School
Module 2: Types, Formats, and Stages of Data	Jen Ferguson, Northeastern University
Module 3: Contextual Details Needed to Make Data Meaningful to Others	Elizabeth Coburn, John Furfey, and Jen Walton, Marine Biologic Laboratory and Woods Hole Oceanographic Institution Alexander May and Alicia Morris, Tufts University
Module 4: Data Storage, Backup, and Security	MJ Canavan, Steve McGinty, and Rebecca Reznik-Zellen, University of Massachusetts Amherst
Module 5: Legal and Ethical Implications of Research Data	Donna Kafel and Lisa Palmer, University of Massachusetts Medical School Lynne Riley, Worcester Polytechnic Institute
Module 6: Data Sharing and Re-use Policies	Matt Sheridan, Laura Quilter, and Aaron Rubinstein, University of Massachusetts Amherst Elizabeth Coburn, John Furfey, and Jen Walton, Marine Biologic Laboratory and Woods Hole Oceanographic Institution
Module 7: Repositories, Archiving, & Preservation	Introduction: Andrew Creamer, University of Massachusetts Medical School David Lowe, University of Connecticut Darla White and Emily R. Novak Gustainis, Harvard University

dressed and consensus was reached amongst all partners. The final challenge was the time it took to develop and finalize the content. Originally the team thought the modules would be developed in less than six months, but partners were all working librarians with busy schedules, working on the curricular content in addition to their regularly scheduled duties, so there were inevitable delays. Despite these challenges, this collaborative approach to research data management curriculum development proved beneficial. Partners brought a variety of subject skills, teaching skills, and librarianship competencies that no one group could have singly possessed. The project coordinators continue to believe that the collaborative approach, despite its challenges, led to a better final product.

New England Collaborative Data Management Curriculum Train the Trainer Workshops

To accompany the launch of the NECDMC site in November 2013, the project team developed a train-the-trainer class "Teaching Research Data Management with the New England Collaborative Data Management Curriculum" to demonstrate how the curriculum could be used and adapted for teaching research data management to diverse STEM audiences. As the class focused specifically on how to teach with NECDMC and not on data management concepts, its forty-one registrants were required to attend a pre-workshop webinar, "Best Practices for Teaching RDM and Consulting on Data Management Plans" a week prior to the November workshop. This webinar focused on key elements of data management including definitions and examples of data, relevance of well-managed data, components of a data management plan, roles for librarians, and used Tufts University's Tisch Library's Data Services Working Group's model to demonstrate how library research data services can be initiated.

The in-person workshop included back-

ground information on the Frameworks and NECDMC projects, pilot and evaluation plan, demonstration of the NECDMC site, a demonstration of teaching with a research case, a presentation on how to write a research case, and a breakout session during which attendees were assigned activities from the curriculum's modules. It was hoped that by the end of the class, participants would decide to use the NECDMC at their local site and adapt it for their teaching needs.

Responses to attendees' evaluations of the workshop revealed that attendees considered the most useful information to be the review of the components of the NECDMC website, the demonstration of teaching with a research case, and discussions on teaching methods for different audiences and settings. Participants appreciated having a suite of teaching materials to start from and not having to develop content from scratch. A second train-the-trainer class was offered in conjunction with the ACRL New England meeting in May 2014. Although the content remained the same, the instructors emphasized customizing the curriculum and its flexibility.

Pilot Sites

During the two workshops, and on the NECDMC website, the project team invited librarians to formally participate in the implementation of NECDMC by becoming pilot partners. Although anyone can use the curriculum, official pilot partners agreed to: 1) use one or more of the modules to teach RDM at their local setting, 2) use one or more of the research cases, 3) collect evaluations from students and submit them to the project team, 4) participate in group pilot partner conference calls, and 5) participate in follow up interviews with the project team. Pilot partners are also encouraged to enhance the diversity of NECDMC's teaching research cases by submitting new cases based on projects at their local settings.

At the writing of this article, there are currently twelve active official NECDMC pilot partners. Details of the NECDMC project and the curriculum's components and uses have been disseminated broadly through e-mail announcements to library associations, schools, librarians in the e-Science Community of Interest, and at annual library conferences including the 2014 University of Massachusetts and New England Area Librarian e-Science Symposium, at a panel presentation at the 2014 Research Data & Preservation (RDAP) meeting, and the American Library Association annual meeting. The lessons learned through the implementation of NECDMC by the pilot partners will be explored in a future Journal of eScience Librarianship article.

Conclusion

The New England Collaborative Data Management Curriculum provides openly accessible and customizable teaching materials for librarians to use in teaching research data management to students and researchers in diverse STEM disciplines. The NECDMC project and its preceding project, Frameworks for a Data Management Curriculum, exemplify how a group of librarians looking for tools to teach data management, banded together and leveraged its project management skills and subject expertise to develop a unique openly accessible and freely customizable case based curriculum for use by anyone. The NECDMC project team envisions the curriculum as a dynamic entity that will evolve through future content contributions and feedback from its community of users.

References

Carlson, Jacob, 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

Creamer, Andrew, Myrna E. Morales, Donna Kafel, Javier Crespo, and Elaine Russo Martin. "Assessment of health sciences and science and technology librarian e-science educational needs to develop an e-science web portal for librarians." *Journal of the Medical Librarian Association* 99, no. 2 (2011):153-156, http://dx.doi.org/10.3163/1536-5050.99.2.007

Creamer, Andrew. 2012. "Creating an Online Research Data Management Course: A Conversation with Data Librarians Robin Rice and Stuart Macdonald." *The e-Science Portal for New England Librarians*. Accessed August 2014. http://esciencelibrary.umassmed.edu/rice_macdonald_interview.pdf

Gabridge, Tracy. 2009. "The Last Mile: Liaison Roles in Curating Science and Engineering Research Data." *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC,* No. 265:15-21. Accessed September 2014. http://old.arl.org/bm~doc/rli-265-gabridge.pdf

Kroll, Susan, and Rick Forsman. 2010. "A Slice of Research Life: Information Support for Research in the United States." Report commissioned by OCLC Research in support of the RLG Partnership. Accessed September 2014. https://www.oclc.org/content/dam/research/publications/library/2010/2010-15.pdf

Qin, Jian and John D'Ignazio. 2010. "Lessons Learned from a Two Year Experience in Science Data Literacy Education." International Association of Scientific and Technological University Libraries, 31st Annual Conference. Paper 5. Accessed August 2014. http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1009&context=iatul2010

Piorun, Mary E., Donna Kafel, Tracy Leger-Hornby, Siamak Najafi, Elaine R. Martin, Paul Colombo, and Nancy R. LaPelle. "Teaching Research Data Management: An

Undergraduate/Graduate Curriculum." *Journal of eScience Librarianship.* 1, no. 1 (2012):46-50, http://dx.doi.org/10.7191/jeslib.2012.1003

Whitmire, Amanda. 2014. "Implementing a graduate- level data information literacy curriculum at Oregon State University." Poster presented at the University of Massachusetts and New England Area Librarian e-Science Symposium, April 9, 2014. Accessed Dec. 4, 2014. http://escholarship.umassmed.edu/cgi/viewcontent.cgi?article=1072&context=escience_symposium

Acknowledgement

Funding Statement