October 2012

A Sample of Research Data Curation and Management Courses

Andrew T. Creamer
University of Massachusetts Medical School

Et al.

Let us know how access to this document benefits you.

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

A Sample of Research Data Curation and Management Courses

Andrew T. Creamer, Myrna E. Morales, Donna Kafel, Javier Crespo, Elaine R. Martin

University of Massachusetts Medical School, Worcester, MA, USA

Abstract

Objective: This paper identifies a sample of research data curation and management courses available at American Library Association-accredited Library and Information Science (LIS) Programs in North America.

Methods: This sample was identified through a content analysis of LIS program course descriptions and syllabi (N=58). Using a framework of research data management and curation competencies, the team gathered a sample of research data curation and management courses offered between fall 2011 and summer 2012.

Results: Only 13 (22%) of LIS programs currently offer a course focused on the management and curation of research data.

Conclusion: Although the literature supports LIS professionals adopting new roles and engaging in eScience and data management, most LIS data-related programs do not have a separate course solely focused on research data management. More LIS programs will need to adapt their curricula in order to help students and practicing professionals develop the needed competencies in research data curation and management.

Introduction

Research data management is a relatively new field for librarians. It involves managing, preserving, providing, or helping to provide storage and access to large data sets accumulated by researchers. It strives to provide ways for these data to be made easily available to other researchers. This may require traditional library skills such as cataloging and developing search procedures for large data sets, and ways to merge sets in a meaningful way. In 2010, Anna Gold forecasted that "a formal curriculum for training and education is emerging, and positions for professional librarians are being advertised to support data curation programs and ser-

vices." In 2011 Kim et al. (2011) analyzed these curricula and professional employment postings and identified the educational requirements for data librarians and a top-ten list of recommended courses to gain these skills.

Indeed, over the last two years there have been a number of educational initiatives funded by universities and funding agencies to support educating librarians to manage and curate research data. The website of the University of Illinois at Urbana-Champaign (UIUC) iSchool's Data Curation Education Program (DCEP), originally established through a grant from the Institute of Museum and Library Services (IMLS) to de-

velop a data curation curriculum, defines data curation as "the active and on-going management of data through its lifecycle of interest and usefulness to scholarship, science, and education. Data curation activities enable data discovery and retrieval, maintain its quality, add value, and provide for re-use over time, and this new field includes authentication, archiving, management, preservation, retrieval, and representation" (2011). It and a few other innovative LIS programs and iSchools have recognized the need for research data management in library and information science education and have been leaders in offering this type of instruction to interested students and practicing librarians.

There are many reasons why research data management should become a part of the LIS curriculum. Writing about the role of libraries in managing data, Perry et al. (2005) noted: "…boundaries are disappearing among the published literature, research data, research databases, and clinical patient data. As research literature increasingly exists alongside repositories of source evidence, large bodies of data can be used to support individual, clinical, or scientific decisions. These datasets may be incorporated and manipulated into knowledge sources through creating application-focused databases. Opportunities exist for both informaticians and librarians to create, maintain, and develop these integrated information resources." Bryan Heidorn (2011) added urgency to these arguments for librarians embracing data management: "Scientists, other scholars, and all of society are now producing, storing, and disseminating digital data that underpin the aforementioned documents in much larger volumes than the text. The survival of this data is in question since the data are not housed in long-lived institutions such as libraries. This situation threatens the underlying principles of scientific replicability since in many cases data cannot readily be collected again. Libraries are the institutions that could best manage this intellectual output."

Cheek and Bradigan (2010) and Creamer et al. (2011) found low numbers of librarians in their samples currently involved in research data management activities. However, Creamer et al. (2012) in a subsequent survey found that a third of a sample of health sciences and science and technology librarians not currently managing research datasets were interested in gaining competencies in research data management because they saw it as a future component of their responsibilities. Thus, there is a demand for research data management competencies for practicing librarians: They would benefit from knowing the options for finding a LIS course to attain them.

Varvel et al. (2012) did a comprehensive study of 476 data-related courses at 55 iSchool and Library and Information Science programs offered in the fall of 2011 and found only 11 institutions offered 5 programs specifically focused on data curation. The study found 12 institutions offered programs that incorporated data-related content, and 15 that offered courses covering digital content but not on data curation. However, this study did not list the names of these institutions, programs, or courses. In the fall of 2011 and spring of 2012 the authors set out to look at these data-related courses offered by LIS programs and collect a sample that could provide instruction in research data management. The purpose of this paper is to present the results of that analysis so that any librarian interested in taking such a course would have a better idea of the current options available.

Methods

A content analysis of course curricula, descriptions, and syllabi related to research data management offered to students at 58 ALA accredited LIS programs in North America was performed. The sample (N=58) included institutions in the United States, Canada, and Puerto Rico. The researchers searched the course catalogues, syllabi, and course descriptions available on the web-

Table 1: LIS Program Course Listings and Links to Data Management Courses

UAlabama	http://www.slis.ua.edu/Course_Des.html	
SUNY Albany	http://www.albany.edu/informationstudies/index.php	
UAlberta	http://www.slis.ualberta.ca/Courses/GraduateCourses/ListedbyNumber.aspx	
UArizona	http://digin.arizona.edu/courses.html	
UBC	http://www.slais.ubc.ca/courses/coursdes/arst/arst575h.htm	
Buffalo	http://gse.buffalo.edu/lis/course-lis	
UCLA	http://works.bepress.com/cgi/viewcontent.cgi?article=1278&context=borgman	
Catholic	http://slis.cua.edu/courses/courses.cfm	
Clarion	http://www.clarion.edu/1317/?directory-query=Library+Science+Department&g.sc.secondaryEntries=true	
Dalhousie	http://sim.management.dal.ca/Courses/MLIS%20Courses%20Offered/	
U of Denver	http://www.du.edu/education/programs/lis/descriptions.html	
Dominican	http://www.dom.edu/gslis/downloads/courses.pdf	
Drexel	https://duapp1.drexel.edu/webcourses/CourseListing.asp?SubjCode=INFO&Levl=GR&univ=DREX	
Emporia	http://slim.emporia.edu/wp-content/uploads/2011/07/slim-course-descriptions-approved-april-2011.pdf	
FSU	http://slis.fsu.edu/Graduate-Program/Graduate-Courses/Master-and-Specialist-Degrees	
Hawaii	http://www.hawaii.edu/lis/courses.php?page=descriptions	
UofI	http://www.lis.illinois.edu/academics/programs/ms/data_curation	
IUs	http://mypage.iu.edu/~xiaji/S604.pdf	
UIowa	http://slis.grad.uiowa.edu/course_descriptions	
Kent State	http://www.kent.edu/slis/forms-downloads/upload/kent-course-rotation.pdf	
UKY	http://cis.uky.edu/lis/syllabi	
LIU	http://www.liu.edu/CWPost/Academics/Bulletins/Graduate-Bulletin.aspx	
LSU	http://sites01.lsu.edu/wp/slis/course-descriptions/	
McGill	http://www.mcgill.ca/sis/programs/mlis/courses/all	
UMD	http://www.gradschool.umd.edu/catalog/courses/lbsc.htm	
UMich	http://www.si.umich.edu/class/sync-151	
UMissouri	http://lis.missouri.edu/courses	
Umontreal	http://cours.ebsi.umontreal.ca/	
UNC	http://sils.unc.edu/courses/special-topics	
Greensboro-UNC	http://lis.uncg.edu/current-students/	
Central NC	http://www.nccuslis.org/	
UNT	http://icamp.unt.edu	
UOklahoma	http://slis.ou.edu/degree-requirements1	
Pitt	http://www.ischool.pitt.edu/lis/courses/descriptions.php	
Pratt	http://www.pratt.edu/academics/information_and_library_sciences/degree_programs/degree_mslis/	
UPR	http://egcti.upr.edu/index.php/programa-academico/maestria	
QueensCUNY	http://www.qc.cuny.edu/Academics/Documents/Grad_Bulletin_09_12.pdf	
URI	http://www.uri.edu/artsci/lsc/Academics/Courses/coursedesc.html	
Rutgers	http://comminfo.rutgers.edu/component/cur,610/option,com_courses/sch,17/task,listing/	
St. Catherine	http://minerva.stkate.edu/gradcatalog10-12.nsf/courses_web?OpenView	
St. John	http://www.stjohns.edu/academics/graduate/liberalarts/departments/lis/110601_academics	
San Jose	http://slisweb.sjsu.edu/classes/coursedesc.htm	
Simmons	http://www.simmons.edu/gslis/academics/courses/all/index.php	
USCarolina	http://bulletin.sc.edu/preview_entity.php?catoid=35&ent_oid=1391#programs_and_courses	
USF	http://si.usf.edu/graduate/courses/descriptions/	
SCSU	http://www.southernct.edu/ils/coursedescriptions/	
USMississippi	http://www.usm.edu/library-information-science/graduate-course-descriptions	
Syracuse	http://my.ischool.syr.edu/Uploads/CourseSyllabus/687syllabus%20brief-1122.61156-cdc092b5-0fe4-4583-a3cf-1923feeb41ae.pdf	
UTN	http://scidata.sis.utk.edu/scidata_curriculum	
UT	http://courses.ischool.utexas.edu/Anderson_Bill/2012/spring/INF382R/schedule.php	
TWU	http://www.twu.edu/slis/masters-courses.asp	
UToronto	http://www.ischool.utoronto.ca/course-descriptions/inf2115h	
Valdosta	http://www.valdosta.edu/mlis/courses.shtml	
UW	http://www.washington.edu/students/crscat/lis.html	
Wayne State	http://slis.wayne.edu/profiles/7620.pdf	
WOntario	http://www.fims.uwo.ca/acad_programs/grad/lis/mlis/Courses/all-mlis-courses.htm	
UWisc-Madison	http://www.slis.wisc.edu/continueed-DataMgmt.htm	
UWisc-Milwaukee	http://www.graduateschool.uwm.edu/students/prospective/areas-of-study/library-and-information-science/#courses	

Table 2: Sample of LIS research data management courses

LIS Program	Course/Program Title
Indiana University	Data Management
University of California at Los Angeles (UCLA)	Data, Data Practices & Data Curation
University of Illinois at Urbana-Champaign	Foundations of Data Curation, Specialization in Data Curation
University of Michigan	Data Curation
University of North Carolina at Chapel Hill (UNC)	Data Management and Curation, Certificate in Digital Curation
University of Arizona	Certificate in Digital Information Management
University of North Texas	Digital Curation and Data Management Fundamentals and Tools, Applications, and Infrastructure for Digital Curation
Pratt Institute	Cultural Heritage Access Research and Technology
Syracuse University	Scientific Data Management, (CAS) in Data Science
University of Tennessee	SciData , Data Curation Education in Research Centers Program, Data Management Concentration
University of Texas at Austin	Introduction to Scientific and Technical Data Collections: An Introduction to Management and Preservation of Scientific Data
University of Toronto	Data Librarianship
University of Wisconsin at Madison	Introduction to Research Data Management

sites of each institution (see Table 1). Similar to Varvel et al. (2012) and an analysis of biomedical and health informatics programs by Kampov-Polevoi and Hemminger (2010), the researchers complemented the analysis with a keyword search on each program's website. These keywords were: data curation, data management, data science, data sets, data librarianship, eScience, research data, and scientific data. The research team compiled the data over the fall semester of 2011 and the spring and summer semesters of 2012.

Courses that were singularly and specifically focused on an overview of the lifecycle of

management and curation of research data were included; thus, excluding courses in metadata, data mining and structures, digital collections, database concepts, data visualization and analysis, digital libraries, and courses on data networks and information systems. The team analyzed the text of the course listings, syllabi, and descriptions and entered the results into a spreadsheet. The researchers reviewed the final sample of courses using Creamer et al.'s (2011) data management competencies as a framework to identify courses that included assignments and lessons that required students to practice these identified competencies: evaluating data formats and provenance, citing data sets, using metadata to describe data sets, preserving and storing data sets, and performing data interviews and developing a data management plan.

Results

Thirteen (22%) institutions teaching courses on data management and curation that met the competencies identified above (see Table 2) were identified. The programs at UIUC and Syracuse University have specializations in data sciences in addition to their courses on research data management. The University of Michigan, the University of North Carolina at Chapel Hill, and the University of Tennessee at Knoxville offer courses and grant-funded fellowships in research data management. The University of Arizona offers a graduate certificate in digital information management. The University of California at Los Angeles (UCLA), Indiana University, the University of Texas at Austin, and the University of Toronto and University of Wisconsin at Madison offer courses on research data curation and management courses. According to Keralis (2012), the Pratt Institute and the University of North Texas have developed programs. The Pratt Institute is rolling out a Cultural Heritage Access Research and Technology (CHART) certificate program and the University of North Texas and its iCURATE program offers courses on Digital Curation and Data

Management Fundamentals and Tools, Applications, and Infrastructure for Digital Curation.

The UIUC iSchool offers a LIS degree that specializes in data curation. Similarly, Syracuse University's iSchool offers a CAS in Data Science, formerly eScience, and an eScience Fellows program. These schools can offer a comprehensive selection of data-related courses. In addition, the University of Michigan, which joins Indiana University as a Council on Library and Information Resources and Digital Library Federation Data Curation Fellowship Postdoctoral Program site, offers a course on the curation of scientific data and supports an Open Data Research Fellows program. The school has also offered graduate training for Data Sharing and Reuse in eScience, and workshops on Data Management and Data Visualization Needs and Priorities for 21st Century Cyberinfrastructure. The University of North Carolina at Chapel Hill is one of a handful of LIS institutions offering a certificate in digital curation that has a course on the curation of research data. It is also the site of the DigCCurr project. This project sponsors multiple LIS initiatives concerning scientific data curation education including the Carolina Digital Curation Doctoral Fellowship Program. The University of Tennessee at Knoxville has created two programs, SciData and Data Curation Education in Research Centers, to teach research data management and curation to its LIS masters-level students concentrating in data management.

UCLA, Indiana University, the University of Texas at Austin, and University of Wisconsin at Madison serve as examples of how an iSchool or LIS program without a data science-related specialization or concentration can still offer a comprehensive research data management course. In Canada, the University of Toronto offers a course on data librarianship that focuses on the curation of research data and is a partner with the University of North Carolina at Chapel Hill on the DigCCur II project.

Discussion

The objective of this content analysis was to gather a sample of courses among the data-related options available at LIS schools that practicing librarians could choose to gain competencies managing and curating research data. Similar to Varvel et al. (2012), the above sample shows that there are a limited number of iSchool and LIS programs teaching data science. The team found about half of data-related courses at iSchools and LIS programs were exclusively offered online, and a quarter was available with an online option. Our data suggests that online courses will be very important in educating future data librarians because of the low number of LIS programs teaching data management and the uneven geographic distribution of these courses.

One of the limits of this sample is that some LIS schools may only have temporary grant funding to sponsor these programs or courses. In addition, some LIS programs may offer multiple courses that cover some aspect of the research data lifecycle. As Helen Tibbo (2007) noted, "there are several LIS programs offering relevant coursework through multiple courses; however, most LIS programs with a concentration in archives offer relevant, but not fully articulated programs." As shown by Varvel et al. (2012), not every data course covers the same or same number of competencies; however, students matriculating at LIS programs without a data curriculum may be able to take multiple courses within their programs to gain the data-related skills needed to work as data librarians upon graduation. As a result, we would like to extend an invitation to this journal's readership to keep us informed of the continuing research data management-related course offerings for future and practicing librarians.

Conclusion

LIS programs have not sufficiently embraced research data management as a key compo-

nent of their curricula. However, this trend may be changing. After Simmons College assistant dean Terry Plum presented at the 2012 Science Boot Camp, a recent Simmons library school graduate seated in the audience asked him why Simmons- at that time- had yet to create a course focused solely on research data management. He invited the attendees to describe to him what they think this course would look like. He made notes of their suggestions. That summer he began collaborating with the University of Massachusetts Medical School Lamar Soutter Library to create a research data management course for Simmons. LIS schools want to evolve in order to better meet the needs of the profession, and we should continue this dialogue and cooperation between practicing librarians and LIS programs to identify and teach the data skills that current library roles demand. Hopefully this paper will provide other LIS administrators interested in developing such courses with a useful sample of the types of research data management courses available today.

LIS schools are just one of several options for librarians interested in data curation. Indeed, there are an increasing number of open access curricula and tutorials online. The University of North Carolina at Chapel Hill and the U.S. National Archives and Records Administration started the DigCCurr project to create "an international digital curation curriculum composed of openly accessible, graduate-level curricular framework, course modules, and experiential and enrichment components and exemplars necessary to prepare students to work in the 21st century environment of trusted digital and data repositories" (DigCCurr). Other options for data curation education are through collaborations with library partners such as Johns Hopkins Data Conservancy initiative or Purdue University Libraries and the University of Illinois at Urbana-Champaign's Data Curation Profiles Toolkit community. This community, for example, describes itself as "an environment where academic librarians of all kinds, special librarians at research fa-

cilities, archivists involved in the preservation of digital data, and those who support digital repositories can find help, support and camaraderie in exploring avenues to learn more about working with research data and the use of the Data Curation Profiles Tool" (Data Curation Profiles).

Lastly, regional library organizations will be of significant importance because they can offer a range of continuing and professional development opportunities. For example, the National Network of Libraries of Medicine New England Region collaborated with its regional partners and brought Jian Qin and Diane Hillmann from Syracuse University to New England to host Scientific Data Management and Metadata Professional Development Days (Professional Development Day for New England Librarians), as well as Christine Borgman from UCLA to present the keynote address at the 2012 New England E-Science Symposium.   In addition, it worked with its partners at WPI to create a data management curriculum framework that is available online (Frameworks for a Data Management Curriculum). At the local level these types of professional development opportunities will be essential to help support practicing librarians interested in adopting new roles managing and curating research data.

References

Accredited Library Science Programs. *American Library Association.* Accessed September 11, 2011.http://www.ala.org/ala/accreditedprograms/directory/alphalist/index.cfm

Carolina Digital Curation Doctoral Fellows. *University of North Carolina*. Accessed September 20, 2011. http://www.ils.unc.edu/digccurr/aboutII.html#cdcdf

Cheek Fern M., Pamela S. Bradigan. 2010. "Academic Health Sciences Library Research Support." *Journal of the Medical Library Association* 98, no. 2 (2010):167–171.

doi: 10.3163/1536-5050.98.2.011. Accessed September 2, 2012. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2859272/

Creamer, Andrew, Myrna Morales, Donna Kafel, Javier Crespo, and Elaine Martin. "Data Curation and Management Competencies of New England Region Health Sciences and Science and Technology Librarians." Poster presented at the University of Massachusetts and New England Area Librarian e-Science Symposium, Worcester, MA, April 2011. http://escholarship.umassmed.edu/escience_symposium/2011/posters/8/

Creamer, Andrew, Myrna Morales, Donna Kafel, Javier Crespo, and Elaine Martin. "An Assessment of Needed Competencies to Promote the Data Curation and Management Librarianship of Health Sciences and Science and Technology Librarians in New England." *Journal of eScience Librarianship* 1, no. 1 (2012): 18-26. doi:10.7191/jeslib.2012.1006.
http://escholarship.umassmed.edu/jeslib/vol1/iss1/4/

Data Curation Profiles Toolkit. *Purdue University*. Accessed September 21, 2011. http://www4.lib.purdue.edu/dcp/

DigCCurr. *University of North Carolina*. Accessed September 20, 2011. http://www.ils.unc.edu/digccurr/

Digital Curation Centre. "What is Digital Curation?" *Digital Curation Centre.* Accessed September 9, 2011. http://www.dcc.ac.uk/digital-curation/what-digital-curation

Digital Curation Education Program. Accessed September 15, 2011. http://cirss.lis.illinois.edu/CollMeta/dcep.html

Escience Fellows Program. *Syracuse University*. Accessed September 15, 2011. http://eslib.ischool.syr.edu

E-Science Portal for New England Librari-

ans. *University of Massachusetts*. http://esciencelibrary.umassmed.edu

Frameworks for a Data Management Curriculum. *University of Massachusetts*. http://library.umassmed.edu/data_management_frameworks.pdf

Gold, Anna. 2010. "Data Curation and Libraries: Short-Term Developments, Long-Term Prospects." *California Polytechnic State University Digital Commons*. Accessed October 1 2011. http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1027&context=lib_dean

Heidorn, P. Bryan. "The Emerging Role of Libraries in Data Curation and e-science." *Journal of Library Administration* 51, no. 7/8 (2011):662-672. doi: 10.1080/01930826.2011.601269.

Kampov-Polevoi, Julia and Bradley M. Hemminger. "Survey of Biomedical and Health Care Informatics Programs in the United States." *Journal of the Medical Library Association* 98, no. 2 (2010):178–181. doi: 10.3163/1536-5050.98.2.014. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2859255/

Keralis, Spencer D.C. 2012."Data Curation Education: A Snapshot." In "The Problem With Data." CLIR Report. http://www.clir.org/pubs/reports/pub154

Kim, Youngseek, Benjamin K. Addom, and Jeffrey M. Stanton. "Education for eScience Professionals: Integrating Data Curation and Cyberinfrastructure." *The International Journal of Digital Curation* 6, no. 1 (2011):125-138. Accessed August 2, 2012. http://www.ijdc.net/index.php/ijdc/article/viewFile/168/236

Moen, William E. "iCAMP Curate, Archive, Manage, Preserve." *University of North Texas*. Accessed May 20, 2012. http://digital.library.unt.edu/ark:/67531/metadc83318/m2/1/high_res_d/20120402_CNI_S2012_iCAMP_Overview.pdf

Open Data: Graduate Training for Data Sharing and Reuse in E-Science. *University of Michigan*. Accessed September 16, 2011. http://www.igert.org/projects/228

Open Data Research Fellows. *University of Michigan*. Accessed September 16, 2011. http://opendata.si.umich.edu/

Perry Gerald J., Nancy K. Roderer, and Soraya Assar. "A Current Perspective on Medical Informatics and Health Sciences Librarianship." *Journal of the Medical Library Association* 93, no. 2 (2005):199–205. Accessed September 1, 2011. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1082936/

Plum, Terry. 2012 Science Bootcamp Capstone Address. Presented at the Science Bootcamp, *Tufts University*, June 2012. http://esciencelibrary.umassmed.edu/science_bootcamp

Professional Development Day for New England Librarians. *University of Massachusetts*. Accessed September 21, 2011. http://libraryguides.umassmed.edu/2011PDD

Professional Development Day for New England Librarians. *University of Massachusetts*. Cited May 20, 2012. http://tinyurl.com/bv2kbq4

Rambo, Neil. "E-science and Biomedical Libraries." *Journal of the Medical Library Association* 97, no. 3 (2009):159–61. doi: 10.3163/1536-5050.97.3.001. Accessed September 1, 2011. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2706433/

SciData. *University of Tennessee*. Accessed May 20, 2012. http://www.sis.utk.edu/SciData

Tibbo, Helen R. "DigCCurr: Building a Digital Curriculum." *DigCcurr.* Accessed September 21, 2011. http://www.ils.unc.edu/digccurr/digccurr2007_tibbo.pdf

Varvel, Virgil E., Elin J. Bammerlin, and Carole L. Palmer. "Education for Data Professionals: A Study of Current Courses and Programs." Poster presented at the iConference, Toronto, ON, Canada, February 2012. http://cirssweb.lis.illinois.edu/DCCourseScan1/Images/iSchoolPoster2012Final.pdf