

The collaborative creation of a Research Data Management Librarian Academy (RDMLA)

Jean P. Shipman^{a,*} and Rong Tang^b

^a*Vice President, Global Library Relations, Global Strategic Networks, Elsevier, 7909 Desert Ridge Cove, Cottonwood Heights, UT, USA*

^b*Associate Professor, School of Library and Information Science, Simmons University, 300 The Fenway, Boston, MA, USA*

Abstract. With the growing momentum of research data services in libraries, a team of librarians and Library Information Science (LIS) faculty members have been working together to develop a free, online Research Data Management Librarian Academy (RDMLA) aimed primarily for practicing librarians and information professionals, but also for researchers to gain knowledge about research data management (RDM) principles and best practices. The training modules will be made available to anyone across the globe. The development team includes librarians from Harvard Medical School, Harvard University, Tufts Health Sciences, Massachusetts College of Pharmacy and Health Sciences (MCPHS), University, Boston University School of Medicine, Brown University, Northeastern University, Elsevier, and Simmons University. Simmons University's School of Library and Information Science will grant continuing education credit for those desiring such. This is a unique partnership between librarians, LIS educators, and a publisher. The RDMLA is hosted online on Canvas under the CC-BY-NC-SA (Attribution-NonCommercial-ShareAlike) licensing.

The need for this training was evidenced through interviews and online surveys which identified gaps in current training offerings and highlighted what skills librarians and researchers need to contribute to their RDM success. An inventory of existing courses was prepared along with a review of job description competencies. This publication outlines the outcomes of the needs assessment, inventories, and the Academy's training unit content.

Keywords: Research data management, training, librarians, RDM, online curriculum, continuing education

1. Background

Librarians' organization skills are transferable to the management of research data as librarians know how to describe data to effectively store it for easy retrieval and discovery for later use and reproducibility. Over the past five years, many library positions in this area have been created and filled by talented librarians. The titles for these positions range from Data Librarian to Data Specialist, Data Visualization Librarian, and Data Scientist. While LIS programs, and i-Schools are starting to add courses to support this important librarian role, practicing librarians and information specialists find it difficult to acquire the needed additional skills, as they cannot easily leave their positions and often do not have the time to participate in formal coursework or academic programs.

* Corresponding author. E-mail: j.shipman@elsevier.com.

To meet this need, Elsevier partnered with a team of health science librarians, library administrators, and LIS educators from the greater Boston area to develop the Research Data Management Librarian Academy (RDMLA). The RDMLA project is led by Dr. Elaine Martin, Director of the Countway Library of Medicine at the Harvard Medical School, Dr. Rong Tang, Director of Ph.D. Program and Associate Professor at the School of Library and Information Science, Simmons University, and Jean P. Shipman, Vice President, Global Library Relations, Elsevier. Dr. Martin is a renowned name with data management, as she created a bootcamp series for librarians when she was the director of the Northeastern National Network of Libraries of Medicine. She also received funding to create a Massive Open Online Course (MOOC) on the topic as part of the BD2K (Big Data 2000) federal initiative. Dr. Rong Tang is a well-accomplished researcher in areas of usability and user experience (UX), information behavior, data science curriculum, and open government data (OGD). Drs. Martin and Tang had organized a team of librarians and LIS faculty members from the following institutions: Harvard Medical School, Harvard Library, Tufts Health Sciences, MCPHS University, Boston University, Brown University, Northeastern University, and Simmons University. RDMLA is a unique partnership between librarians, LIS educators, and an information analytical company and publisher. Elsevier is proud to sponsor the RDMLA. But is an RDMLA needed?

2. Needs assessment

This was the question that Elsevier posed to the project team. Initial funding for a needs assessment and inventory of training opportunities within this topical area was granted by Elsevier to the project team. The needs assessment was conducted with both librarians around the world and library educators. A survey was issued via *Library Connect*, Elsevier's outreach program for librarians. Another survey was distributed via the Association for Library and Information Science Education (ALISE) and JESSE (LIS faculty) listservs. Alyson Gamble, a Simmons University Ph.D. student, was hired as a data fellow to coordinate the needs assessment and training inventory and to analyze the results. All results are available at <https://rdmla.github.io/about/>.

3. Library connect survey results

The *Library Connect* survey consisted of nineteen questions developed by Jean Shipman and Dr. Elaine Martin. The survey was issued using SurveyMonkey. The survey covered four months, from April 23, 2018 to September 4, 2018. Responses totaled two hundred and forty-one, and respondents were from all around the globe. Responses to open-ended questions were coded and themed; numbering 628. The survey revealed that the top challenges librarians faced in providing RDM services included "capacity/bandwidth; limited staffing," "marketing and outreach of RDM service," "collaborative understanding among campus departments," "upskilling staff," "providing consistent service in terms of quality and options," and "handling researcher's and faculty's misconception of RDM and library services." The most essential RDM skills and knowledge, according to the respondents, were "data services/science skills," "skills regarding data reference interviews" "basic training," and "connecting with the researcher and faculty."

4. LIS educator and i-School email survey results

Six responses were received for the short email survey sent via the ALISE membership listserv and summarized. All respondents reported that their schools offered some courses in RDM, ranging from data analytics, data curation, data visualization, data mining, data science, and text mining, through to metadata. One of the schools offered a separate degree focused on RDM. Several schools were partnering with computer science departments to create such coursework. One respondent was interested in exploring “How practitioners with the expertise and interest in teaching in this area could become involved in existing academic programs”.

5. Training inventory

A review of different training opportunities was conducted. Workshops, MOOCs, webinars, as well as data science courses offered via LIS programs and i-Schools were analyzed to determine topical areas of instruction. Topics covered across training opportunities included data curation, disciplinary/funder requirements, data management plans, data storage, infrastructure, security, software, ethics, consultations/instruction/outreach, and metadata. Results are also available on the project’s github site.

6. Job description analysis

Job descriptions for positions relating to RDM were analyzed to determine the key competencies requested by employers. Professional positions were identified from the American Library Association (ALA) and Association of College and Research Libraries (ACRL) JobList, Association for Information Science & Technology (ASIS&T) listservs, the Code4Lib jobs site, and more. Of 44 jobs identified, “consultation, instruction, and outreach” was the most frequently mentioned roles, followed by “disciplinary funder requirement,” “data curation,” “data management plans (DMPs),” and “metadata.”

7. RDMLA curriculum

7.1. Overview of curriculum

Using the results of the needs assessment and training inventory, the project team discussed the most glaring gaps within the current training opportunities and the topics most requested in the surveys. Two key library data service-related training needs were identified. These included: (1) being able to articulate the need for a data research management service by librarians to both the library and university administration, and (2) learning how to establish such a service within a library based on the experience and lessons learned by others. Other instruction requested addresses the ability to learn more about common coding software and data management platforms. A foundation course provides the setting and an overview of the curriculum. There are eight total units within the curriculum. Each unit may be taken alone or if all the units are completed, continuing education certificates can be obtained from Simmons University’s School of Library and Information Science. The eight units include:

- (1) Foundations of research data management
- (2) Navigating research data culture
- (3) Advocating and marketing the value of RDM in libraries
- (4) Launching data services in libraries
- (5) Project management and assessment
- (6) Overview of data analysis and visualization tools
- (7) Overview of coding tools
- (8) Overview of platform tools

8. Audience

The main audience for this training is librarians of all types: academic, medical, special, public, government, and school. However, the key audience is the practicing librarian who is unable to leave the workforce to obtain additional formal training. Professionals who work in a data-intensive working environment may also benefit from the RDMLA. Several of the units in the RDMLA curriculum will also assist researchers to learn more about how to manage their laboratory data. The RDMLA can be taken by any individual around the globe who can comprehend the English language. There is no registration fee required for completing the units.

9. Curriculum content developers

The eight units comprising the RDMLA were created by teams of librarians and LIS faculty with expertise in the specific topic areas. Each team worked with Simmons Online's (<http://internal.simmons.edu/faculty-staff/general/strategic-initiatives-simmons-online/>) instruction designers to create a uniform framework for the units to be offered via the Canvas learning management software. The units are self-guided and asynchronous for convenience. Videos were created to illustrate key aspects of the topics. Assignments are available as well as interactive quizzes for individuals to assess their comprehension of the content. References to required and supplemental readings are provided. As an example, the last unit addresses RDM software platforms, both open source and commercial ones. A grid comparing the different solutions is provided, as well as a link to Elsevier's Mendeley Data informational website which also contains a promotional toolkit.

10. Licensing and continuing education credit

The RDMLA units are hosted online on Canvas under a CC-BY-NC-SA (Attribution-NonCommercial-ShareAlike) licensing. Simmons University's School of Library and Information Science will provide continuing education credit to participants who wish to pursue such credit. A fee will be required to cover the costs of providing such documentation. Individuals completing the RDMLA units are encouraged to indicate such on their resumes to indicate to potential employers their specialization in RDM. Registration is required of participants who wish to receive a certificate of their Academy completion.

11. Public dissemination

The RDMLA has been and will continue to be promoted via presentations at many library, i-School and library school conferences, both domestic and international. Other conferences such as those held by NFAIS and Researcher to Reader also were venues employed to share information about the RDMLA. A press release will be issued by Elsevier as will a *Library Connect* article when the RDMLA is available. The listservs used to distribute the needs assessments will receive word of the Academy's availability. Referrals will also be encouraged. The Academy developers plan to partner with the Association of College & Research Libraries, the U.S. Medical Library Association, the Special Library Association, and the National Library of Medicine's National Network of Libraries of Medicine to further spread the news. Each project team partner will also communicate the curriculum's availability within their institutions and professional associations. A professional publication is in press.

Acknowledgements

The RDMLA is a product of many individuals' contributions. A roster of individual contributors is available on the project's github site (see: <https://rdmla.github.io/partners/>). This article is based on a lightning talk presented at the February 2019 NFAIS Conference in Alexandria, Virginia.

About the First Author and Presenter

Jean P. Shipman, MSLS, AHIP, FMLA, is the VP, Global Library Relations for Elsevier. Prior, she was the Executive Director, Knowledge Management and Spencer S. Eccles Health Sciences Library; Director of the Mid-Continental Region and National Training Office of the National Network of Libraries of Medicine; Director for Information Transfer, Center of Medical Innovation; and Adjunct Faculty of the Department of Biomedical Informatics, School of Medicine; all at the University of Utah. She has also been employed by the John Hopkins University, Greater Baltimore Medical Center, University of Maryland, University of Washington and Virginia Commonwealth University. She served as president of the Medical Library Association for 2006–2007 and on the Board of Directors for the Society for Scholarly Publishing from 2013–2016. She was a member and co-chair of the Chicago Collaborative, a group of publishers, librarians and editors, who met to discuss issues regarding scholarly communications. She is the co-editor of two books: *Information and Innovation: A Natural Combination for Health Sciences Libraries*, and *Strategic Collaborations in Health Sciences Libraries*. She has also authored many journal articles, book chapters and given numerous professional presentations. Telephone Number: +1 385 347 1829; Email Address: j.shipman@elsevier.com.