

Building the Digital Curation Workforce: Advancing Specialized Data Curation

Washington University in St. Louis (Cynthia Hudson-Vitale) in partnership with the University of Minnesota (Lisa Johnston), the University of Michigan (Jacob Carlson), Cornell University (Wendy Kozlowski), Pennsylvania State University (Robert Olendorf), and the University of Illinois, Urbana-Champaign (Heidi Imker) (known as the “Data Curation Network”), propose a three-year project (July 1, 2018 - June 30, 2021) to develop an intermediate to advanced level data management and curation training program open to all academic library staff. The project team is requesting \$249,424 in support from IMLS. This program seeks to build and extend capacity for advanced data curation and discipline-specific data management among academic library staff nationwide.

The past few years have seen an increased demand for subject liaisons, library archivists, digital curators, and library repository managers to incorporate data management and curation activities into the suite of services provided to faculty and students. Subject liaisons possess specialized discipline knowledge and expertise that can be leveraged to support the management and curation of research data. Yet, in practice, many lack the confidence and training to appropriately support researchers demand. Additionally, repository managers, archivists and others tasked with curating research data in academic libraries are required to curate across many different disciplines and data types, often without extensive knowledge or expertise in all of those areas. This program will address both the need for front line library staff to gain confidence in research data-related services, and upskill the discipline knowledge of research data staff, through a peer-to-peer training model.

Statement of Need

Results from our recent [SPEC Kit \(#354\)](#) indicate that library staffing for data curation is one of the most significant challenges that academic libraries face in providing curation services. Many institutions rely on the equivalent of less than one full time person to provide research data management and curation support. Despite these lack of resources, data curation services have been repeatedly identified in reports and publications as a growing service area for the transformation of academic libraries.

Currently, there are several digital curation training opportunities available as continuing education for library and archive professionals (IMLS funded DigCCurr, SAA’s DAS certificate, LoC’s DPOE program, Data Science for Librarians, Data Carpentry). Yet, each of these programs focus on foundational skills related to data curation and lack the much needed advanced, discipline, and functional training required to appropriately curate and manage research data. According to a recent [CLIR](#) report, these existing curriculum materials “could be adapted, extended, or built upon to expand and scale up current educational and training offerings to address increasing needs.” Specifically, this program will seek to address the [CLIR report](#) recommendation to collaboratively develop interdisciplinary, competency-building curriculum and training. The National Research Council (NRC) report, [Preparing the Workforce for Digital Curation](#), provides an analysis of the curricula learning outcomes needed for training a digital curator workforce, in which one of the key skills highlighted was a discipline specificity. Our work will build on this by connecting data curator skills with data management skills and domain and functional expertise. Our co-PI instructors bring expertise in this area as directors of institutional data repository and curation programs across six major universities, as well as leads on past projects such as the Data Curation Profiles ToolKit and the Data Information Literacy (DIL) project that developed 12 competencies for teaching STEM graduate students data management skills.

National Impact

Ultimately, the training provided by this program will extend the expertise of all library staff into data management and curation, while building discipline-based competencies among data curators and research data management staff found across the library landscape. Additionally, this program will build and make widely available much needed functional and discipline specific curation and management resources for other librarians to use and build upon.

Project Design

Our approach will leverage the in-depth digital data curation training for data curators established for the [Data Curation Network](#) (with hands-on curation procedures for a variety of disciplinary subject areas). The project grant is unique as it will build on successful digital curator training competencies developed in past IMLS projects ([CLIR pub. 174](#)) but innovate these past approaches by deepening the focus in the following ways:

- Focus on intermediate to advanced skills for research data management (RDM) and data curation from a discipline perspective;
 - Prepare participants to work with the researcher much earlier in the research process
 - Provide mock training of what could happen in the faculty's office; build confidence among curators and library staff to talk about data issues and to be proactive.
 - Tie back to researchers process and workflow to ensure reproducible research.
- Incorporate the experiences and expertise of participants to enrich the curriculum and iterate the training content each subsequent session that is offered.

Workshop Format & Content: This program will consist of a 3 day training event every 6 months over the course of the three-year project. Members of the Data Curation Network will provide instruction and facilitation of the core data curation training activities. The training will focus on in-depth *functional* and *discipline-based* data curation skills and competencies (rather than on the administrative or managerial skills needed to develop and execute a robust digital curation program). As defined by the [CLIR](#) report, “functional areas of curation [are] essential for long-term curation, including appraisal, acquisition, identification, research, development, forensics, accessibility, monitoring, destruction, tools and workflows, preservation planning, digitization, storage management, records management, rights management.” Specifically, the project team will draw upon the results of our [previous research](#), in which faculty and researchers identified the most important curation activities for their work.

Capstone Project: Participants of the program will pair up (one data curation/management librarian and one functional or discipline-expert) and complete a capstone project, due 6-months later, in the form of a training module or object that will be used in the next iteration of the training. To prepare for this capstone at the time of registration, attendees will be asked to identify a functional or discipline specific area of curation activity or support that is currently lacking from the existing curation training available. Pairs will then have dedicated time during the 3 day workshop to create a plan to develop the curriculum over the next six months. Before leaving the workshop each pair will present their plans to the cohort – who will provide feedback and make suggestions for improvements. One required component of the capstone will be for participants to interview 1-2 researchers about their research workflows - thus ensuring the modules are based on real-world processes. We will iterate this approach four times building a more comprehensive training program garnered from the combined skills and expertise of the practitioner participants who attend each session. The project will output a final set of training modules that will be shared with the community at large as open educational resources.

Diversity Plan

We will work with ARL diversity initiatives and the CLIR Fellows program to recruit a diverse cohort of participants. Early career, non-traditional librarians, paraprofessionals, and curators will be encouraged to attend. The Data Curation Network is committed to maximizing the recruitment and retention of women, underrepresented racial and ethnic groups, and people with disabilities. All of the partnering institutions have ongoing programs that focus on the recruitment and retention of trainees and faculty from diverse ethnic backgrounds that are typically underrepresented.

Budget

The total budget for this project is \$249,424 with \$183,400 in direct costs and \$66,024 for indirects. This amount includes \$96,000 for 64 participants travel to participate in this program. A budget of \$42,000 to cover the costs of the four trainings (meals, room rental) and \$39,000 to support the travel and provide stipends to trainers. An honorarium of \$100 will be provided to each researcher who completes an interview as part of the capstone project (\$6,400).