Learning about research data in the lab at the Pitt iSchool

BY LIZ LYON | OCTOBER 23, 2014

Skilling up for RDM

Today’s research environment features data-intensive science, funding mandates for enhanced research data management (RDM), and high-level policy directives towards a more open sharing of data products from research projects. This has led many university and college libraries to take on new responsibilities; libraries are repositioning, restructuring and reskilling their staff to take on RDM roles.1, 2

In response, iSchools and library schools are introducing data curation and RDM courses into the curriculum to prepare the next generation of information professionals and to retrain existing librarians for these new roles.

Highlighting a domain disconnect

Many of these new roles embrace elements of data stewardship, data management, data science and data visualization.3 Job titles vary from data librarian, data archivist and data steward to data analyst, data engineer and data journalist. While some of these roles may not immediately be considered applicable to libraries, others have potential to become mainstream library positions.

It is also important to recognize the diversity of data in terms of standards, formats, ontologies and cultural practices that characterize different disciplines or communities. One challenge for librarians is how to best address this breadth; there is a real risk of a domain disconnect between the data requirements and support services offered. New courses for MLIS and doctoral students and librarians in the iSchool at the University of Pittsburgh seek to bridge this disconnect by taking an innovative, immersive approach — placing students within an active research environment such as a science laboratory.

Heading to the lab

Pitt’s new RDM courses are based on an immersive informatics pilot study, a novel RDM training course that was developed and delivered by a partnership between UKOLN Informatics, University of Bath and the University of Melbourne.4
The first new data course, “Research Data Management,” focuses on funder requirements, data management plans, legal and ethical issues, data centers, data advocacy and data sustainability. The second, “Research Data Infrastructures,” explores data storage, data publication and citation, data discovery, data standards, data repositories, data preservation and data science.

As part of these courses, the iSchool students have immersive placements in a range of labs at the University of Pittsburgh, including the Departments of Public Health and Medicine (specifically, the Renal-Electrolyte and Epidemiology laboratories) and the Swanson School of Engineering (at the RFID center of excellence, which studies radiofrequency identification technology). The students work in pairs and are assigned to a doctoral student, instructor or post-doctoral researcher. In the lab, students can observe day-to-day research workflows and have access to live datasets, laboratory notebooks and sophisticated instruments — experiences most MLIS students would not otherwise encounter in their curriculum.

One challenge for librarians is how to best address this breadth; there is a real risk of a domain disconnect between the data requirements and support services offered.

Bilateral learning

The goal is an exchange of RDM experience, questions and knowledge between the iSchool student and the research scientist. Feedback from both students and researchers has been positive.

From students:

“It was great to see a real-life example of how a lab generates and uses data.”

“We learned not only about the specifics of their research but also about the lifecycle of data.”

“This was a valuable experience. It was very practical and illuminated some of the struggles that one may encounter in discussing data as its own area of research.”

From researchers:

“Showed them data backup on three drives — they asked a question about fire risk.”

“Explaining what one does to a new person is instructive, since it shows you what you do not understand and cannot explain. Discussion with the (LIS) student exposed some weaknesses in my own thinking.”

Realizing the value and benefits for libraries

Immersive RDM courses have many potential benefits:

- Library-faculty partnerships deepen by having library and informatics experts bring their knowledge into the laboratory and work alongside researchers.
- Librarians gain firsthand understanding of the laboratory setting, bench-based workflows, instrumentation outputs and researchers’ day-to-day data challenges.6
- Faculty researchers learn from library students about data mandates, software, methodologies and data handling.
- The enhanced knowledge and new data skills, which are in short supply, empower LIS students and library staff to adopt new RDM roles.6
- Library senior managers gain intelligence about how to structure their organization to deliver research data services.
- LIS students enter the job market better positioned for new career opportunities working with data.

References


