

# Northwest Five Consortium Data Management and Curation Workshop

*Final Report, Summer 2015*

## Project Summary

The Northwest Five Consortium (NW5C) Summer 2015 Data Curation Workshop was developed to promote learning and collaboration across NW5C institutions in the areas of research data management and curation (RDMC). Organizers used a model bringing together research teams from each institution; teams consisted of a faculty member, one or two student research assistants, a librarian, and (optionally) an information technology staff member. This team model allowed for shared learning and collaborative development of best practices. Learning was supported by a highly-skilled facilitator and a practical, project-based focus. In total, twenty faculty, students, librarians and professional staff representing all of the NW5C schools registered for the workshop. Amanda Whitmire, Assistant Professor and Data Management Specialist from Oregon State University, served as facilitator, and the workshop was held on June 18-19 at Reed College. Teams generally conformed to the envisioned model, with some minor modifications based on individual needs.

The workshop was designed as pilot test for a model of librarian-researcher collaboration across campuses and to allow teams to undertake a preliminary assessment of current RDMC practices. To this end, prior to the workshop, each librarian conducted a detailed interview with the faculty researcher at their institution regarding the data management practices related to a current project. Librarians drafted Data Curation Profiles (DCP) reflecting these interviews, which were submitted to the facilitator before the workshop to guide the crafting of curriculum relevant to the specific data management landscape and challenges facing the participants. These DCPs were based on the protocol developed at Purdue University Libraries and currently supported by an Institute of Museum and Library Services Planning Grant.<sup>1</sup>

The workshop curriculum was crafted by the librarians in collaboration with the facilitator with the aim of developing participants' awareness and understanding of data management issues. Topics covered included data management planning, file-naming and organization, storage and backup, metadata, data sharing, and long-term preservation. As organizers recruited participants and gained a clearer understanding of their immediate needs, the focus shifted to include skill-building in areas such as developing workflows and clarifying laboratory protocols. The workshop included various active learning exercises and time for self-reflection and discussion. Time was set aside for the teams to work together on developing strategies to meet their own needs, with the facilitator available for assistance by request. Time was also set aside for the librarians to meet and share their local experiences drafting their DCP and brainstorming steps for moving forward and addressing RDMC issues at their campus beyond the needs of the specific research team with which they were working.

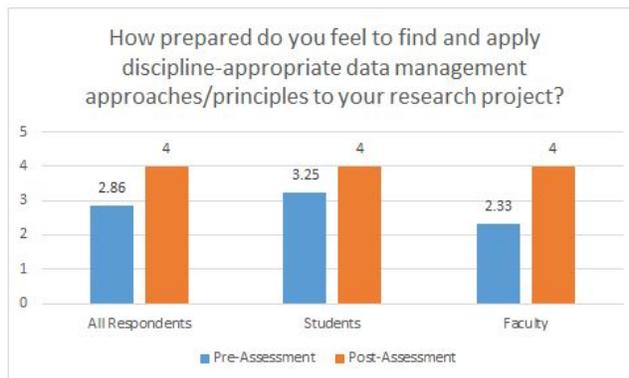
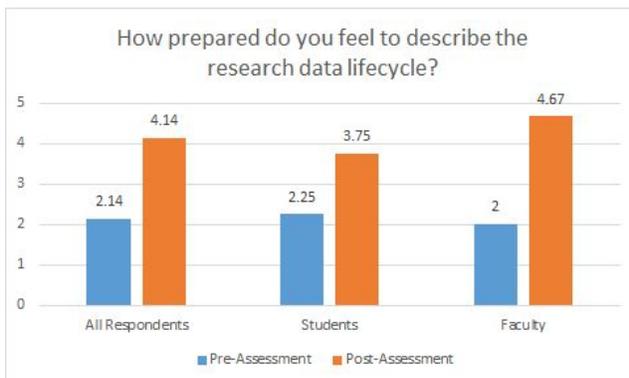
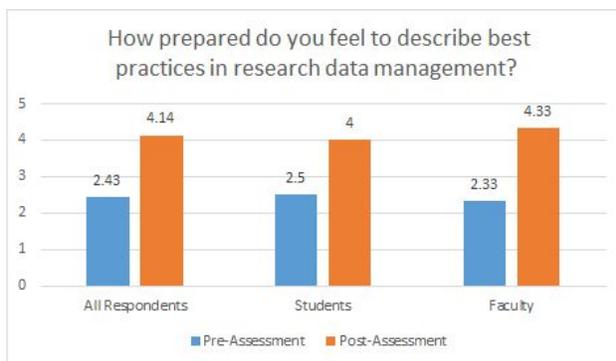
---

<sup>1</sup> Data Curation Profiles website <<http://datacurationprofiles.org/>>, accessed August 20, 2015.

## Outcomes and Assessment

The organizers see the team model as a success in developing new learning opportunities for students while also increasing faculty and staff communication in a collaborative cross-institutional environment. The workshop successfully delivered on the proposed objectives:

- Increasing participant awareness and understanding of data management topics
- Piloting a model for librarian-researcher collaboration across campuses
- Providing time and support to undertake a preliminary assessment of current practices
- Introducing effective techniques for locating and incorporating domain-specific best practices and requirements in RDMC



A pre- and post-workshop survey was distributed to all participants to assess base levels of understanding of and familiarity with the material. Separate assessment instruments were used for the faculty/student research teams and the librarian participants. Total registered attendees of the workshop included 14 faculty/student researchers and 6 librarians; pre-workshop completion was 79% for the faculty student/research teams (n=11) and 100% for the librarians (n=6) while post-workshop completion was 57% for the faculty/student teams (n=8) and 67% for the librarians (n=4).

All respondents from both respondent groups agreed that the workshop format and team model were effective. One faculty member wrote, "Working in teams was KEY. We need a follow up either at our own institution or another NW5C workshop!" A student wrote, "I found it very helpful to be introduced to the concepts behind data management... Being aware of the advantages of data management, I can progress in my career and find various applications for them. This seems far more advantageous than simply being told by a PI that the lab follows XYZ protocols, which I may or may not continue to apply in the future."

In general, faculty feedback focused on appreciation for newly learned skills which could be immediately implemented in their labs

for data backups, efficient data storage, and consistency in recording & formatting data.

Faculty members commented that the preliminary DCP work was helpful in preparing for the workshop and for further RDMC planning. Researchers also discussed potential ways to share data beyond this workshop. Faculty supported extending the workshop's mission to other groups on their home campuses. Librarians raised some issues that were not within the explicit scope of the workshop, such as building a shared data repository, and posed questions about how data management plans might be applied on an individual level to undergraduate research. One of the student researchers voiced a desire that future workshops provide more space for interaction and discussion with student workshop participants from other institutions.

### **Key Findings from Student & Faculty Responses**

- 100% of respondents agreed workshop model was an effective way to learn about research data management
- All faculty referenced the importance of working in teams and fostering collaboration
- Specific skills (e.g. file naming conventions, file types, readme files) as well as larger concepts (e.g. data curation, metadata) were noted as important takeaways from the workshop
- Participants expanded their definitions of "data management" to emphasize the importance of planning
- When defining the "research data life cycle," faculty were more likely to provide more detailed explanations, specifically emphasizing data storage and preservation issues

### **Key Findings from Librarian Responses**

- All librarians reported a significant increase in comfort levels with the subject matter and feeling of prepared to do outreach to faculty or colleagues
- All librarians mentioned networking or collaboration with NW5C peers when asked about the most important outcome of the workshop

### **Observed Outcomes at Participating Institutions**

- At one institution, the lab group "created a file naming protocol that is now in use, modified...data storage workflows... created a more complete metadata sheet...and created a discarded data log that includes justification statements"
- Another team described new practices within the lab including new file-naming conventions, a modified workflow for documenting data collection processes, as well as an overall increased emphasis on the importance of keeping data management at the forefront of experimental design
- The librarian and faculty researcher at a third institution have begun collaborating on a proposal for increased data services support from the library, while student researchers from the team presented on the topic of data management to other undergraduates engaged in summer research

## **Future Programming Possibilities**

With the positive response to this workshop and the continued interest in developing RDMC services at the NW5C campuses, organizers discussed several future programming possibilities, outlined below.

### **Future Workshops**

- If funding allows, repeat the Summer 2015 model of an in-person workshop with teams from all institutions and an external facilitator centralized at one of the Portland NW5C institutions
- As a lower cost, alternative, hold a virtual 5-campus RDMC workshop based closely on the Summer 2015 model, with a single shared presenter, using videoconferencing to bring participants together
- “Train the trainer” diffusion of the Summer 2015 model, where the participating librarians extend workshops to their own campuses by acting as local facilitators
- Building on the “Train the trainer” model, participating librarians create more focused discipline-specific workshops across their campuses

### **Virtual Communities of Practice**

- Implement a RDMC email list, where 2015 participants could opt-in to share ideas and practices; this list could easily be expanded with new cohorts of workshop participants or other interested parties
- Participating librarians could collaborate on creating resources to support RDMC in the form of research guides, handouts, workshop curricula, and other materials using the network developed in planning and implementing the 2015 workshop

### **Shared Infrastructure**

- The possibility of a shared data repository for NW5C institutions was raised during the closing discussion of the workshop; any such project would require significant collaboration between campuses, as well as sustained financial and administrative support across the NW5C institutions

### **Continued Follow-up Assessment**

- As outlined in the project proposal, organizers will again survey workshop participants at 6 months following the workshop (December 2015); in this follow-up, organizers will again ask participants for ideas for future support for RDMC on the NW5C campuses