Data Curation Education in Research Centers: Formative Evaluation Findings from Years One and Two

Carole L. Palmer1, Cheryl A. Thompson1, Matthew S. Mayernik2, Valerie Williams2, Karon Kelly3, and Suzie Allard3

1Center for Informatics Research in Science & Scholarship (CIRSS), Graduate School of Library & Information Science, University of Illinois at Urbana-Champaign (UIUC)
2National Center for Atmospheric Research (NCAR), University Corporation for Atmospheric Research (UCAR)
3School of Information Sciences, College of Communication & Information, University of Tennessee at Knoxville (UT)

Data Workforce Problem
The current data curation workforce consists primarily of:
• Scientists and technical experts who receive little formal training in data management
• Data managers educated through on-the-job training

Few education programs exist that provide formal training specifically related to data management and curation.

DCERC Model
DCERC is developing a new model for educating data professionals that:
• Introduces library & information science students to data curation practices and issues in a research center environment.
• Designs and delivers foundational courses in data curation.
• Provides Masters and Doctoral internships at a data-intensive research center.
• Builds community among students, science and data mentors, and faculty.
• Contributes to data curation research results, documentation of curation best practices, and advances in practitioner training.

Masters Internship Experience
• Masters students intern at NCAR for two months in summer.
• Summer internships are launched with an introductory workshop.
• Students learn current data management practices and challenges while developing expertise and conducting research.
• Each student is paired with two mentors:
  - Data mentor – provides data management expertise
  - Science mentor – provides expertise of the scientific domain, and real-life data management challenges

Formative Evaluation Questions
• How well does the internship experience match the needs of the students? The mentors?
• Which elements of the internship did the students find most valuable? Least valuable?
• Which areas of the internship experience can be improved?

Data Collection Methods and Sources
• Participants: 5 Master’s students and 12 mentors from 2012 & 2013 internships.
• Student feedback: Pre-internship survey; a post-workshop survey, student e-journal entries; post-internship survey and focus groups.
• Mentor feedback: Post-internship survey and focus groups.

Year One Feedback
• Positive experiences:
  - Foundational course and internship experience were extremely valuable, and in line with student career objectives.
  - Students described their mentors as great, accessible, supportive, and enhancing the learning experience.
  - Mentors appreciated the information science perspective.
  - Students valued professional development opportunities.
• Adjustments needed:
  - Opening workshop, while informative, was overwhelming.
  - Student project organization needed improvement:
    - Students needed better orientation to NCAR or mentors.
    - Mentors had limited understanding of student skills.
  - Internship timelines were very tight.

Year Two Feedback
• Positive experiences:
  - Projects had better alignment between student/mentor goals and student skills.
  - Students valued the additional professional development opportunities.
  - Mentors appreciated the students’ analytical approaches.
• Adjustments needed
  - Mentors noted the significant time commitment.
  - Data mentors stressed importance of computing skills, which Library & Information students may not have.

Testimonials
“I had a chance to follow the complete data curation process rather than doing one piece... I had a chance to see the big picture and appreciate the time and energy it takes to do the job right.” - Student
“I learned how the new generation of data curators thinks, designs solutions, and how they perceive the expected audience of their work.” - Mentor

Next Steps
• Additional cohorts of Master’s students at NCAR in 2014.
• Evaluate the sustainability of the internship model over time within a research center environment (e.g. NCAR).
• Investigate need for data curation expertise & viability of internship program with broader set of scientific data centers.
• Continue to inform curriculum development with evaluation results.
• Extend and sustain the DCERC model beyond the initial partnership.

Acknowledgments: The DCERC program was funded by the Institute of Museum and Library Services (RE-02-10-0004-10). The DCERC program is led by CIRSS at the University of Illinois, in collaboration with the School of Information Sciences at the University of Tennessee and the NCAR Library.