Enriching Education with Exemplars in Practice: Iterative Development of Data Curation Internships

Matthew S. Mayernik & Valerie Williams
National Center for Atmospheric Research, UCAR

Carole L. Palmer
Information School, University of Washington

Cheryl A. Thompson
Graduate School of Library and Information Science, University of Illinois at Urbana Champaign

Suzie Allard & Carol Tenopir
School of Information Sciences, University of Tennessee at Knoxville
Data Curation Education in Research Centers (DCERC)

PI: Carole L. Palmer
Co-PIs: Mary Marlino, NCAR & Suzie Allard, UT
Collaborators: Carol Tenopir, UT & Matt Mayernik, NCAR

Dates: 2010-2015

Goals

- Model for Library and Information Science education
  - Enrich curriculum and build capacity to strengthen expertise in LIS
  - Internships in research and data centers -- current state-of-the-art practices and front-line challenges

- Evaluate model
  - Effectiveness
  - Viability of scaling to network of partner data centers and iSchools
DCERC program

- Graduate level curriculum focused on data curation
  - Master’s level enrichment of foundations course
  - New Doctoral seminar on research problems

- Academic research experiences with faculty mentors

- NCAR field experiences
  - 8 week internship projects guided by mentors
  - Begins with kick-off workshop
  - NCAR-wide Intern Final Presentations
NCAR internships

• Three cohorts offered summer 2012 -2014

• 9 Masters students participated, 10 total internships

• Students paired with multiple mentors
  • Data and Science mentors
  • 16 NCAR staff members served as mentors
  • 4 data mentors served multiple years
Example student projects

Climate model metadata

Archiving data

Data and metadata organization

Archiving temporal and spatial data

Assess analog data for digital access

Translator and Facilitator

Understands and articulates the needs and goals of scientist
Understands and articulates data manager needs for curation
Creates guidelines to enhance communication and efficiency between scientists and data managers

Sensor data, data processing, file migration

Cross-disciplinary data curation; data subsetting

Time series data, high resolution, provenance, NetCDF

Metadata harvesting, data quality, standard compliance

International data, 50 collections, OAIS, DOI assignment
Formative evaluation questions

• Which processes are functioning well or need improvement from the student and mentor perspectives?

• To what extent are the primary activities yielding intended short term outcomes (e.g., application of theory to practice)?

• To what extent does the DCERC model improve education for students? Contribute to NCAR’s operations and mission?
Internship evaluation methods

• Questionnaires
  • pre- & post-internship with students and mentors
  • post-workshop

• Weekly student reflections

• Post-internship focus groups of students and mentors

• Employer perspective interviews
Hitting the ground running

- Increase pre-summer communication
  - Conference calls orienting to NCAR organization and science
  - Project specification between students and mentors

- Streamline the kick-off workshop
  - Reduce the length but increase project discussion time
  - Reduce external speakers

- Provide more preparation
  - Students: exposure to science, research projects, visit to research center
  - Mentors: exposure to LIS, student background, coursework, and scoping projects
More professional development and networking

- Attending GeoData, ESIP meeting, The HDF Group, etc.
- Informal mentoring
Key considerations

*in applying the model*

- Multi-mentorship from scientists, data professionals
- Collaborative, peer-to-peer project design
- Recognition of LIS contributions to data management
- Formative evaluation & iterations
Acknowledgements

Carole L. Palmer, PI, Cheryl A. Thompson
University of Illinois Urbana-Champaign
Graduate School of Library and Information Science

Suzie Allard, Carol Tenopir
University of Tennessee
School of Information Sciences

Mary Marlino, Matthew S. Mayernik
National Center for Atmospheric Research

Students, Scientists and Data Mentors at NCAR

DCERC is funded by the Institute for Museum and Library Services
Laura Bush 21st Century Librarian Program

Many images in this presentations are from the UCAR Digital Image Library. Available at: http://www.fin.ucar.edu/ucardil/