Curating data collections in the classroom: lessons learned

RUTH DUERR[A] & TIFFANY CHAO[B]

[A] NATIONAL SNOW AND ICE DATA CENTER (NSIDC), UNIVERSITY OF COLORADO AT BOULDER (RDEURR@NSIDC.ORG)

[B] CENTER FOR INFORMATICS RESEARCH IN SCIENCE AND SCHOLARSHIP, GRADUATE SCHOOL OF LIBRARY AND INFORMATION SCIENCE, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN (TCHAO@ILLINOIS.EDU)

The *data curation project* was piloted in Fall 2013 as part of the Foundations of Data Curation course, a required class for the Data Curation Specialization at the Graduate School of Library and Information Science (GSLIS) at the University of Illinois, Urbana-Champaign.

The course was conducted in a synchronous online environment with students at different stages of the graduate program who were generally novices to data curation principles and data management.

Project objectives

Each student individually curated a digital data collection by identifying curation needs, designing and implementing a curation plan, and producing products suitable for use by a repository.

ACKNOWLEDGEMENTS:

Special thanks to the Inter-University Consortium for Political and Social Research (ICPSR), Advanced Cooperative Arctic Data and Information Service (ACADIS), National Oceanic and Atmospheric Administration (NOAA), and the Antarctic Glaciological Data Center (AGDC). This project was made possible in part by the Institute of Museum and Library Services, grant #IMLS RE-02-10-004-10.

Data collections for curation

Available collections were categorized as:

- Data upgrade– accessible data and documentation (easiest to curate)
- Original ingest– no/minimal documentation (intermediate)
- Data rescue– limited access to data due to obsolescence or other risk factors (challenging)

Disciplines included:

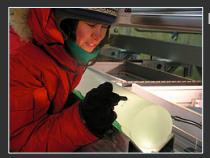
- History (personal collection)
- Local & Traditional Knowledge
- Chemistry
- Hydrology
- Social sciences
- Glaciology
- Climate Studies

Project Milestones

Select data collection

- Collections came primarily from domain repositories including: ICPSR, ACADIS, NOAA@NSIDC, and AGDC
- Report on status of collection based on National Science Board (NSB)[1] digital data collections types

EXAMPLES OF DATA TYPES IN COLLECTIONS FOR CURATION



Ice cores from the West Antarctic Ice Sheet Divide.[2]

Narwhal tusk research integrating interviews with local communities in Canada and Greenland.[3]

Development of draft Curation Plan

- Assess curation needs of available data and documentation
- Establish contact with repository and data producer

Progress presentation

- Share data collection and discuss curation work with colleagues (in-person)
- Additional space provided for discussion via online course website

Final report

- Include new data products and updated Curation Plan based on implementation experiences
- Discuss lessons learned from curation experience

[1] National Science Board. (2005). *Long-lived Digital Data Collections.* Retrieved from http://www.nsf.gov/pubs/2005/nsb0540; [2] Photo credit: Dr. Kendrick Taylor (http://www.waisdivide.unh.edu/Gallery/Best-of-WAIS-Divide.shtml); [3] Photo credit: Glenn Williams (https://eloka-arctic.org/communities/narwhal/index.html)

Success requires…

- Data repository and researcher participation and support
- Instructor background with
  - hands-on data curation experience
  - adequate disciplinary knowledge
  - familiarity with data center requirements

Lessons learned

INSTRUCTOR INSIGHTS:

- Adequate time needed for students to work on the project (2.5 to 3 months)
- Provision of time and space for students to discuss work in progress (synchronous and asynchronously)
- Challenge of assigning 'grades' for work completed

STUDENT INSIGHTS:

- Data curation involves "detective work" and "strategic thinking"; it is a creative, evolving, complex process that encompasses mediation and collaboration
- Curation requires understanding the data and research being done
- There is no one-size-fits-all solution for curation

Improving the project

Need better ways to ensure students share progress and results (i.e., maximize learning). Possibilities for project improvement include:

- Incorporating verbal reports into weekly class sessions
- Integrating additional milestones to mark project progress, which allows instructor to provide additional guidance as needed
- Proposing group rather than individual projects to mirror real-world collaborative work