

A Residential Data Curation Internship: Opportunities and Challenges

Kelly Gordon, Temporary Research Data Specialist

Albert R. Mann Library, Cornell University, Ithaca, NY, 14853, USA

THE MESSAGE: LIS interns can act as “data curation emissaries” to researchers, helping to promote awareness of library data curation services as well as data sharing and archiving practices.



Intern housing, research vessels, and the gazebo at CBFS

THE SETTING: During the internship, I lived and worked with researchers, graduate students, and undergraduate research interns at the Cornell Biological Field Station (CBFS) on Oneida Lake in central New York State. My residence at the station:

- Allowed daily consultation with researchers regarding their dataset
- Promoted awareness of data management issues amongst graduate students and undergraduate research interns via informal discussions and a talk at the summer’s end.
- Facilitated the transition of responsibility for data archiving tasks from library staff to CBFS staff.

OPPORTUNITIES:

For LIS Interns: Day-to-day involvement with research activities promotes a deeper understanding of the data management issues researchers face.

For Researchers: Working with LIS interns leads to a greater awareness of the availability of data repositories for archiving datasets, the services that information professionals can provide to assist with the process, and the tasks and tools involved in preparing a dataset for archiving.

For Data Curation Initiatives: Partnerships forged between library staff and researchers can be a powerful marketing tool in promoting data curation services.

CHALLENGES:

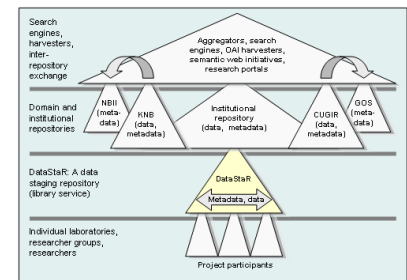
•**Low man on the totem pole:** researchers’ other activities may take priority over intern’s needs to keep project moving forward.
Solution: Good communication and sufficient lead time on requests for help

•**Can open, worms everywhere:** Data needs more cleanup than initially anticipated.
Solution: Clearly delineate which cleanup tasks are beyond the scope of the internship.

•**A rocky transition:** researchers may have trouble finding funding, staff time, expertise to continue data curation tasks after the internship ends.

Partial solution: Continued support from library staff can allay this, but ultimately responsibility rests with researchers!

THE PROJECT: This position was part of the DataStaR (Data Staging Repository) project. DataStaR is an NSF-funded effort by staff at Cornell’s Albert R. Mann Library to create tools and partnerships to facilitate data sharing and preservation by Cornell researchers.



DataStaR’s place in the data curation universe

DataStaR’s principal aims are to develop:

- a staging repository that researchers can use to prepare datasets for transmission to a long-term data repository.
- a set of user-friendly tools with which researchers can create metadata in a number of domain-specific formats

DataStaR personnel have partnered with a number of data owners, including CBFS, to explore the research community’s needs for data curation services and support.

For more info:

<http://datastar.mannlib.cornell.edu/about>

ACKNOWLEDGEMENTS:

Sincere thanks to Gail Steinhart of Cornell University’s Albert R. Mann Library, and to Ed Mills, Lars Rudstam, Kristen Holeck, Tom Brooking, and the staff of Cornell Biological Field Station for their invaluable assistance. This work is supported by National Science Foundation grant number III-0712989.

