

6th International Digital Curation Conference

December 2010

How to Re-Tool Librarians for Data Curation

Sherry Lake (shlake@virginia.edu), Andrew Sallans (sallans@virginia.edu)

Scientific Data Consulting Group
Brown Science and Engineering Library
University of Virginia

ABSTRACT

The University of Virginia Library has begun a process of “re-tooling” subject librarians to meet emerging demands of scientific data management. In leading this new initiative, the Scientific Data Consulting Group (SciDaC) focuses on two main activities: 1) “Data Curation Brown Bag” discussions, and 2) data interviews. Through this process, we hope to gradually help traditional subject librarians develop literacy in issues surrounding scientific data management. Each brown bag session focuses on a very specific topic (i.e. the NSF data management plan policy, the NSF Data Net Program, etc.), offers a short presentation and white paper, and then concludes with an informal discussion.

These sessions are expected to help subject librarians become conversant in the issues and to promote the discussion among their departments and faculty. In parallel, the sessions also prepare the subject librarians as partners for the data interviews. Our interview strategy is based upon the recent interview models from JISC, Purdue/UIUC, Oregon and others, with local customizations. We are performing three interviews per month, and each interview consists of two Scientific Data Consultants, one subject librarian, and the researcher. After several months of leading brown bag discussions and conducting data interviews, we have found that this process is very effective in helping our subject librarians “re-tool” for the new responsibilities. We hope that sharing this model will benefit other institutions as they encounter similar training challenges.

THE NEED FOR ACADEMIC LIBRARIANS TO EVOLVE

The library of the future will need subject librarians who are skilled in the area of data curation. Faced with shrinking budgets and staffs, institutions may attempt to repurpose existing staff to support researchers and their data services needs (Salo, 2010). Anna Gold’s 2007 paper on cyberinfrastructure, data, and libraries, demonstrates that there is a change in academic libraries, with greater emphasis to support science and engineering scholarship. Garritano and Carlson (2009) offer additional examples of how subject librarians should support researchers and their research data. They suggest that librarians need to go beyond introductory articles and reports on data curation to actively collaborate with researchers. They list five sets of skills that librarians need to engage in this new area: 1) library and information science expertise, 2) subject expertise, 3) an ability to build partnerships and develop outreach efforts, 4) a willingness to participate in sponsored research, 5) and an ability to balance workload.

To demonstrate that libraries can provide the needed data curation solutions for researchers, libraries will need to create credible and valuable data services using subject liaisons and

data liaisons (Gabridge, 2009). At the University of Virginia, we have focused on preparing subject librarians for the research data conversation. We have built upon our existing services and developed new skills, roles and organizational structure. Our model does not include adding the full suite of data services to the subject librarian’s list of responsibilities, but instead developing a level of data literacy and then approaching the problems as a team of experts.

UNIVERSITY OF VIRGINIA LIBRARY MODEL

The University of Virginia Brown Science & Engineering Library has been exploring data services models for faculty researchers since 2004. The Library’s Research Computing Lab, in partnership with science and engineering subject librarians, developed a number of service areas that allowed researchers to get assistance with the complete research process (from experimental design to publication and data management) in one single location. (Hunter, Lake, Lee, & Sallans, 2010) The close partnerships between the Research Computing Lab team and Science and Engineering Librarians allowed the researcher to more easily switch between one research activity and another, without having to develop new working relationships and navigate a new set of support services.

In the spring of 2010, the Library decided it was time to go in a different direction and focus more on data management support. This move involved the closing of the Research Computing Lab, and development of a new strategy and plan for scientific research support. The new vision and strategy, falling under the name Scientific Data Consulting Group (SciDaC), builds upon the lifecycle framework, but now puts all the attention on data. Some of the main targets for this new work include support for new data management requirements from funding agencies and publishers, development of institutional repository services, and coordinating policy development.

ROLES IN THE MODEL

UVa Subject Librarian

Serve as subject matter experts and provide opportunities for the scientific data consultants to work with researchers and graduate students. Necessary skills: collaborative, ability to build partnerships, communication, content expertise.

Scientific Data Consultant

One of the main responsibilities of the Scientific Data Consultant is to work with subject librarians to assess maturity of researcher’s data management practices, make recommendations for how to improve, and shepherd process improvement. Other duties include: serve as the Science and Engineering Library’s data specialist and consultant in the areas of data management, metadata production, data organization and

2 How to Re-Tool Librarians

preservation; provide systems analysis and design services for data-driven research projects; create liaison efforts to partner with engineering & sciences faculty & graduate students on their research; and assist researchers writing grant-required data management plans.

CURRENT TRAINING MODEL AND SUPPORT

Brown Bag Discussions

As part of the plan for the data curation initiative, the SciDac group started hosting "Data Curation Brown Bag" sessions every two weeks. Each brown bag has a very specific topic (i.e. the NSF data management plan policy, the NSF Data Net Program, etc.), offers a short presentation and a 1-page white paper, and then allows time for casual discussion.

The objective of these brown bag lunches was to:

- Gradually educate the subject librarians on the pressing issues and trends taking place in the emerging data curation space.
- Provide an opportunity for a short presentation and white paper (1-page, very straightforward overview of the issue), and then casual discussion.
- Help the subject librarian to become comfortable with data curation issues, to then promote the discussion among their departments and faculty as they interact with them.

Brown Bag Topics:

- A High-Level Analysis of scientific Data at the University of Virginia
- The Data Interview Initiative
- Top 10 List for Data Interview
- The Importance of Managing Research Data
- The U.S. National Science Foundation DataNet Program
- Data Interview Methods
- UVA's Institutional Repository Implementation
- New Developments in Citing Data
- First 3 Data Interview Observations

Data Interviews

The goal of our data interviews is to develop an understanding of how our science and engineering researchers manage their research data and initiate a discussion about how to simplify processes and improve practices. Each interview is scheduled for 60 minutes and includes the Scientific Data Consultants, the Subject Librarian and the researcher. A final interview report is then distributed to all Subject Librarians helping to give them a better understanding of research data processes beyond their own fields.

- Identify common research data problems and needs,
- Identify the types of digital "data" that being created,
- Identify communities and individuals who are under the most pressure from emerging grant regulations,
- Identify potential partnerships for institutional repository data deposit, and
- Develop opportunities to provide data management recommendations and training.

Interview models consulted:

- Data Audit Framework (DAF) for background reading (Jones et al., 2008),

- University of Oregon for consultation and information on implementation and buy-in (Westra, 2010),
- University of Glasgow for interview questions and format (Jones, 2008 & Jones, 2009) and
- Purdue's Distributive Data Curation Center (D2C2) for question refinement (Carlson & Witt, 2007).

Our interview protocol was based on the protocol from Wisconsin's *Summary Report of the Research Data Management Study Group* (Wolf et al., 2009).

GOALS AND OBJECTIVES

With our model for re-tooling librarians for data curation, we hope to achieve the following:

- Build data Literacy for subject librarians
- Develop knowledge of how researchers actually manage their data at U.Va.
- Create opportunities to consult and collaborate
- Establish the Library as the place for preservation of data

REFERENCES CITED

- Gabridge, T. (2009). The Last Mile: Liaison Roles in Curating Science and Engineering Research Data. *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC*, (265), 15-21. Retrieved from <http://www.arl.org/bm~doc/rli-265-gabridge.pdf>.
- Garritano, J. R., & Carlson, J. R. (2009). A Subject Librarian's Guide to Collaborating on e-Science Projects. *Issues in Science & Technology Librarianship*, (57), 5.
- Gold, A. (2007). Cyberinfrastructure, Data, and Libraries, Part 2: Libraries and the Data Challenge: Roles and Actions for Libraries. *D-Lib Magazine*. Retrieved from <http://www.dlib.org/dlib/september07/gold/09gold-pt2.htm>.
- Hunter, C., Lake, S., Lee, C., & Sallans, A. (2010). A Case Study in the Evolution of Digital Services for Science and Engineering Libraries. *Journal of Library Administration*, 50(4), 335-347. doi:10.1080/01930821003667005.
- Salo, D. (2010). Retooling Libraries for the Data Challenge. *Ariadne*, (64). Retrieved from <http://www.ariadne.ac.uk/issue64/salo/>.

DATA INTERVIEW/ASSESSMENT REFERENCES

- Carlson, J., & Witt, M. (2007). *Conducting a Data Interview*. Presented at the 3rd International Digital Curation Conference, Washington DC. Retrieved from http://docs.lib.purdue.edu/lib_research/81.
- Jones, S. (2008). Digital Preservation Policy Scoping Study. *University of Glasgow: Digital Preservation Policy Scoping Study*. Retrieved September 9, 2010, from <http://www.hatii.arts.gla.ac.uk/research/prespolicy.html>.
- Jones, S. (2009). Digital preservation study: interview template. University of Glasgow. Retrieved from http://www.gla.ac.uk/media/media_126658_en.pdf.
- Jones, S., Ball, A., & Ekmekcioglu, C. (2008). The Data Audit Framework: A First Step in the Data Management Challenge. *International Journal of Digital Curation*, 3(2). Retrieved from <http://www.ijdc.net/index.php/ijdc/article/view/91>.
- Westra, B. (2010). Data Services for the Sciences: A Needs Assessment. *Ariadne*, (64). Retrieved from <http://www.ariadne.ac.uk/issue64/westra/>.
- Wolf, A., Simpson, M., Salo, D., Flee, D., Cheetham, J., & Barton, B. (2009). *Summary Report of the Research Data Management Study Group*. University of Wisconsin. Retrieved from <http://digital.library.wisc.edu/1793/34859>.