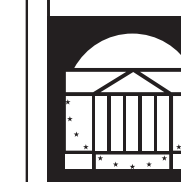# How to Re-Tool Librarians for Data Curation

Andrew Sallans, Librarian for Digital Services, als9q@virginia.edu
Sherry Lake, Scientific Data Consultant, sah@virginia.edu
Scientific Data Consulting Group, Charles L. Brown Science and Engineering Library, University of Virginia, Charlottesville, Virginia, U.S.A.

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.

Background on 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
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
Scientific Data Consultants: Work with subject librarians to assess maturity of researcher's data management practices, make recommendations for how to improve, and shepherd process improvement. Necessary skills: data management, metadata production, data organization/preservation, and systems analysis/design.
Subject Librarians: 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.
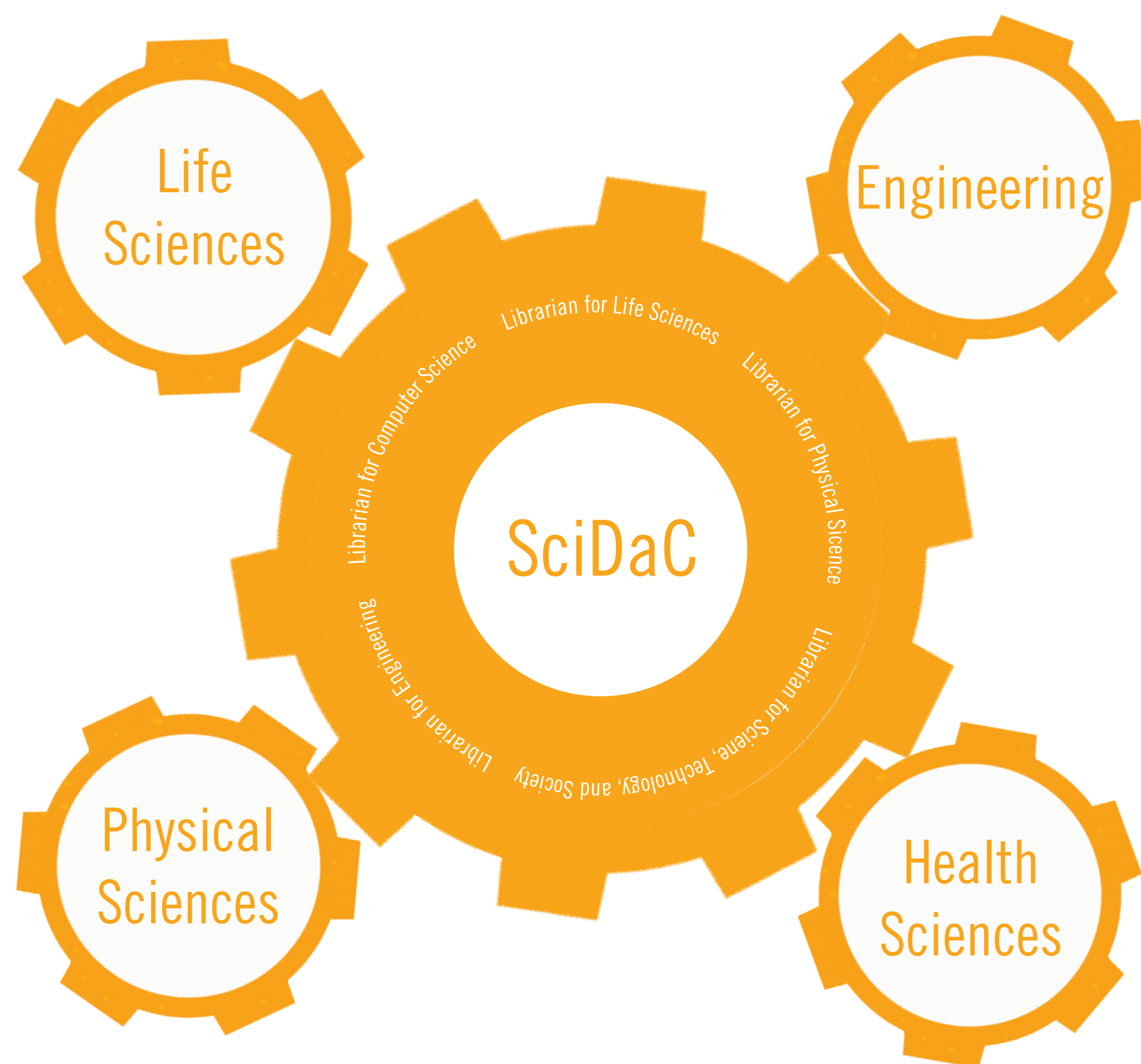
Current Training Model and Support
1) "Data Curation Brown Bag" Discussions: The SciDac group hosts "Data Curation Brown Bag" sessions every two weeks. The objective of these brown bag lunches was to:
a. Gradually educate the subject librarians on the pressing issues and trends taking place in the emerging data curation space (see topics on bar at bottom).
b. Provide an opportunity for a short presentation and white paper (1-page, very straightforward overview of the issue), and then casual discussion.
c. 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.

2) "Data Interviews": Goal 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.
a. Identify common research data problems and needs,
b. Identify the types of digital "data" that being created,
c. Identify communities and individuals who are under the most pressure from emerging grant regulations,
d. Identify potential partnerships for institutional repository data deposit, and
e. Develop opportunities to provide data management recommendations and training.

Goals and Objectives
1) Build data Literacy for subject librarians
2) Develop knowledge of how researchers actually manage their data at U.Va.
3) Create opportunities to consult and collaborate
4) 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.

Salo, D. (2010). Retooling Libraries for the Data Challenge. Ariadne, (64). Retrieved from http://www.ariadne.ac.uk/issue64/salo/.

UNIVERSITY of VIRGINIA LIBRARY

Brown Bag Discussions: Every 2 weeks, building subject librarian data literacy

| A High-Level Analysis of Scientific Data at U.Va. | The Data Interview Initiative | Top 10 List for Data Interview | Importance of Managing Research Data | U.S. National Foundation Data-Net Program | Data Interview Methods | U.Va.'s Institutional Repository Implementation | New Developments in Citing Data | First 3 Data Interview Observations |