Good Reads: Cyberinfrastructure Primer

A. Gold, 2007. Cyberinfrastructure, Data, and Libraries, Part 1. D-Lib Magazine 13(9/10). 13(9/10). http://www.dlib.org/dlib/september07/gold/09gold-pt2.html
and
A. Gold 2007. Cyberinfrastructure, Data, and Libraries, Part II. D-Lib Magazine 13(9/10). http://www.dlib.org/dlib/september07/gold/09gold-pt2.html

The author provides a comprehensive yet straight-forward introduction to scientific cyberinfrastructure for science data from a library perspective, summarizing both library infrastructure issues of today and E-Science promises for the future. Gold provides a timely overview of important elements amidst what appears as an ongoing transition in scientific research data and data practices: “To be able to exchange data, communicate it, mine it, reuse it, and review it is essential to scientific productivity, collaboration, and to discovery itself.”

The reader is provided a path through important community reports that document the ongoing processes: the NSF Atkins report (2003; see Databits Spring 2005 Good Read) to the National Science Board Long-Lived Data report (2005) and finally the Association of Research Libraries (ARL) on Long Term Stewardship of Digital Data Sets. The Task Force on E-Science sponsored by the ARL has taken into consideration the changing roles of libraries and librarians in terms of data and document artifacts. Discussion ranges broadly from supercomputers and grid science to data preservation and curation. Though mention of life cycles and raw data seems to imply near-linear and well-bounded entities rather than complex concepts, the multiple dimensions of infrastructure are covered including data policies and business models. For libraries anchored by physical collections and archive traditions, considering digital data represents an entirely new aspect of their work. An important case is made for a wide range of new roles and supportive services associated with this postulated work with data.

Part 2 of this primer continues with the services offered by data library efforts today: social science, GIS, bioinformatics and archival data services. Gold makes clear that data librarianship requires not only a new framework because it differs significantly from artifact-based collections but also a re conceptualization of roles and collaborative undertakings. She wisely mentions the need to build both capacity and understanding of technology and scaling as well as of new roles. Librarians are stretching from curating physical artifacts to curating digital materials as their focus on cataloging and archiving broadens to include a new service called data management. The primer ends with suggestions for actions that librarians and library leaders may take as they venture into this new territory. This two-part primer includes a suggestion, a tap on the should, to remind the library community to “encourage conceptual dialogue regarding data and informatics efforts”, that is to say, some scholarly thinking to accompany the library plunge into oceans of data.

- Karen S. Baker, CCE-LTER & PAL-LTER