

4th International Digital Curation Conference

December 2008

Purposeful Curation: Research and Education for a Future with Working Data

Carole L. Palmer

Allen H. Renear

Melissa H. Cragin

Graduate School of Library and Information Science,
University of Illinois at Urbana-Champaign

November 2008

Abstract

It has been suggested that digital libraries are more like archives than libraries and that library science cannot provide a theoretical foundation for handling digital data. We assert, however, that library and information science (LIS) provides much of the theoretical framework needed for digital libraries, and for data curation, and is a major source of vital research findings and professional education. This is a natural outcome of the LIS focus on the purposeful collection and organization of information and provision of services to support users, and on the considerable scale of relevant research activities long underway. The need for LIS contributions to the field is evident in results from our current research on scholarly and scientific data, digital collections, and our experiences with the Data Curation Educational Program (DCEP) masters and continuing education activities.

The 4th International Digital Curation Conference taking place in Edinburgh, Scotland over 1-3 December 2008 will address the theme *Radical Sharing: Transforming Science?*. Please ensure you use guidance in this template to produce your paper. Please submit your paper in one of the following formats: Microsoft Word (.doc), Open Office (.odt), or Rich Text (.rtf).



2 Purposeful Curation

Libraries collect, curate, organize, preserve with a purpose—the current and future use of our scholarship, science, history, and heritage. Library and Information Science (LIS) is concerned with understanding and developing the means to accomplish this. Through its theory, principles, research, and professional practice, it builds systems of content, services, and technologies that respond to the needs of users to access and use information of value over the long term.

Ross (2007) has argued that,

...library science has not demonstrated that it has the theoretical foundations and knowledge base that are capable of providing the framework for handling digital entities and for underpinning digital libraries. Moreover, as digital libraries are more akin to archives than they are to traditional libraries we need to seek their theoretical foundations in the domain of archival science and their practices in archival and records management environments (Ross, 2007, pp. 18-19).

However, just like physical libraries, digital libraries exist to support the information needs of user communities, and therefore must be grounded in the core values of librarianship—organizing and providing access to information in ways that add value and enhance use. Moreover, while archival science has much to contribute to a theory of digital libraries, and probably deserves a special place in the list of foundational fields, this does not minimize the contributions being made by LIS. The core theoretical contributions from both of these disciplines are essential for developing high functioning digital libraries, repositories, and curation services.

LIS, in contrast to archival theory, focuses on a scientific understanding of users, access, and functions of information. “The true essence of librarianship... is the maximization of the effective use of graphic records for any purpose... .” (Shera, 1971, p. 57). Its theories and research are aimed at adding value to improve use potential (Taylor, 1986) and coordinating and integrating information in alignment with complex social structures (Shera, 1972).

Understanding what is required to support information use goes well beyond archival science, to, for instance, the sciences of (i) user communities and their information behavior, (ii) data representation and retrieval, and (iii) collection and service development and management, all of which are essential for high functioning digital data curation systems. No one field is equipped with the range of theory and practice necessary for managing the entire lifecycle of digital data, but the large and rapidly growing body of existing work in LIS on scientific and scholarly user communities, interoperability, digital preservation, data modeling, ontologies, digital aggregation, information architecture, sustainability, and others offer a broad foundational base. A glance at the current tables of contents of LIS research journals (JASIST, JDOC, IP&M, and others) will make this immediately evident.

Results from our current research on data curation needs in the digital humanities show the vital contributions to be made by LIS.

- Curation must be informed by our best understanding of how data will be used in research, scholarship, and general access, by current and emerging methods of use.
- More and more, metadata become data and play a particular role in connecting data across domain boundaries as well as object boundaries.
- Curation must scale in ways that accommodate changing formats and data models.
- Markup variation creates interoperability and transformation difficulties.
- We do not yet understand how to manage different levels of data granularity for users.

◆ —◆

These and other information problems are aligned with LIS research and education, as seen in our current studies of data curation and management requirements across the sciences and our Data Curation Educational Program (DCEP), which is focused on curation for scientific and scholarly research data. This program includes a masters concentration and summer institutes for practicing academic librarians and other research data practitioners. These new information professionals will build and maintain not only digital libraries and curated data sets, but also the associated indexing systems, metadata standards, ontologies, and retrieval systems. They will assure that these data systems will work in concert with the many other existing digital libraries, archives, and repositories to support experiments in scientific laboratories, the interpretation of texts by scholars in special collections, the development of exhibits in museums, and other purposeful work with data over time.

References

- [proceedings] Ross, Seamus. (2007). Digital preservation, archival science and methodological foundations for digital libraries. Proceedings of the 11th European Conference on Digital Libraries (ECDL), Budapest (17 September 2007). Available: http://www.ecdl2007.org/Keynote_ECDL2007_SROSS.pdf. Accessed July 24, 2008.
- [book] Shera, Jesse H. (1971). *The Complete Librarian and other essays*. Cleveland, OH: The Press of Case Western Reserve University.
- [book] Shera, Jesse H. (1972). An epistemological foundation for library science. In J. H. Shera, *The Foundations of Education for Librarianship* (pp. 109-134). New York: Becker and Hayes.
- [book] Taylor, Robert S. (1986). *Value-added processes in information systems*. Norwood, N.J.: Ablex Publishing Corporation.