

## **Managing the Digital World: the Role of Digital Curation**

Joyce Ray, Associate Deputy Director for Library Services

For the

United Kingdom's Digital Curation Centre (DCC) Annual Meeting

December 1-4, 2008, Edinburgh, Scotland

We all recognize that the whole world is “going digital.” This phenomenon is affecting every form of the human record, from family photographs to television footage, from music recordings to scientific data. New works are now being created in digital form only, and many of the physical collections held by libraries, archives and museums are being digitized for online access by scholars, students, and the general public. How can we ensure that this global torrent of digital data will be preserved in meaningful form into the future and managed so that information can be found when needed by those who want it?

Until a few years ago, long-term digital preservation received relatively scant attention. The cost of digital storage was continually declining, and many thought it would be possible to “just save everything.” Now we are a little sadder and a little wiser. We have learned that storage costs are not negligible, because even though per-bit storage costs are low, the sheer volume of data being generated—by environmental sensors, for example, as well as by creators of You Tube videos—ensures that storage costs, not to mention capacity, are considerable. We now also recognize that preservation cannot be accomplished by technology alone; we need to create preservation environments to actively manage digital repositories within a framework of policies and procedures. We will need to make decisions about which data to save and which to discard. And we need to invest in good practices to ensure that needed information can be found at the right time. It makes economic sense to share information widely once the initial investment in creation has been made.

Creating pathways to information is like laying a trail of breadcrumbs back to a source. It is essential for long-term preservation, yet it is expensive, especially when human effort is involved. For this reason, information scientists are working to automate as much of the breadcrumb-laying trail as possible. They are also working to improve information discovery for the current and future users who may have completely different needs from those who created the information and who, therefore, will look for it in different ways and need additional paths.

Users will need a variety of tools to enable them to

- search simultaneously across a wide range of disparate resources in order to perform computations, discover relationships, and detect discrepancies;
- drill into mountains of data to find valuable gems of useful information;

- document the provenance and understand the context of information items within a larger collection; and,
- understand their rights and responsibilities regarding data sharing and re-use.

These considerations form the basis of digital curation.

The UK's Digital Curation Centre (DCC) was founded in 2004 as a consortium including the Universities of Edinburgh and Glasgow, in addition to other organizations and centers. **The DCC defines digital curation as “maintaining and adding value to a trusted body of digital information for current and future use; specifically, [it is] the active management and appraisal of data over the life-cycle of scholarly and scientific materials.”** [<http://www.dcc.ac.uk> ]

The Centre notes that “The scientific record and the documentary heritage created in digital form are at risk from technology obsolescence, from the fragility of digital media, and from lack of the basics of good practice, such as adequate documentation for the data.” The Centre aims to be:

- a center of excellence in digital curation and preservation,
- an authoritative source of advocacy and expert advice and guidance,
- a key facilitator of an informed research community, and
- a service provider of a wide range of resources, software, tools and support services.

The DCC has hosted an annual international conference on digital curation since 2005. The third International Digital Curation Conference was held in Washington, DC, in 2007 (IMLS was a sponsor), and the fourth will be held December 1-3, 2008, in Edinburgh, Scotland [see <http://www.dcc.ac.uk/events/dcc-2008/> ].

US institutions have also been actively involved in the development of digital curation. Purdue University, for example, has established a Distributed Data Curation Center to investigate and develop solutions to curation issues of organizing, facilitating access to, archiving and preserving research data and data sets [see <http://d2c2.lib.purdue.edu/index.php> ] In 2007, IMLS funded a research project at Purdue, in partnership with the University of Illinois Urbana Champaign, to develop case studies and curation profiles for researchers in a variety of scientific disciplines. The profiles will identify how different disciplines collect, use and share data, with a goal of developing library services to help researchers manage their active data and of preserving data with long-term research value.

It is apparent that libraries need to develop new services to meet the demands created by the digital environment, and to do this they need staff with skills that have not traditionally been part of library education. Increasingly, libraries, archives and museums find that they need a new generation of information professionals to manage digital resources. This need exists across all types of institutions, from libraries that need to preserve research data or electronic publications; to archives that need to preserve administrative e-mail or student theses and dissertations; to museums that acquire digital artworks, sound recordings, or scientific data. These institutions also have internal data management

needs, from databases to manage information about their collections to digitized copies of content from their physical collections.

To address these growing needs, IMLS has made significant investments in the education of librarians, archivists, museums studies and information professionals. In 2003, we created the 21<sup>st</sup> Century Librarian Program and the 21<sup>st</sup> Century Museum Professional to support professional education and development. In 2005, we added a category within the library program to build institutional capacity in graduate schools of library and information science around the management, preservation and use of digital assets. A number of schools have used this funding to develop new courses and programs of study relating to digital libraries and digital preservation. The following year, we invited applications to develop educational programs in data curation, and new programs were funded at the University of Illinois Urbana Champaign [ <http://www.lis.uiuc.edu/> ] and the University of North Carolina at Chapel Hill [ <http://sils.unc.edu/> ] that have developed masters' degree concentrations and doctoral fellow programs in data curation. Programs have also been established at the University of Arizona [ <http://www.sir.arizona.edu/> ] and the University of Michigan [ <http://www.si.umich.edu/> ]; others are in development. In many cases, new data curation programs are integrated with archival education programs, which are evolving and expanding in response to the deluge of born-digital data.

Recognizing the need to engage museum professionals as well as librarians and archivists in discussing the education of information professionals, IMLS collaborated with the Florida State University (FSU) College of Information and the Ringling Museum of Art to organize a workshop on the education of cultural heritage information professionals in Sarasota, Florida, last April [ <http://chips.ci.fsu.edu/> ]. The principal organizer, FSU professor Paul Marty, worked with a program committee including Michele Cloonan, Graduate School of Library and Information Science, Simmons College; Phyllis Hecht, Museum Studies Program, Johns Hopkins University; Ray Santiato, Miami-Dade Public Library System; Marjorie Schwarzer, Department of Museum Studies, John F. Kennedy University; Helen Tibbo, School of Information and Library Science, University of North Carolina at Chapel Hill; and John Wetenhall, The John and Mable Ringling Museum of Art. Nearly 50 educators and practitioners representing library, archives and museum programs across the US, as well as one Canadian institution, convened to discuss the information needs of institutions and their users and the relevant information skills needed by professionals working in cultural heritage institutions. Given their differences in experience, education, and professional identity, the degree of consensus among the participants was rather remarkable.

Attendees agreed that they have common needs, and that users will be best served if institutions work together to address them. They recognized that information seekers increasingly work online and want information online; in many cases, information that is not online will not be used. Although the public continues to value the physical experiences of museums and libraries, users' information needs cannot be met without an online presence.

Participants noted a growing interest in archival education, and they believed that new archivists and data curators with skills in data management will be valuable to museums as well as to libraries and

archives. A new program at Pratt Institute's School of Information and Library Science, for example, offers post-masters' certificates in archives and museum libraries [ <http://www.pratt.edu/sils/> ].

Actions called for by workshop participants included:

- Develop a better understanding of the relationships that exist among libraries, archives, and museums as well as established and emerging institutions that combine elements of one or more of these traditional venues;
- Encourage research across curricula and encourage practice-based research in academic environments to build a stronger relationship between research and practice and improve understanding between students and faculty;
- Promote interaction among library and information science, archival studies, and museum studies educational programs, as well as among educators and professionals, researchers and practitioners, administrators, faculty and students;
- Recognize that information is central to the work of all libraries, archives and museums, from collection management to engagement with their communities, and build this recognition into academic curricula as well as into research and practice; and,
- Prepare information professionals who can assist their organizations to meet the evolving needs of both internal and external users in the 21<sup>st</sup> century.

With this foundation, and following a planning meeting in Washington, DC, last May with representatives from the DCC, IMLS is sending a number of its grantees to the International Digital Curation Conference in Edinburgh in December, as well as to a working group meeting on education and training following the conference. The grantees are currently working on projects funded under the Laura Bush 21<sup>st</sup> Century Librarian Program to develop curricula relating to digital preservation, archiving, and curation. The working group meeting will explore opportunities to advance the development of international data curation training and education. Watch this space for more information!