ISQ

INFORMATION STANDARDS QUARTERLY

SPECIAL ISSUE: DIGITAL PRESERVATION

DIGITAL PRESERVATION
METADATA STANDARDS

TRUSTWORTHY
DIGITAL REPOSITORIES

UNIFIED DIGITAL
FORMATS REGISTRY

AUDIO-VISUAL
DIGITIZATION GUIDELINES

DIGITAL PRESERVATION
EDUCATION

NISO
How the information world
CONNECTS

OP [OPINION]

A judgment formed about something;
a personal view, attitude, or appraisal

Mary
Molinaro

MARY MOLINARO

How Do You Know What You Don't Know?
Digital Preservation Education

Two Scenarios: Scanning Projects Gone Bad

Imagine this scenario: a curator for a local history museum is approached by the museum director to scan some of the photo collections and make an online exhibit. The museum has a web page and the director suggests the photos be put on that page somewhere. The museum has a flatbed scanner and the curator goes to work scanning. The collection of 100 photographs takes quite a bit of time to scan, but within a couple of weeks the images are scanned. The curator has some experience with webpages and places low-resolution copies of the images on a webpage linked from the museum's main page. The JPEG copies are on the hard drive of the computer attached to the scanner and are numbered sequentially starting with IMG001.jpg. The curator realizes that the images should be preserved and so copies the files onto gold CDs so they will be safe. In reality, the curator clearly does not understand archival file formats, the intricacies of content management systems, issues with file naming conventions, or that CDs are an unstable and impermanent storage media.

In another scenario the Press Association of a medium-sized state is interested in having the state's newspapers made available onl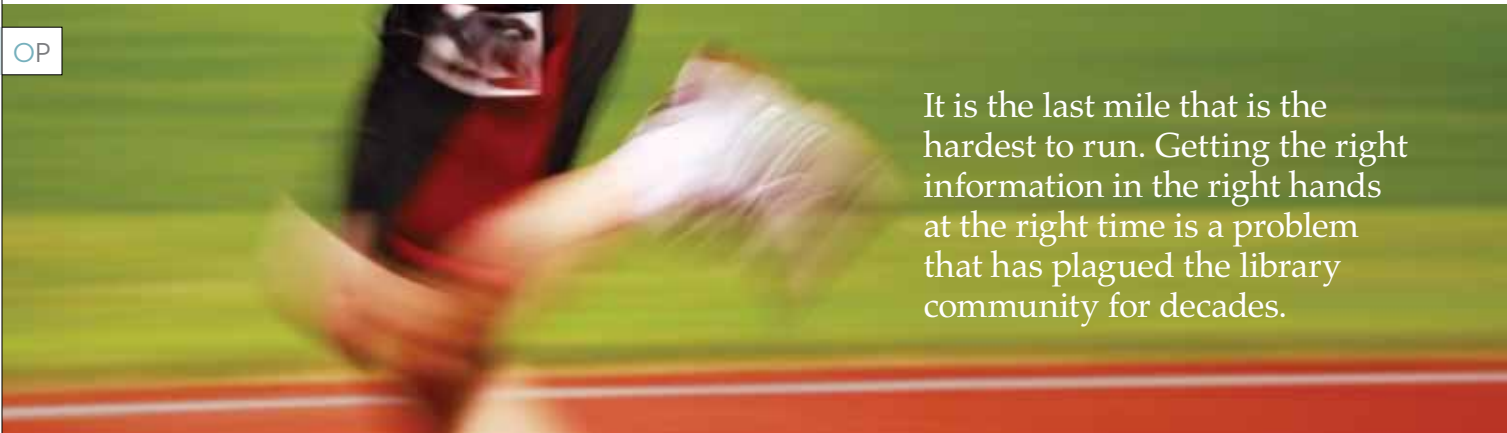ine. They are aware that there are large runs of microfilm in the state historical society. In addition, there are large numbers of other state documents that would also be useful. The historical society recently purchased a state of the art microfilm scanner and has tested it enough to know that the scanner is very fast and very good. When approached by the Press Association about scanning the film, they estimate how long it will take them to scan all 200,000 reels of film and with the new scanner realize that it will not take very long at all. They agree to do the job for $200,000. Once they start the project they quickly realize that the files that are created are quite large; so large they can't afford the storage to store the TIFF images. They also realize that they have not planned for a way to present the pages to users other than as a series of JPEG images. There is also no preservation plan for the images. Rather than go back to the Press Association to re-scope the project, the director of the society decides to do the best they can now and make improvements later—after all it is digital access and it is better than nothing.

In reality, a poorly conceived plan is not better than nothing. Spending limited resources on projects that will have little hope of being sustainable is a tremendous waste that serves no one well. Unfortunately, scenarios similar to these are playing out all across the country. Yes, there are many well thought-out projects with preservation plans in place, but in so many organizations a little knowledge about scanning and webpages can be a dangerous thing. Every institution with responsibility for the stewardship of materials in digital form has some interest in long-term digital preservation. How are the staff members in organizations across the country expected to have the knowledge and skills to ensure that their projects and programs are well conceived, feasible, and have a solid sustainability plan? In short, how does the staff know what they do not know?

CONTINUED »

How are staff members in organizations across the country expected to have the knowledge and skills to ensure that their projects and programs are well conceived, feasible, and have a solid sustainability plan?

It is the last mile that is the hardest to run. Getting the right information in the right hands at the right time is a problem that has plagued the library community for decades.

Overview of efforts

Since 2000 there have been at least two surveys of preservation readiness in cultural heritage organizations in the U.S., both published in 2005. In 2003 Cornell University began surveying 100 institutions participating in its Digital Preservation Management Workshops and in April 2005 the Northeast Document Conservation Center (NEDCC) surveyed 169 cultural heritage institutions on a wide variety of topics related to digitization and digital preservation. Those surveys revealed that barely a third of the respondents had policies in place for the management and preservation of digital content. (As a point of fact, since 2007 the Inter-University Consortium for Political and Social Research (ICPSR) has partnered with Cornell to offer the Digital Preservation workshops and tutorial with support coming from the National Endowment for the Humanities since 2008. They continue to collect data from participants on digital readiness.)

Since those surveys were completed there have been numerous educational opportunities that include significant information in scanning standards and preservation planning that enable practitioners to gain experience in scoping digital projects. Notably the School for Scanning, Digital Directions, and the Persistence of Memory workshops offered by the Northeast Document Conservation Center (NEDCC) and the Digital Preservation Management Workshop formerly at Cornell and now sponsored by the ICPSR.

There are also numerous conferences such as the International Conference on Preservation of Digital Objects (iPRES) or the International Digital Curation Conference (IDCC) that are held for practitioners and center on the topic of digital preservation.

There are some excellent programs being offered by the iSchools around the country to educate new librarians and archivists to teach the skills needed as they move into professions steeped in digital expectations. The programs at the University of Arizona, the University of Michigan, and the University of North Carolina at Chapel Hill are examples of those who are turning out grads who both understand the issues and who will be prepared to lead the way as they move into positions across the country. These new professionals will be highly desirable for the skills they bring to the table in terms of digital acumen.

Additionally, the University of Arizona's graduate certificate program in Digital Information Management (DigIn) and the University of North Carolina's DigCCur program that has the tag line, "Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum" are offering both education and the development of communities of practice for working practitioners. The School of Information at the University of Michigan has a program to create internship opportunities in digital preservation, administration, and curation. All three programs receive support from the Institute for Museum and Library Services (IMLS).

Standards, Tools, and Projects

Over the last decade there has also been movement in terms of standards and best practices for digitization and sustainability. There are a number of resources available to provide guidance for those undertaking digital projects. In fact, when looking for guidance there are many "imaging guidelines" available from a wide variety of organizations. But most are highly technical and many are out of date. It is understandable if people actually undertaking digital projects set these aside in favor of the manual that came with the scanner or the advice of a well meaning colleague. One must know that their current practice is lacking to even look for improvements.

Additionally, a great deal of effort has gone into the development of tools to assess the strengths and weaknesses of existing repositories of digital data. Of note are the Trustworthy Repositories Audit and Certification (TRAC) criteria and checklist that was developed by RLG and the National Archives and Records Administration (NARA) and the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA). These tools can be used for repository planning as well as assessment.

The National Digital Information Infrastructure and Preservation Program (NDIIPP) has as its mission "to develop a strategy to collect, preserve and make available significant digital content, especially information that is created in digital form only, for current and future generations." With this mission comes the realization that this will require effort at the local level so that material is available to preserve. NDIIPP is funded by Congress and is leveraging the weight of the Library of Congress to begin to mobilize at a local level. Partners across the country have been involved in many worthy and important initiatives including MetaArchive, the Internet Archive, LOCKSS, and Portico. With the Digital Preservation website, NDIIPP presents an excellent set of resources for librarians, archivists, and the public.

The Last Mile

With all of these resources, why is it that the librarians in the local libraries are still making grave and costly errors in building sustainable collections when asked to do a "scanning project"? Why are the tools and standards being largely ignored at the most basic level in so many institutions across the country? How is it with the wealth of information and training available to the library community that it seems so elusive to so many people?

I believe the analogy often given to describe so many types of projects is very true in this case: it is the last mile that is the hardest to run. Getting the right information in the right hands at the right time is a problem that has plagued the library community for decades. When adding in the incredible pace of change in the digital environment, limited resources for training and travel, and work days that are already overburdened, it is not surprising that at the local level people forge ahead on projects blissfully unaware of standards and best practices.

In this area of rapid change it is those who are already heavily involved in the development of the tools and services that are best able to leverage their use. Unfortunately there are still vast numbers of people and project managers who have no idea of where they should even start.

Moving Forward: A call to action

Something has got to give. If we have any hope to preserve the digital record of our lives and collections there must be a coordinated effort that takes advantage of the years of work that has been put into the development of the practices that will provide the best shot at sustainability. People at the local level must be encouraged and supported to represent their collections and communities in a digital form that has a very good chance to persist over time. We must leverage the expertise that exists and make it easy for people at the local level to know what to do.

To this end the Library of Congress, through the NDIIPP program, is taking a leadership role once again. Initial steps have been taken to establish a broad-based education program to reach practitioners across the country through a program dubbed Digital Preservation Outreach and Education (DPOE). This program is in the planning stage, but the idea of taking training and education for digital preservation into the heart of the country will make all the difference in empowering the front lines in the fight for sustainability of our digital heritage. Updates will be available on the NDIIPP website as the planning unfolds.

It is important for those who are knowledgeable to participate in ways that will make a real difference. Partnerships and collaborations will fit hand in glove with education programs offered at the local level. In the digital preservation community we have talked around these issues for many years. Increased visibility at a local level supported by national organization will finally make it possible for all of the talk to become reality. | OP | doi: 10.3789/isqv22n2.2010.08

MARY MOLINARO <molinaro@uky.edu> is Director, Preservation and Digital Programs, in the University of Kentucky Libraries.

Clareson, T. "NEDCC Survey and Colloquium Explore Digital Preservation Policies and Practices." RLG DigiNews Feb 15, 2006.
worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file1339.html#article1.

DigCCurr, University of North Carolina at Chapel Hill,
www.ils.unc.edu/digccurr/

Digital Curation Centre, Digital Repository Audit Method Based On Risk Assessment (DRAMBORA).
www.dcc.ac.uk/resources/tools-and-applications/drambora

Digital Preservation Workshop, Inter-University Consortium for Political and Social Research (ICPSR)
www.icpsr.umich.edu/dpm/

International Conference on Preservation of Digital Objects (iPRES) 2010
www.ifs.tuwien.ac.at/dp/ipres2010/

International Digital Curation Conference (IDCC), 2010 (6th)
www.dcc.ac.uk/events/conferences/6th-international-digital-curation-conference

Kenney, A. and E. Buckley. "Developing Digital Preservation Programs: the Cornell Survey of Institutional Readiness, 2003-2005." RLG DigiNews, Aug 15, 2005.
worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file1088.html#article0

National Digital Information Infrastructure and Preservation Program (NDIIPP), Library of Congress.
www.digitalpreservation.gov/library/

National Digital Information Infrastructure and Preservation Program (NDIIPP) Initiatives
www.digitalpreservation.gov/library/initiatives.html

Northeast Document Conservation Center (NEDCC) workshops
www.nedcc.org/education/workshops.introduction.php

Trustworthy Repositories Audit and Certification (TRAC) criteria and checklist, Chicago: Center for Research Libraries; Dublin, Ohio: OCLC Online Computer Library Center, Inc., 2007.
www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying

University of Arizona, Graduate Certificate Program Digital Information Management
digin.arizona.edu/index.html

RELEVANT LINKS