Library of Congress Digital Preservation Newsletter

Digital Preservation Pioneer: Nan Rubin

Since its start in 2005, the NDIIPP Preserving Digital Public Television Project has made significant progress in building infrastructure, creating standards and obtaining resources, and much of the credit goes to its project director, Nan Rubin.

Rubin has been a social activist since her teenage years and is accustomed to fighting for causes, especially empowering under-represented people through broadcast media. "I've really focused on trying to expand people's voices using the media that's been available," Rubin said. She started community radio stations in Cincinnati, Ohio, and Denver, Colo., and has been a radio producer for 30 years.

Her work eventually led her to public television station Thirteen/WNET New York. A project to move the station's 60,000 videotapes resulted in creating an archive for the station, which was a revelation for Rubin. "I realized how important it was to community groups to save their legacy and put it online and make it available to people in new ways," she said. Her passion for preservation extends to other projects, including helping a large network of bilingual public radio stations in California develop an archive to make 20 years of news programs available on-line.

Preserving Digital Public Television focuses on creating a consistent approach to digital curation among stations that produce national programming for PBS. In addition to Thirteen, there are three other partners in the project: PBS; WGBH in Boston; and New York University.

The project has produced new tools and improved practices. "We've chosen a set of metadata schema that includes four elements: PBCore (a standard developed by and for public media organizations), METS rights, MODS and PREMIS." Rubin has also facilitated several successful data transfer tests with the Library, and is working to develop transfer tools.

The goal is ultimately for PBS and other national content to arrive in digital form at the Library's National Audio-Visual Conservation Center. There the content will be preserved on servers and data tapes. This goal is a key element of the American Archive, a new national initiative which will support digital archiving and access for public television and radio programs nationwide.
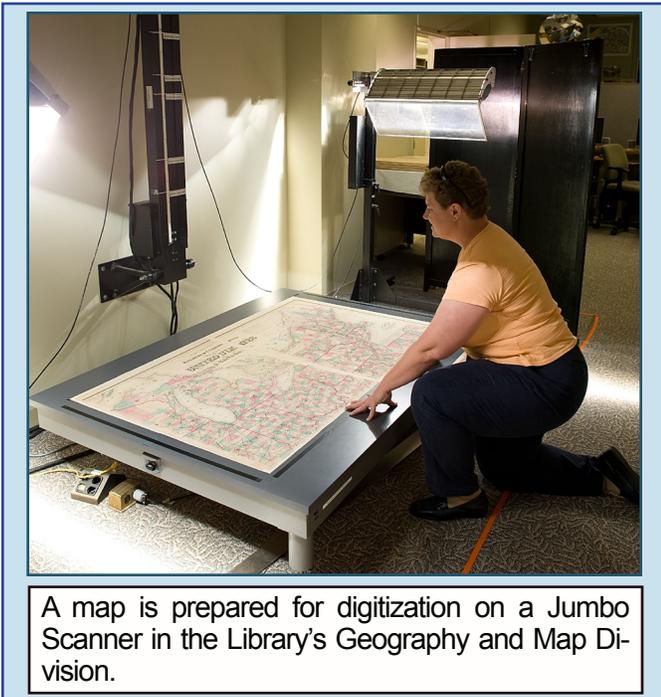
Rubin has high hopes for the American Archive but she is concerned about getting equal representation for all public broadcasters. "WGBH and Thirteen create 60 percent of the national programming seen on public television, but the American Archive has to include small stations as well as independent and minority producers in both television and radio," she said. "I consider it a really important to represent those community producers. And they're not going to be brought to the table unless they have advocates. It's not just Great Performances and Nova that need to be preserved; it includes all these other parts of public broadcasting – local stations, independent and small producers, who represent the heart of our rich collective program sources."

Many stations are unsure about what to do with their programs for the long term and Rubin promotes the American Archive as a solution. Still, she is frank about the many challenges that must be addressed, and knows that it will take time and effort to bring about change. But Rubin also has faith that the public-broadcasting world will recognize the achievements of the NDIIPP Preserving Digital Public Television Project and will elevate digital preservation as a priority. "We need to demonstrate how it's really important to production, to be able to get your stuff later, and that it can be integrated without being overly burdensome or unmanageably expensive," Rubin said. "We're working on that now."

Read the complete profile of Rubin at: http://www.digitalpreservation.gov/partners/pioneers/detail_rubin.html

LIBRARY OF CONGRESS

Meeting the Challenge: Federal Agencies Collaborate on Digitization Guidelines

A map is prepared for digitization on a Jumbo Scanner in the Library's Geography and Map Division.

The Library of Congress is one of twelve federal agencies participating in a new government collaboration to develop a common set of digitization practices and guidelines.

To deal with different types of formats, two working groups have been formed. A still-image working group will establish appropriate guidelines for the scanning of text, maps, photographic prints and negatives. An audiovisual working group will set standards for digitizing audiovisual materials—sound recordings, videos and film. They will post their recommendations on two Web sites:

www.digitizationguidelines.gov/stillimages/ and
www.digitizationguidelines.gov/audio-visual/

These guidelines, which cover such criteria as evaluating image characteristics and establishing metadata elements, will serve not only the agencies, but also digitization service providers, equipment manufacturers, and other technologists. The products of the working groups are not considered to be static, but will continue to be updated and posted to the Web site.

Although the Library has served as convener during the formative stages, no one institution is a leader in this collaborative effort. Joining the Library are the Defense Visual Information Directorate, National

Agricultural Library, National Archives and Records Administration, National Gallery of Art, National Library of Medicine, National Transportation Library, Smithsonian Institution, U.S. Geological Survey, the U.S. Government Printing Office and the Voice of America.

"All the government agencies are in a similar situation; there are certain types of materials that you know exactly what to do with," said Steve Puglia, preservation and imaging specialist for the National Archives, and involved in this effort. "In the still-imaging area, there are certain baseline approaches that already work well. One challenge we are currently facing is to create good procedures for quality assurance and quality control. Fixing things upfront in the process is really important, and will result in the best-quality final product."

In addition to the guidelines, other resources are available on the Web site. One of these is a glossary of digitization-related terms—a wide-ranging list of more than 200 terms and definitions, from common terminology, such as megapixel, to the more arcane, such as scotopic sensitivity. In addition, more than 500 other terms have been identified for future updates, so look for the list to grow into a comprehensive one-stop glossary for digitization.•

International Data Archive Technology Alliance

The Library of Congress hosted a meeting on October 14-15, 2008, of social science data stewards from seven nations to discuss forming a Data Archive Technology Alliance. Organizers of the meeting included the Interuniversity Consortium for Political and Social Research, which is the lead partner for the NDIIPP Data-PASS project. Read more at :

http://www.digitalpreservation.gov/
news/2008/20081028news_article_
DataTechAlliance.html

To subscribe to this newsletter, go to https://service.govdelivery.com/service/multi_subscribe.html?code=USLOC&origin=http://www.loc.gov type in your e-mail address, scroll down and click on "Digital Preservation." Past newsletters are archived in PDF/A at: http://www.digitalpreservation.gov/news/archive.html