Library of Congress Digital Preservation Newsletter

State Information Executives "Fly-In" to the Library

April 30, 2008 -- Chief information officers and staff from 10 states came to the Library on April 30 to talk about the NDIIPP effort to help preserve state government digital information. The National Association of State Chief Information Officers arranged the visit as part of its annual Fly-In; an initiative that helps state information technology executives discuss critical policy matters with federal officials in Washington.

NDIIPP staff provided details about the status and goals of the four recently launched states projects. There was also a lively discussion centering on the role of CIOs in connection with long-term access to digital content.

"We have always seen state IT executives as essential players for digital preservation in the states," said Bill LeFurgy, NDIIPP states initiative manager. Doug Robinson, NASCIO executive director and Fly-In coordinator, agreed. "We all know that CIOs are playing a much more visible role in managing the life cycle of electronic records, including their preservation," he said.
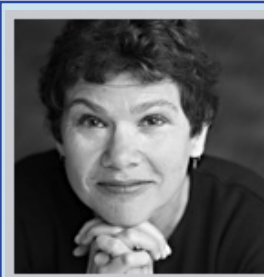
Several CIOs expressed strong desire for improving digital preservation in their states and called for enhanced partnership with libraries, archives and other entities with cultural heritage mandates. Other attendees noted a need for improved tools and services to conduct preservation activities. All agreed the Library's work with the states was providing a major boost to an area in critical need of attention.

CIOs and staff from Alabama, Hawaii, Massachusetts, Minnesota, Ohio, Oregon, South Carolina, Washington and West Virginia attended the meeting.

For more further information on the NDIIPP States Initiative, see ▪ www.digitalpreservation.gov/partners/states.html.

Digital Preservation Pioneer: Fran Berman

Fran Berman, director of the San Diego Supercomputer Center

Dr. Francine Berman's interest in digital preservation is as personal as it is professional; in the same breath she might muse about curating exabytes of NASA data and preserving photos from cell phones.

Fran is the director of the San Diego Supercomputer Center at the University of California, San Diego, which hosts both powerful computers and America's largest academic data center. For decades SDSC has played a major international role in scientific and engineering research. In past years the center has become an ally with the world of archives and libraries. Information technology is blurring institutional boundaries.

Until recently, scientific bodies such as NASA have been the undisputed data archiving heavyweights. But many archives and libraries now work routinely with terabytes of digital content and data volumes are rapidly increasing. The cultural heritage community needs partners like Fran to help tackle digital data stewardship on a massive scale. And she is eager to help.

In past interviews, Fran referred to 2000-2010 as the "data decade." Now, in 2008, she reflects on the events of the past eight years and stresses the need for sustainable digital preservation. "Digital preservation is critical for research and education in the information age." she said. "The research community is cognizant that we need to preserve our most important digital collections – from nightly surveys of the skies in the National Virtual Observatory collection to the critical longitudinal information on families in the Panel Study of Income Dynamics. The conversation is moving beyond 'why is it important to preserve data?' to sustainability issues: 'how will

LIBRARY OF CONGRESS

we do it?' and even more challenging: 'how will we pay for it?'"

From 2006 to 2007, the Library of Congress sought answers to these questions by collaborating with SDSC to test data transfer and storage. Since then other NDIIPP partners have conducted similar test projects and each project has resulted in a step forward in the creation of the infrastructure – or, more accurately, the cyberinfrastructure – required for the large-scale data transfer, storage, retrieval and interoperability among stewardship institutions.

This work is similar to laying railroad track across the frontier or stringing telephone lines into homes. "It is an interesting time," Fran said. "People often think that the infrastructure is the boring stuff. But all of this 'boring stuff' has to happen in order for you to do what you may consider the exciting stuff. For you and I to work in our offices, somebody had to make sure that we were wired for reliable lighting and the light bills would be paid. But the fact that the lights are working is non-memorable. They are just part of a functioning environment.

"In the information age, our digital data needs to be accessible where and when we want it, which means that data preservation infrastructure has to become a basic component of our cyberlandscape. To be part of this cyberlandscape, someone has to make sure that the 'data bill' is paid. Funding for data cyberinfrastructure in the cyber-world is still not assumed or incorporated in the same way we incorporate funding for physical infrastructure in the natural world."

SDSC and its partners the UCSD Libraries, the National Center for Atmospheric Research and the University of Maryland Institute for Advanced Computer Studies have developed a digital-preservation approach called Chronopolis, based on a data grid (a system of sharing and managing data distributed among participating computers). Current Chronopolis collections include copies of the Inter-university Consortium for Political and Social Research (ICPSR) and California Digital Library collections. "Multiple copies of collections are critical," Fran said, "to mitigate the risk of data loss due to error, power outage, natural disaster or other unforeseen circumstances at one of the sites."

"In addition to digital research collections, many of us also care about saving digital family photos and the contents of our hard drives, although we may not think of this as digital preservation," she said. "Virtually everyone cares about preservation of some digital entity but we have not achieved the broad societal awareness of the issues of digital preservation in the same way that the general public is aware of other issues like global warming or stem cell research."

Part of Fran's job is to provide SDSC resources to enable the effective use of data for research and education. She loves problem solving and working with researchers to use technologies in new, interesting, effective and empowering ways. "Leading SDSC has given me the opportunity to come to work every day and make a difference for the research community," she said. "SDSC's terrific staff and powerful cyberinfrastructure have helped scientists do everything from creating new approaches to the design of cancer drugs to understanding how the Universe formed after the Big Bang. It's also introduced me to the world of digital data and helped create an abiding intellectual passion for its issues."

As for digital preservation in general, Fran's vision is both grand and rooted in reality. "Data is the natural resource of the information age," Fran said. "Data is fragile and needs to be stewarded in the 'cyberworld' just like we need to take care of rain forests and the environment in the physical world. Preserving valued data in the information age is fundamental to ensure that it will continue to inform and enrich our world for the foreseeable future."

Read more about Fran's pioneering vision and work at: http://www.digitalpreservation.gov/importance/pioneers/detail_berman.html. ▪

2008 NDIIPP Partner Meeting

In early July, NDIIPP partners will meet in the Washington, DC area to share updates on their work, learn about newly developed tools and services, meet new partners in the NDIIPP network, and discuss policies to ensure institutional commitment to providing long term access to digital materials. A full report on the meeting's outcomes will be available on http://www.digitalpreservation.gov in mid-July. ▪

To subscribe to this newsletter, go to http://service.govdelivery.com/service/subscribe.html?code=USLOC_13, type in your e-mail address, scroll down and click on "Digital Preservation."