

POLICY

RECORDS MANAGEMENT

NARA faces daunting records task

Agency plan to create electronic records archive still a work in progress

BY DIANE FRANK

The National Archives and Records Administration is taking on a challenge that has no guaranteed solution — attempting to preserve every electronic record the federal government has ever created.

The Electronic Records Archives program is forcing officials and industry to contemplate storage of enormous amounts of information, as large as terabytes' and petabytes' worth of data.

digital photos taken in the 1990s to geospatial information systems that will be designed in 2030. Furthermore, NARA has to maintain the original look, feel and context of the records.

This task is complicated by rapid changes in technology and a lack of understanding about what records actually are, said Kenneth Thibodeau, Electronic Records Archives program director.

"We need to find a way to free electronic records from the hardware and

"When we figure out what the system will be, we will go out to industry and ask them to build it," Thibodeau said.

Observers, including the General Accounting Office, which is watching the electronic archive program very closely, are "positive but skeptical" that the program can be successful, said J. Timothy Sprehe, a records management expert and president of Sprehe Information Management Associates Inc.

"It's just mind-boggling," he said. "You have to take as a fundamental premise that everything is changing and will continue to change."

The idea being discussed now is to

take a page from the Extensible Markup Language community's playbook. During the past several years, various communities of interest have developed special XML schemas for particular uses, such as e-commerce transactions and legal information.

NARA officials believe part of the solution is to store information in basic templates, which provide standardized ways of describing the context

and presentation of records, said Dan Jansen, a project manager for the electronic archive program.

Although NARA will establish the basic terminology, agency officials plan to ask different user communities to help refine that language for particular kinds of records, Jansen said.

That terminology will be the language the system uses to communicate, but the electronic archive staff is still working to develop the system itself. The idea is to create "virtual workspaces" that will provide storage, backup, security, and the ability to search and access the information from any location or user system, whether someone is on the move, in an office or at a NARA facility, Thibodeau said. ■

SEARCHING FOR TECH

The National Archives and Records Administration's Electronic Records Archives program is looking for technologies that officials say do not currently exist to create the archives of the future. To do so, NARA is working with research projects led by other agencies and organizations, including:

- The Digital Libraries Initiative, led by the National Science Foundation.
- The Open Archival Information System reference model project, led by NASA.
- The International Research on Permanent Authentic Records in Electronic Systems project.



But the real challenge is not just storing that information, but being able to make those records accessible in the same form they were originally created.

"Records created just a few years ago are already unreadable by today's technology," U.S. Archivist John Carlin said Nov. 8 at the Electronic Records Archives users conference. "Building [the archive] is not and will not be easy, but there is simply no alternative."

Anyone who has tried to open a document created in a word processing program from the 1980s already has an idea of one part of the problem. But NARA's problems go well beyond that. The agency has to preserve and allow access to every type of record stored, from punch cards generated in the 1950s to

software that created them," said Reynolds Cahoon, assistant archivist for human resources and information services and chief information officer at NARA. "As yet, there is no known on-the-market solution."

The development of the system that will perform this task started in 2001. Thibodeau said NARA is asking for ideas from outside the agency, namely by getting feedback from the user conference, issuing requests for information to industry, and working with government and private-sector researchers.

But officials do not expect to award the contract until fiscal 2004, and initial deployment is not envisioned until fiscal 2007. The electronic archive solution is still very much under development.

ILLUSTRATION/KEVIN CURRY