

FORMAL STATEMENT

L. REYNOLDS CAHOON

**Assistant Archivist for Human Resources and Information Services,
and Chief Information Officer
National Archives and Records Administration**

before the

**Subcommittee on Technology Policy, Information Policy,
Intergovernmental Relations, and the Census
of the
Committee on Government Reform
U. S. House of Representatives**

July 8, 2003

I am L. Reynolds Cahoon, Assistant Archivist for Human Resources and Information Services, and Chief Information Officer of the National Archives and Records Administration (NARA). I wish to join the Archivist in thanking you, Chairman Putnam, for recognizing the importance of Electronic Records Management and holding this hearing.

Two Perspectives on Records Management

As the Chief Information Officer of NARA, I have to maintain a dual perspective on the needs and responsibilities of this agency. In one direction, I have to see to it that NARA has the information technology that enables it to accomplish its mission and serve the citizens. In the other direction, I must take into account NARA's government-wide responsibility for leadership in the lifecycle management of records, which are critical information assets needed to make the U. S. Government work, to ensure that it is accountable to the citizens, and to document our national experience.

These two perspectives must coalesce in a comprehensive vision encompassing both how NARA accomplishes its own business and the direction NARA gives to other agencies. Only concerted, complementary efforts in both directions will enable the Federal Government to address the challenges posed by electronic records and make the transition to e-Government in a way that takes full advantage of the opportunities offered by networking and information technology while preserving what is best in our democratic traditions.

While my responsibilities as CIO focus on issues and topics related to information technology, there is not a shadow of a doubt that the goal of my efforts, and those of my staff, is not to acquire or deploy technology, but to help realize the contribution that effective records management can make in the transition to e-Government, in interagency

and intergovernmental collaborations, and in the value that American citizens can realize in the information assets of their Government.

Legal Mandates for Records Management

The Archivist has described NARA's need to preserve and provide sustained access to key records of the U.S. Government in the National Archives and the Presidential libraries. Equally daunting is NARA's responsibility for guiding other agencies' management of the electronic records they create. This guidance must provide a sound and comprehensive basis for achieving the objectives of records management as set out in the Federal Records Act. Those objectives do not make records management a goal in its own right, and they do not divorce the preservation of historically valuable records from the on-going performance of the Government's business. The objectives of this law include:

- Accurate and complete documentation of the policies and transactions of the Federal Government;
- Effective and economical government operations;
- Simplification of activities, systems, and processes; and
- Prevention of unnecessary Federal paperwork; as well as
- judicious preservation and disposal of records. (44 U.S.C. 2902)

While these objectives were set out in the Federal Records Management Amendments of 1976 (H.R. 13828), they resonate well with key purposes of the E-Government Act of 2002. Among other worthy goals, the E-Government Act aims

- To improve the ability of the Government to achieve agency missions and program performance goals.
- To reduce costs and burdens for businesses and other Government entities.
- To promote better informed decisionmaking by policy makers.
- To make the Federal Government more transparent and accountable.
- To transform agency operations by utilizing, where appropriate, best practices from public and private sector organizations; and
- To provide enhanced access to Government information and services in a manner consistent with laws regarding protection of personal privacy, national security, records retention, access for persons with disabilities, and other relevant laws.

Government performance, decision-making, and accountability, as well as access to Government information all depend on how well the Government creates, retains, and manages the records that document its decisions and performance; in other words, how well agencies achieve the objectives of Federal records management.

NARA's responsibility for government-wide direction in records management makes it imperative for us to ensure that our guidance on managing electronic records is appropriate and effective. This is a strategic requirement which we address with as much seriousness and commitment as we apply to the task of preserving key electronic records of the U.S. Government for the citizens. In addressing this strategic requirement, we

must remember that what makes something a federal record has nothing to do with its form – whether it’s a report or a web page, a letter or an email – and nothing to do with the information technology used to create, communicate or store it. What makes it a government record is either its connection to the conduct of government business or the value of the information it contains.

The often intricate connection between the technology and electronic records and the potential for improved service to the citizen and better return on investment offered by appropriate selection and use of information technology, compels us to adapt and enrich our records management guidance to make it highly useful to individual agencies as they implement IT solutions, and to align it with the government-wide objectives of the E-Government Act, the Government Performance and Results Act, and other key statutes.

Responding to the challenges of electronic records

Our response is comprehensive, multi-faceted, and future oriented. Like the dual vision I mentioned earlier, this orientation aims both at delivering electronic records to future generations and at positioning electronic records management as an integral element in effective e-Government. It is essential to tackle both of these objectives in parallel. The solutions NARA adopts for preservation and sustained access must complement the solutions other agencies implement to improve the conduct of the Government’s current business.

NARA’s actions in the short-term to improve the lifecycle management of electronic records, from the moment of creation, are critical for its long-term goal of ensuring that the history of our time is accessible to future generations through authentic records. Without neglecting the importance of transmitting those electronic records that already exist to future generations, continuing expansion of the application of IT in government, combined with continuing expansion in the kinds and the complexity of information recorded in digital form mean, that we have to anticipate that the challenges entailed by electronic records will explode to a level that makes everything we have faced up to now look small by comparison.

NARA looks for a coherent solution to the challenge of electronic records because both electronic records management to support current business and archival preservation of electronic records for the long-term face the same, fundamental technical problem: the need to make digital information assets independent of the specific hardware and software used to process, store, or communicate them at any time. In the archival context, this problem manifests itself over time in the twin problems of obsolescence and progress. In the context of e-Government, this problem manifests itself in the need for systems to interoperate across agency lines, and more broadly for agencies to interact electronically with citizens, with businesses, and with state, local, and tribal governments, and with other nations.

For e-Government to achieve the goals which the Congress and the President have established in the areas of efficiency and timeliness, it is necessary that the systems involved in a transaction be able to interoperate and automatically exchange required

information. Both far-reaching and fundamental differences in requirements and simple differences in the frequency and timing of when institutions and individuals acquire and replace technology make it extremely unlikely that all the stakeholders with whom the U.S. Government needs to interact will ever have fully compatible systems. Moreover, developments in information technology make that unnecessary. For example, an online transaction in e-government might involve, on one end, an agency application running on a sophisticated database management system and, on the other either a company using a system that is fine-tuned to its business, or a citizen using a personal computer.

Rather than requiring that all parties use the same, or compatible, systems, the technology is providing generic products that translate among disparate systems. Notably, many commercial products implement the eXtensible Markup Language (XML). The XML standard provides a simple, open way to identify what kinds of data are being exchanged online. For different systems to participate in electronic transactions, each one only needs to recognize the XML markup of the data it receives, and translates the data into its own internal format. Conversely, it needs to be able to export its data in the shared format. As predicted by industry analysts 5 years ago, this mediated approach is proving to be one of the major enablers of e-business and e-government.

The Need for Authenticity

But for quality of service, and ultimately for the trustworthiness of the Government, information assets must not suffer any diminution in authenticity or reliability as they go through repeated transformation from one internal format, to an intermediate common format, and to another internal form in the course of a transaction. Effective records management is in fact the discipline which can ensure that authenticity and reliability remain intact in transmission of information across time, space, and institutional and technological boundaries, and it can also contribute to timeliness and efficiency.

NARA is uniquely positioned and legally required to ensure that federal records management does make this contribution to e-government. Of itself, records management cannot satisfy the requirements for authenticity and reliability, but it provides a well-established framework for defining the requirements, for evaluating the solutions that technology offers, and for ensuring that the real-world results achieve their objectives.

In broad terms, we regard the mediated approach that technologists have developed to solve the problems of sharing information in real time as the most viable approach for sustained access to authentic electronic records over the long-term. No one can pretend to know what information technology will be like 25, 50, or more years in the future. But it is a reasonable assumption that Americans will want to use the best technologies available to them to access the electronic records of their Government, regardless of when those records were created. In effect, the solution NARA implements to provide sustained access to these records for the long-term must mediate between the myriad digital formats of federal electronic records that exist now, and that will be created in years to come, and the technologies that best serve the needs of citizens at any time in the future.

The most basic requirement for the architecture and the technology that NARA adopts to accomplish its core mission is that the system be able to evolve over time, keeping pace with progress in IT, and responding to citizens' expectations for the best available service. Building this solution will not be easy. Based on market research conducted over the past two years, an extensive dialogue with the IT industry, and continuing collaborations with leading researchers in computer science and engineering, we have learned that key technologies enabling such a solution are available today, and that it will be possible to develop a full solution gradually over time.

NARA's Initiatives on Electronic Records Management

NARA's comprehensive approach involves three major initiatives: the Records Management Initiatives, the Electronic Records Management project in the Administration's e-Government portfolio, and the Electronic Records Archives Program, as well as a number of other activities. It starts close to home with an open, everything-on-the-table review of our records management guidance.

Under the leadership of the Deputy Archivist, Dr. Lewis Bellardo, our Records Management Initiatives aim to create mutually supporting relationships with agencies, whereby

- records management is viewed as a normal part of asset and risk management;
- records are managed effectively for as long as they are needed;
- NARA's records management program adds value to the agency business processes; and
- records of archival value are preserved and made available for future generations.

We will adapt our guidance to the digital environment; make it adaptable by agencies to suit the specific needs entailed by their missions and the different constituencies they serve; and we are supplementing our government-wide guidance with targeted assistance to individual agencies.

Government-wide Records Collaborations

We are coordinating our efforts with other government-wide initiatives in furtherance of the President's Management Agenda, especially his e-Government goals, notably with the Office of Management and Budget (OMB) and the Federal CIO Council. Our work with the Federal Enterprise Architecture (FEA) Program Management Office in OMB is reflected in the position of records management in the June 2003 release of the Business Reference Model (BRM) of the FEA.

Lifecycle management of records is positioned in all four of the main areas of business defined in the BRM. In the Services for Citizens area, it appears under the heading of "Cultural and Historic Preservation." In the Mode of Delivery area, it is part of "Information Infrastructure Management." The Support Delivery of Services area includes "Central Records and Statistics Management" under the General Government

line of business. And the Management of Government Resources area includes Records Retention as a sub-function of Information and Technology Management. Thus the Business Reference Model simultaneously promotes records management and positions records management to contribute to e-Government.

Our coordination is not only at the intellectual level of elaborating the framework of the Federal Enterprise Architecture. It also has practical elements. Working with other agencies, NARA has undertaken four Electronic Records Management projects in the Administration's e-Government portfolio, providing specific examples of the contributions of good records management to e-Government.

The Electronic Records Management Initiative will provide the tools that agencies will need to manage their records in electronic form, addressing specific areas of electronic records management where agencies are having major difficulties. This project will provide guidance on electronic records management applicable government-wide and will enable agencies to transfer electronic records to NARA in a variety of data types and formats so that they may be preserved for future use by the government and citizens. Ultimately, to be successful, the management of electronic records must be an integral part not only of the FEA, but also of the information architecture and infrastructure of each agency.

The FEA includes a Technical Reference Model (TRM) which provides a foundation to support the construction, delivery, and exchange of business and application components, called 'Service Components', that may be used and leveraged in many different applications. Components can be large or small, written in different development environments, and may be platform independent. Components can be executed on stand-alone computers, or a LAN, Intranet, or the Internet. The FEA Service Component Framework provides for the definition, development and deployment of common service components which will be available for use and reuse of applications running in the FEA.

I have the privilege of serving as the co-chair of the Components Subcommittee of the Architecture and Infrastructure Committee, chartered by the CIO Council. The subcommittee's objective is "to foster the identification, maturation, use and reuse of Component-Based Architectures and Architectural Components in the federal government." In this context, NARA will lead the acquisition of records management components which will be included in the FEA's Service Component infrastructure and available to any application. This will be a significant improvement over the current situation where records management is implemented as a stand-alone application which is not integrated with the applications which are used to conduct business.

Electronic records are, in fact, created in such business applications. Agencies will be better equipped to optimize the use of these information assets in e-Government if they can be retained and managed within the systems they actually use to transact business. Records management component services will make this possible, by identifying electronic records wherever they exist, by enabling both government officials and agency

customers to navigate among records, and by ensuring that the integrity of the records is protected and that they are properly retained and destroyed.

Electronic Records Standards

We are also strengthening records management through support for the development of relevant standards and alignment with the emerging national information infrastructure.

First, standards. We have worked with NASA and representatives of other American and international organizations since 1995 in the development of the ISO standard for Open Archival Information Systems. NARA has also contributed to the development of the ISO standard on records management.

We are also working with the Association for Image and Information Management, International, to establish integrated EDMS/ERMS functional requirements for electronic document management system (EDMS) and electronic records management systems (ERMS), and on a proposed standard for the use of the Adobe Portable Document Format (PDF), for the long term storage of multi-page documents that may contain a mixture of text, raster images and vector graphics.

NARA has made substantial contributions to the development and the success of the DoD standard for Records Management Applications (DoD 5015.2-STD), the de facto standard for records management software, adopted by private companies, as well as by other governments, such as the State of Michigan. Currently there are over 40 commercial off-the-shelf software products certified as compliant with the Department of Defense standard. And NARA is supporting DoD in its efforts to update and enhance this standard.

NARA is also one of the original members of the Federal Geographic Data Committee (FGDC). We have contributed to the development of the FGDC's policies on data management and on historical data, as well as its data transfer and content standards. We are proud to be part of the nationwide effort to realize the vision of the National Spatial Data Infrastructure, an effort actively embraced by all levels of government in this country.

National Information Infrastructure

Second, development of the next generation national information infrastructure. The challenges NARA faces in the realm of electronic records are immense and complex, and rendered more so by the prospect of continuing, open-ended change in information technology. But developments in the technology also offer improving prospects for viable solutions. Rather than seeking to surmount these challenges alone, or expecting the market to develop *specific* solutions for all the problems we face, we seek solutions in mainstream developments that promise to become major elements in the emerging national information infrastructure needed to support not only e-Government and e-business, but also education and research in a networked world.

Rather than looking for *specific* archival and records management solutions, we have sought to identify mainstream developments that, with modest redirection, might be adapted to *specific* archival and records management requirements. We have pursued this approach since 1998. In doing so, we have benefited from the prior work of agencies such as NASA, the Defense Advanced Research Projects Agency, the National Science Foundations, the National Institute of Standards and Technology, and the Army Research Laboratory.

One of the major assets we have is the very nature of the challenge we face: its great complexity and the immense size. NARA's electronic records challenge has attracted the attention of world-class researchers in institutions around the country, such as the San Diego Supercomputer Center, the University of California /Berkeley, the Georgia Tech Research Institute, the University of Illinois at Urbana-Champaign, Ohio State University, and the University of Maryland's Institute for Advanced Computer Systems.

The importance of the dynamic challenge NARA faces has been recognized at the highest levels of government. Every year the White House issues what is known as the 'Blue Book.' This is an official supplement to the President's Budget which explains how the Executive Branch coordinates the Government's investment in networking and information technology research and development. In the Blue Books for the last two Fiscal Years, 2002 and 2003, the White House has identified digital preservation in general, and the appraisal and management of electronic records in particular, as significant areas of research. The Administration's recognition of the need for research in this area is critical because, as long as the technology continues to develop, and as agencies find new ways to apply it in conducting the public's business, NARA will have to continue to work with leaders in research and development to find new solutions to new problems.

But we cannot delay action. NARA needs to move ahead aggressively to acquire and implement solutions that will enable us to ensure that important records of our era are not lost. Records such as:

- the State Department's worldwide diplomatic correspondence, which was converted to digital form starting in 1972;
- the digital returns from the 2000 Census of Population;
- the digital map of the United States;
- operational and intelligence records of the war in Iraq; and,
- the automation of Congress's legislative process and the planned coverage of its proceedings in High Definition Television.

The Electronic Records Archives Program

The insights we have gained and the lessons we have learned to date from our research collaborations have emboldened us to launch a frontal assault on the technological challenges entailed in preserving and providing sustained access to valuable electronic records in the National Archives and in the Presidential libraries. This strategic initiative is our Electronic Records Archives (ERA) Program. This program includes three

elements. It is responsible for the research partnerships I have just mentioned. It is managing the effort to acquire a system which will enable us to manage the lifecycle of electronic records and to preserve them for American citizens, and it supports business process redesign and change management activities to ensure that NARA and its staff are able to use the ERA system successfully.

NARA's vision for the ERA system is that it "will authentically preserve and provide access to any kind of electronic record, free from dependency on any specific hardware or software, enabling NARA to carry out its mission into the future." The key features of the Electronic Records Archives will be (1) an architecture which is scalable to expected growth in the volume of electronic records and evolvable both to accommodate new types of electronic records and to take advantage of improvements in technology, and (2) a system with the ability to accept, preserve, and provide sustained access to any kind of electronic record for as long as needed.

The information and expert advice we have received on the state of the art supports the conclusion that it is possible to develop such an architecture now. With respect to preserving and providing sustained access to electronic records, the proven methods that are currently available are limited to relatively few formats. But we are optimistic that we will be able, over the course of time, to expand the varieties of electronic records that we can preserve indefinitely. There are two elements to this optimism. One is that the requirements of e-Government and of e-business are driving the emergence of standards, products, and services that make information assets independent of any specific hardware or software. The other is that these same requirements will lead agencies increasingly to adopt open-standard, infrastructure-independent formats, such as XML.

In the interim, we will implement a flexible approach to preservation and access. We will start with guaranteed physical preservation of all electronic records in their native formats. While it is desirable to convert the records to formats which would be both infrastructure independent and immune to obsolescence over substantial lengths of time, until the market for products supporting such formats matures, we will adopt methods for replicating these records in formats that are either less sensitive to obsolescence or better for meeting short term demands for access. A number of open formats that are available now include the Portable Document Format for textual information, the Tagged Image File Format for scanned paper, and the relational or other standard database formats not only for data files, but also for semi-structured records, such as e-mail.

Our judgement that it is possible to start building the Electronic Records Archives is supported by market research and by our dialogue with the IT industry. Over the past three months, we have had representatives of 72 companies come to NARA, to tell us about their products and services, which they believe could contribute to the solution. We have had extensive discussions with a dozen major systems integrators who are candidates to lead the development of the Electronic Records Archives system.

What we heard from these companies is, in sum, that it is possible to build a digital archives, which not only can be sustained over the long-term, but which can grow and

evolve in response to the dynamic challenge NARA faces. The need for the Electronic Records Archives system, the possibility of developing it, and the feasibility of an evolutionary development solution has also been affirmed recently in a study we have commissioned from the Computer Science and Technology Board of the National Academy of Science.

The Electronic Records Archives Program focuses on the archival end of the records lifecycle, on delivering historically valuable electronic records to Americans in the future. But it is not ignoring the front of the lifecycle in doing so. Just as, in generic terms, we are seeking archival solutions in mainstream developments in the National Information Infrastructure, we are specifically seeking to build an Electronic Records Archives that is optimally integrated with approaches and technologies used to manage records in active support of the Government's current business.

Partnering to Find Solutions

I would like to close with a final point. Throughout my remarks, I have mentioned many of NARA's partners across the Government and the Nation. NARA's Strategic Plan repeatedly recognizes the need for us to partner with others in addressing the challenge of electronic records. It not only recognizes the need, but also commits us to working in collaboration. We have wholeheartedly embraced this collaborative strategy in the Electronic Records Archives Program. The partners I have named do not by any means exhaust the list, and merely naming them does not do justice to the benefits NARA has derived from working with them.

Our research partners have opened our eyes to possibilities we had not conceived; they have expanded our horizons and strengthened our eagerness to move ahead. The managing partners in our research activities – DARPA, the National Science Foundation, and the Army Research Lab – have opened the doors to a wealth of knowledge, expertise and creativity while enabling us to avoid the need to develop within NARA the full administrative overhead needed to manage high technology research.

We have also been enriched by peer collaborations with other agencies and institutions that also need to preserve digital information for long periods of time. Our earliest partners in this area were NASA, the Patent and Trademark Office, and the InterPARES project. The InterPARES project focuses squarely on the preservation of authentic electronic records. Headquartered at the University of British Columbia, the project brings together archivists, records managers, information scientists, computer scientists and engineers, chemists, conservators, museum directors, and even artists from around the world.

We have also developed profitable partnerships with the library community. We are pleased to be collaborating with the Library of Congress in its initiative to preserve America's cultural and intellectual heritage in digital formats. We also profit from interactions with librarians and technologists from leading universities as a member of the Digital Library Federation.

We are also grateful to the many companies that have not only informed us about their products and capabilities, but have also shared with us the benefits of their experience and insights into the risks, the complexities, and the best practices for developing large, complex IT systems. We look forward to working even more closely with industry in the development of the ERA system in the near future.

Finally, I need to mention our relationship with the General Accounting Office. Two year ago, the Congress asked GAO to review the ERA Program, and they have been with us ever since. We are benefiting from GAO's understanding of systems development, gained in decades of experience across the Government. It is an enviable situation to have GAO going along with us in each step of the program, sharing its insight and recommendations so that we can avoid problems before they happen.

The Archives of the Future

One sign of the vitality and success of our commitment to approaching the challenge of electronic records through extensive partnerships is in the fact that, in 2001, for the first time in history, the Archivist of the United States was invited to give the keynote address at a Symposium of the Institute of Electrical and Electronics Engineers. I would like to close my testimony with a quote from Mr. Carlin's speech to the IEEE:

“The ‘Archives of the Future’ will not consist of many buildings scattered across the country. Instead, the ‘Archives of the Future’ will be available on the desktop of any American who chooses to explore the records of his or her country.... Building this new, digital archives is not and will not be easy. But we have no alternative....

“A society whose records are closed cannot be open. A people who cannot document their rights, cannot exercise them. A nation without access to its history cannot analyze itself. And, a government whose records are lost cannot accountably govern.”

Building the Electronic Records Archives is a difficult, serious endeavor, one which NARA could not achieve without the support contributions of our many partners, not the least of which is the Committee on Technology, Information Policy, Intergovernmental Relations, and the Census. I thank you for the opportunity you have given us to explain our strategy for ensuring that electronic records management becomes a critical and successful element in the infrastructure of e-Government, and that Americans will continue to have access to valuable records.

Mr. Chairman, I would be happy to answer any questions that you or your subcommittee might have.