

ABOUT US

- [Visit Us](#)
- [Vision & Mission](#)
- [Organization](#)
- [History](#)

BUDGETS, PLANS, & REPORTS

- [Strategic Plan](#)
- [Performance Plans](#)
- [Performance Budgets](#)
- [Performance & Accountability Reports](#)
- [E-Gov Report](#)
- [All Reports & Plans](#) ➔

RULES & REGULATIONS

- [Laws & Authorities](#)
- [Regulations](#)
- [Significant Guidance](#)

FEEDBACK

- [Contact Us](#)
- [Comment on Draft Policy & Regulations](#)
- [Inspector General Hotline](#)
- [Customer Service Commitment](#)

EMPLOYMENT

- [Jobs, Internships & Volunteering](#)
- [Equal Employment Opportunity](#)

RESOURCES

- [Press Releases](#)
- [Publications](#)
- [Information Quality](#)
- [Our Records Schedule](#)
- [Electronic Records Archives \(ERA\)](#)
- [Services to Members of Congress](#)
- [Support the National](#)

 [Print](#)  [E-mail](#)  [Bookmark/Share](#)

Preserving Our Federal Heritage in the Digital Era: What is NARA's Role in Creating the Government's Digital Archive?

Lewis Bellardo
Deputy Archivist of the United States

Presentation at Federal Library and Information Center Committee Forum on Preserving Electronic Record
March 27, 2001

Introduction

Good morning. I'd like to begin by sharing a [chart](#) with you that shows the exponential growth of electronic records. This growth presents challenges for everyone, not just archives. Interoperability across disparate platforms and different software is needed by government and industry to conduct current business. The challenges to accomplish this are similar to the challenge to preserve authentic and reliable electronic records over time.

Description of NARA's Electronic Records Archives Project

Our Electronic Records Archives project (ERA) is focused on identifying and developing ways to preserve and provide access to electronic records within a comprehensive and stable architecture that will be:

- infrastructure independent,
- modular,
- scalable, and
- extensible.

This architecture will not be dependent on any specific hardware and software. It will also be modular: we intend to leverage obsolescence to work for us - we will use plug-in components that can be replaced as technology changes. By scalable, I mean that the approach can be used to manage the electronic records of a small repository - such as a State archives or historical society - as well as the National Archives of the United States. Extensible means it will be able to handle additional kinds of electronic records over time; it will not be limited to specific types of records that exist today.

There is a [conceptual diagram](#) of the Electronic Records Archives in your handout package. Records will be handled as collection-based persistent objects, using XML DTDs, style sheets, and topic maps. The Electronic Records Archives will have three integrated components:

[Archives](#)[FAQs](#)

Learn why
Democracy
Starts Here

watch the
video



[Go to the
Speeches and
Writings by the
Archivist Main
Page](#)

NARA will gain intellectual control and "wrap and containerize" the records with metadata necessary to preserve their authenticity on the Accessioning Workbench. This control could also be achieved while the records are in agency custody.

The collections will be stored in the Archival Repository and described in the Archival Research Catalog. The Archival Research Catalog will be an integrated online catalog of NARA holdings in all media. We will have the Catalog available for use later this fiscal year.

The records will be accessed by researchers through the Reference Workbench. High-use records will be maintained online or near-line. Other less frequently requested records will be "reassembled" from the Archival Repository for presentation when they are requested.

To develop the Electronic Records Archives, we are working with the National Partnership for Advanced Computational Infrastructure (NPACI) and the San Diego Super Computer Center, and other research partners and collaborations, as I will discuss later in this presentation. We plan to complete the research, development, prototype, and pilot for the Electronic Records Archives by 2004.

How ERA Fits in with NARA's Responsibilities for the Life Cycle

We have been preserving electronic records - primarily database records - since 1968. We have just taken legal custody of the email and other electronic office files of the Clinton Administration, and are about to accession the first segment of the Department of State electronic cable files. And the types of electronic records that we are accessioning is expanding dramatically as evidenced by the Federal agency web site "snapshots" taken at the end of the last Administration.

To NARA, "Government digital archive" has a specific meaning that goes beyond the concept of a collection of government information in electronic form. The National Archives of the United States is *the* Government's official repository for the *records* of the Federal Government that have continuing value. Just as we preserve Federal records in paper, on microfilm, on audio tapes and motion picture film, we preserve Federal records that were created electronically.

NARA has statutory responsibility - 44 U.S.C. chapters 21, 29, and 33 - for overseeing the life cycle of Federal records. Under this authority we assist agencies with the management of their records and provide disposition authority for the records, which means eventual transfer to the National Archives of the United States of permanent records and destruction of temporary records. We work in partnership with the agencies to protect their records. We must protect privacy, trade secrets, national security and other legally privileged information in records.

It is this responsibility for the life cycle of records that sets us apart from the Library of Congress, the Superintendent of Documents, and other collectors of Government information. There are two important points to keep in mind.

The Government information dissemination products that LC and SupDocs collect begin their life as Federal records that provide evidence of the programs and activities of the agency. One of these activities is information dissemination.

NARA does not collect records - we accession them as a legal process that is part of the life cycle. Records are transferred to NARA from the originating agency, not a third party. And agencies execute transfer of legal custody to NARA when we accession the records. This chain of legal custody is critical for maintaining the authenticity and reliability of records as evidence.

Records are authoritative. Their legal status as records is important. Government, citizens, businesses rely on records for documentation of rights and to carry out Government business. It is critical to be able to use the records in the future with that future day's technology to conduct current Government business.

Example: Veterans need assurance that their medical files created during their military career can be used by VA hospitals to which they may be admitted for treatment 30 years later.

Agencies and NARA need to be able to guarantee the integrity of the record over time. Much of what we are doing with our research on the Electronic Records Archives has applicability to our life-cycle responsibility to assist agencies in ensuring that they can use their records over time. Where an agency must be able to use their electronic records 10, 20, or 30 years down the road, they face the same migration and accessibility issues that NARA is dealing with for permanent records.

Examples:

FAA has to be able to access aircraft safety records for as long as that model of aircraft is still being used.

EPA's records on toxic waste clean up have long-term use to both EPA and citizens.

FDA needs the ability to retain reports of adverse drug reactions for the life of the drug.

DOE must retain long-term records on disposal of nuclear waste.

Where Are We Now With Building ERA?

Our research last year with the National Partnership for Advanced Computational Infrastructure was directed to two key areas:

Scale, security and the ability to use multiple computers working together spread across the country and

Effective collection level preservation, management, and long-term access to diverse, heterogeneous collections of records.

The Prototypes have demonstrated key technologies necessary for multiple high performance national computer systems to work together to securely preserve and then to provide access to electronic records. We have explored using "knowledge-based information management" approaches that will enable us to support collection level preservation, management, and long-term access. This is important because we have to find ways of organizing immense volumes of interconnected material in a way that will allow us to find all that we are looking for - but "just" what we are looking for without requiring us to scroll through many thousands of query results. This, of course, is critical for libraries as well. To give an example, a recently conducted narrowly defined search of the Clinton Administration White House email that we just accessioned yielded tens of thousands of "hits."

Our research has also explored the issues around and demonstrated the ability to a manage an increasingly broad variety (that is, technically diverse and from a wide range of sources) of collections (not just unconnected records) including:

F.D.R. Presidential Library Web Site Collection

Office of the Secretary of Defense Gulf War "Gulf-link" Web Site Collection

Geospatial Records Collections...multiple sources/varying formats

Vietnam War Herbicide Records Collection

Vietnam War Psychological Operations Information System (PSYOPSIS) Sortie (Air Psychological Warfare Activities Data) Records Collection

Survey of Household Food Consumption, and

UC Berkeley Digital Library Project Collection, which contains a huge variety of record types and formats.

Our Approach: Partnerships and Collaborations -- ERA and Other Electronic Records Initiatives

NARA Strategic Plan, first issued in 1996, calls for NARA to build on collaborations with external partners. Our work today has built on a number of earlier partnerships and collaborative efforts. Let me list a few of them here. A [complete list](#) is provided in your handout.

NARA has been an active member, U.S. Committee of OAIS since 1995. NARA has hosted 16 of the 19 U.S. committee workshops. *The Reference Model for an Open Archival Information System (OAIS)* is currently a draft ISO standard.

We are members of the Federal Geographic Data Committee, whose Historical Data Working Group is chaired by NARA.

Beginning in 1998, we have worked with the Defense Advanced Research Projects Agency (DARPA) and Patent and Trademark Office (PTO) on their Distributed Object Computation Testbed (DOCT) to handle 2 million electronic patent application case files.

NARA has also been very involved in the InterPARES Project, a major international research initiative begun in 1998. The participants in this project are archival scholars, computer engineering scholars, national archival institutions and private industry representatives from Europe, Asia, Canada, the United States, and Australia. The InterPARES partners are collaborating to develop the theoretical and methodological knowledge required for the permanent preservation of authentic records created in electronic systems. The peer review committee for InterPARES reported recently that "The importance of the InterPARES research for the archival world, academia, industry, government, and society in general, both in Canada and around the globe is considerable."

We have viewed these collaborations as a means not only of furthering NARA's program but also fostering further collaboration among our partners.

For example:

The OAIS model is being used for the InterPARES project

SDSC is now in partnership with InterPARES for research on high-level models for archival preservation

The National Historic Publications and Records Commission (NHPRC), NARA's grant-making arm, has funded a project at the SDSC to scale ERA for smaller archives.

Other NARA Initiatives with Electronic Records

I'd like to take a few minutes to outline some of NARA's other electronic records initiatives. Our responsibilities at the front end of the life cycle of records - and the fact that the permanent records being created today will feed into the Electronic Records Archives - dictate that we work with agencies and their electronic records at the front-end. We need to ensure that electronic records are created and managed in a way that will ensure their availability and usability at the end of their active use in the agency. We have taken a multi-tier approach to the front end of the life cycle:

We endorsed the DOD standard for records management applications - DOD STD 5015.2 - and the certification process that DOD uses to ensure that commercial RMA software is compliant with the standard. We are exploring with DOD further work on extension of the standard.

We are in the middle of our own prototype RMA test within NARA. The lessons we will learn from this test will help us provide better guidance and assistance to agencies.

We are involved with the XML Working Group of the CIO Council's EIEIT Committee

We have been actively involved with development of the International Standard for Records Management. I am pleased to tell you that it was recently balloted overwhelmingly for adoption.

We have issued [records management guidance](#) on implementation of the Government Paperwork Elimination Act - GPEA.

We are working with agencies to develop records management guidance for Federal web sites and records.

These are all little pieces that must be stitched together - with additional pieces - to provide the overall solution to managing electronic records. I would also like to tell you about our Office of the Federal Register initiative, eDOCS. This project, when completed, will allow agencies to create authentic regulations online, pass the completed regulations to the Federal Register staff for processing and submission to GPO for publication in the *Federal Register*. This all-electronic process will make it easier for users to research the *Federal Register* - to find, for example, all of the retention requirements Federal agencies impose on a State or regulated industry.

In closing, I'd like to draw your attention to the NARA web site for additional information about ERA and other NARA programs and initiatives. For more information on ERA, go to http://www.archives.gov/electronic_records_archives/index.html. For other information, go to <http://www.archives.gov>.

NARA Research Partnerships

Open Archival Information System (OAIS) Reference Model

With NASA and Consultative Committee on Space Data Systems since 1995

Distributed Object Computation Testbed (DOCT)

With Defense Advanced Research Projects Agency (DARPA) and Patent and Trademark Office since 1998

National Partnership for Advanced Computational Infrastructure

With National Science Foundation/San Diego Supercomputer Center

Digital Libraries Initiative - NSF grant program

International research on Permanent Authentic Records in Electronic Systems

(InterPARES)

With 7 international, multidisciplinary research teams, 10 national archives since 1998

Presidential Electronic Records Processing Operational System

With Army Research Laboratory and Georgia Tech Research Institute

Membership in organizations developing standards and best practices

World Wide Web Consortium, Digital Library Federation, National Information Standards Organization, Open eBook Form, National Libraries

NHPRC electronic records research grants

See http://www.archives.gov/nhprc_and_other_grants/electronic_records/electronic_records.html

Other NARA initiatives with electronic records

Endorsement of DOD standard for records management applications - DOD STD 5015.2 - and the certification process that DOD uses to ensure that commercial RMA software is compliant with the standard. See http://www.archives.gov/records_management/initiatives/dod_standard_5015_2.html

Membership on the Federal Geographic Data Committee, whose Historical Data Working Group is chaired by NARA

Involved with the XML Working Group of the CIO Council's EIEIT Committee

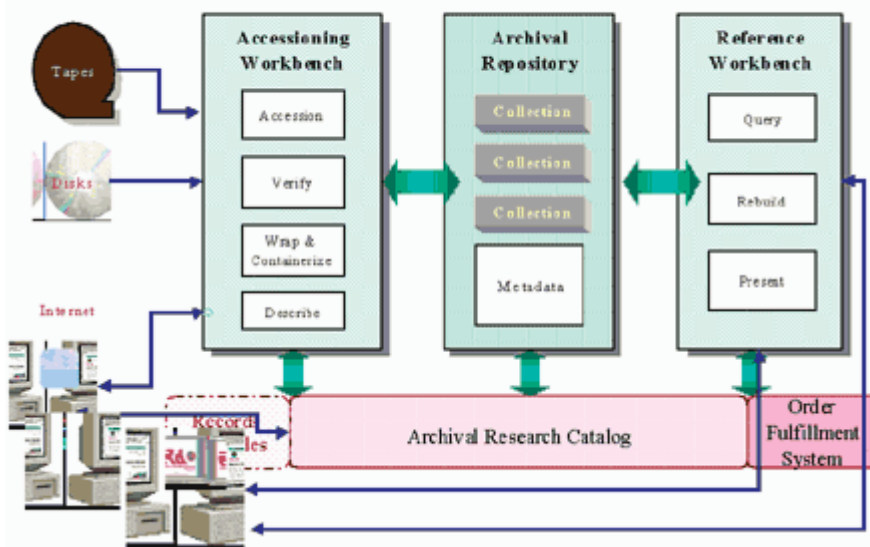
Actively involved with development of the International Standard for Records Management.

Issued records management guidance on implementation of the Government Paperwork Elimination Act - GPEA. See http://www.archives.gov/records_management/policy_and_guidance/electronic_signature_technology.html.

Working with agencies to develop records management guidance for Federal web sites and records.

Digital Strategies Conference, November 16-17, 2000. See http://www.archives.gov/electronic_records_archives/research_room/digital_strategies_conference.html.

Electronic Records Archives: Virtual Workspace Model



NARA Electronic Records Holdings

