
Electronic Records Archives Support Digital Preservation

**NAGARA 2008
Atlanta, Georgia**

John T. Phillips

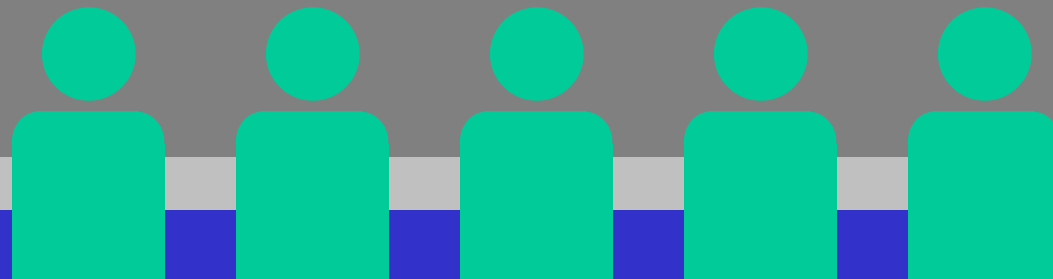
Information Technology Decisions

john@infotechdecisions.com; 865-966-9413

-
- **1:30 p.m.–2:45 p.m. Concurrent Sessions**
 - **C-13: The Electronic Records Archives Program at the National Archives**
 - ERA is the National Archives and Records Administration's strategic initiative to preserve and provide long-term access to uniquely valuable electronic records of the U.S. Government, and to transition government-wide management of the lifecycle of all records into the realm of e-government. Join Rita Cacas, ERA Communications Officer, and John T. Phillips, a member of the Advisory Committee on the Electronic Records Archives, to learn how the project is progressing.
 - ***Presenters: Fynnette Eaton, National Archives and Records Administration; John T. Phillips, Information Technology Decisions***

ERA and e-Records Protection

- ERA Program
- ERA System
- Digital Archives Challenges in the Federal Government
- E-Records Protection



ERA Overview

- **A Program of the US National Archives and Records Administration (NARA)**
- **Electronic Records Archives System**
 - Provides central access to Federal e-records
 - Assures preservation of Federal records
 - Redesigns the life cycle of Federal records
 - Supports easy delivery of all records types
 - Is to be hardware and software independent
 - \$130,000,000 + in program support to date
 - ERA has OMB, GAO, Congressional oversight

Information Resources

The screenshot shows the main National Archives website. At the top, there is a search bar and navigation links. The left sidebar contains a vertical menu with categories like 'Democracy Starts Here', 'WHAT WE DO', 'EXPLORE & INTERACT', 'RESEARCH & ORDER', 'FEDERAL RECORDS', 'PARTICIPATE IN GOVERNMENT', and 'APPLY FOR...'. The main content area features a large banner with the text 'You can apply for: a job, an internship, and volunteer work.' Below this, there are sections for 'News and Events', 'Most Requested' (listing 'Declaration of Independence' and 'Constitution of the United States'), 'FEDERAL HOLIDAYS' (announcing the closure of research rooms on September 3), 'FROM THE ARCHIVIST TO YOU', 'NATIONAL ARCHIVES LOCATIONS', and 'AAD Access to Archival Database'.

www.archives.gov/

The screenshot shows the Electronic Records Archives (ERA) website. The header includes the National Archives logo and the title 'Electronic Records Archives (ERA)'. The main content area is titled 'Electronic Records Archives (ERA)' and features a 'View ERA's Award Winning Video!' section with a video player and a 'Test Transcript' link. Below this, there is a section for 'The National Archives and Records Administration's (NARA) Electronic Records Archives Program (ERA) will sponsor a Global Digital Forum Seminar (GDFR) Governance Workshop to be held in Washington, D.C. on November 14-15 2007'. A list of bullet points describes the program's goals and benefits. The right sidebar contains a search bar, 'Archives by For...' dropdown, and 'ERA Press Information' and 'RMS Record Request Service' links. The bottom of the page has a 'Most Requested' section.

www.archives.gov/era/

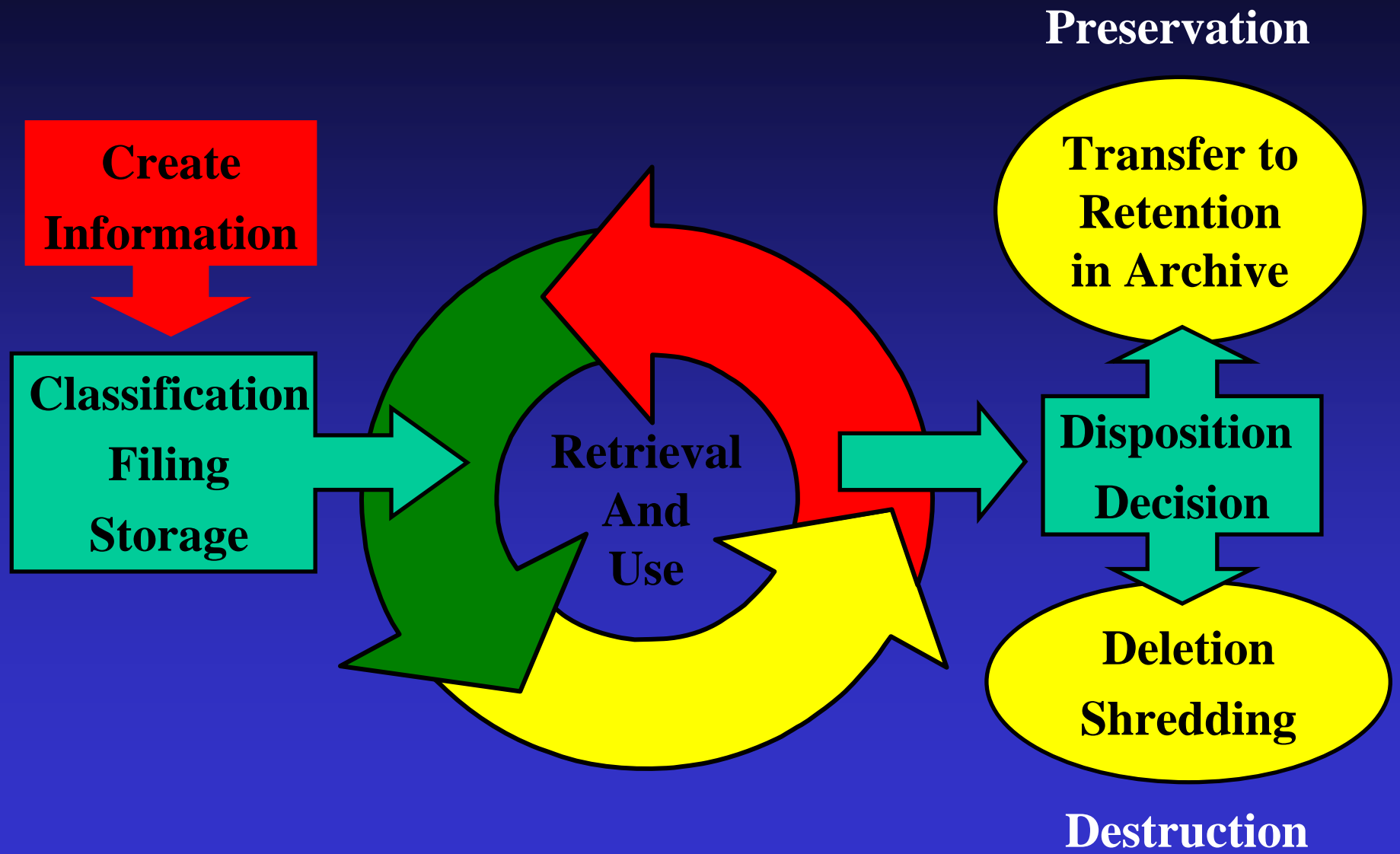
ERA Project Initiation

- September 8, 2005 -- NARA awarded Lockheed Martin a \$308 million contract to build a permanent archives system to preserve and manage electronic records created by the federal government.
- A six-year program was initiated to **incrementally** develop the Electronic Records Archives (ERA) with work performed by Lockheed Martin in coordination with NARA.

ERA Return on Investment (ROI)

- **Makes Federal records more accessible**
- **Enables Information Life Cycle Management of both physical and electronic records**
- **Provides a preservation environment to protect public assets against loss**
- **Allows addressing complex privacy, confidentiality, and security issues for records**
- **Promotes collaboration and communications regarding Federal programs and initiatives**
- **Relieves agencies of some IM responsibilities**

Information Life Cycle

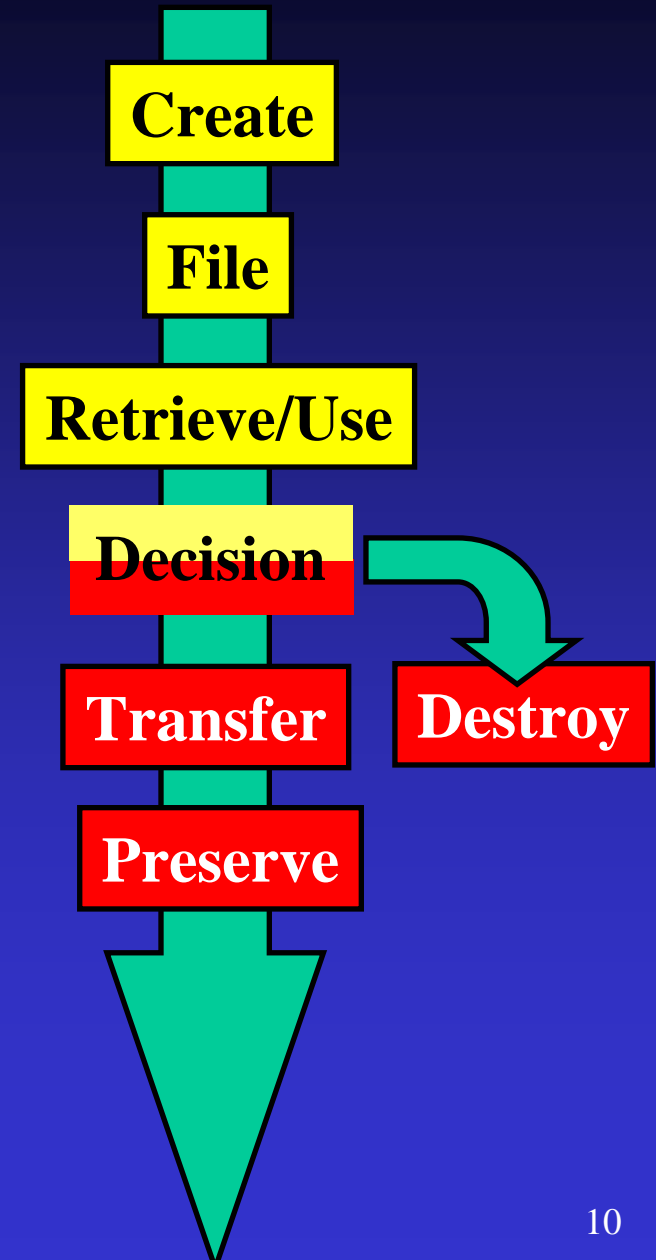


ERA System Challenges

- **Is an infrastructure to automate the Federal Information Life Cycle process**
 - Is to serve as a means of receiving, managing, and preserving electronic records for the future
- **Must work with a framework of agencies exercising their own responsibilities**
 - Many agencies have unique missions and records
- **Must provide public access to records**
 - The public has a very divergent computer base
- **Very unique in system complexity and scale**

Records Responsibilities

- **Records Creator**
 - Identify and organize based on business needs
 - Make initial retention decisions and transfer
- **NARA**
 - Revise retention based on more global requirements
 - Make long-term transfer and retention decisions
 - Enable preservation



ERA Records Activities at NARA

- After “active” organizational use at the “end of the life cycle”
- Take possession of transferred records
- Describe the records
 - Classification, metadata assignment
- Provide access to descriptions
 - Provide access to records content as appropriate
- Provide copies of records
 - When appropriate based on privacy, legal, etc.

ERA Status

- **Design the ERA System and Infrastructure**
 - Includes research in computer science, engineering and archival theory
 - Development of an ERA archival business model
 - Incorporating these components into the ERA system design
 - Build on key research partnerships
- **Build the ERA System**
 - Approve comprehensive requirements and design plans
 - Approve a plan for development, implementation, operation, and maintenance
- **Immediate near-term goal**
 - Have a functional subset of the system operational in 2007

ERA Project Timeline

- Increment 1 will require two years of development to provide system functions for managing, preserving and providing online access to NARA's electronic records
- **Initial Operating Capability (IOC) of the ERA System occurred on June 27, 2008!**
- Four more annual increments will each provide additional functions, technology enhancements, and advanced preservation techniques

ERA Functional Timeline

- **Increment 1**
 - Workflow for records scheduling and transfers
 - Transfers of electronic records, ingest, archival storage
- **Increment 2**
 - Workflow for processing presidential records
 - Content search of electronic records
 - Review of electronic records
 - Classified electronic records
- **Increment 3**
 - Public access to descriptions of all records in ERA
 - Access to electronic records
 - Electronic records preservation

ERA Functional Timeline

- **Increment 4**
 - Expansion of capacity and enhanced functionality
 - Redaction of electronic records
- **Increment 5**
 - Expansion of capacity and enhanced functionality
- **Full use (Full Operating Capability) of the ERA System is expected by 2011**

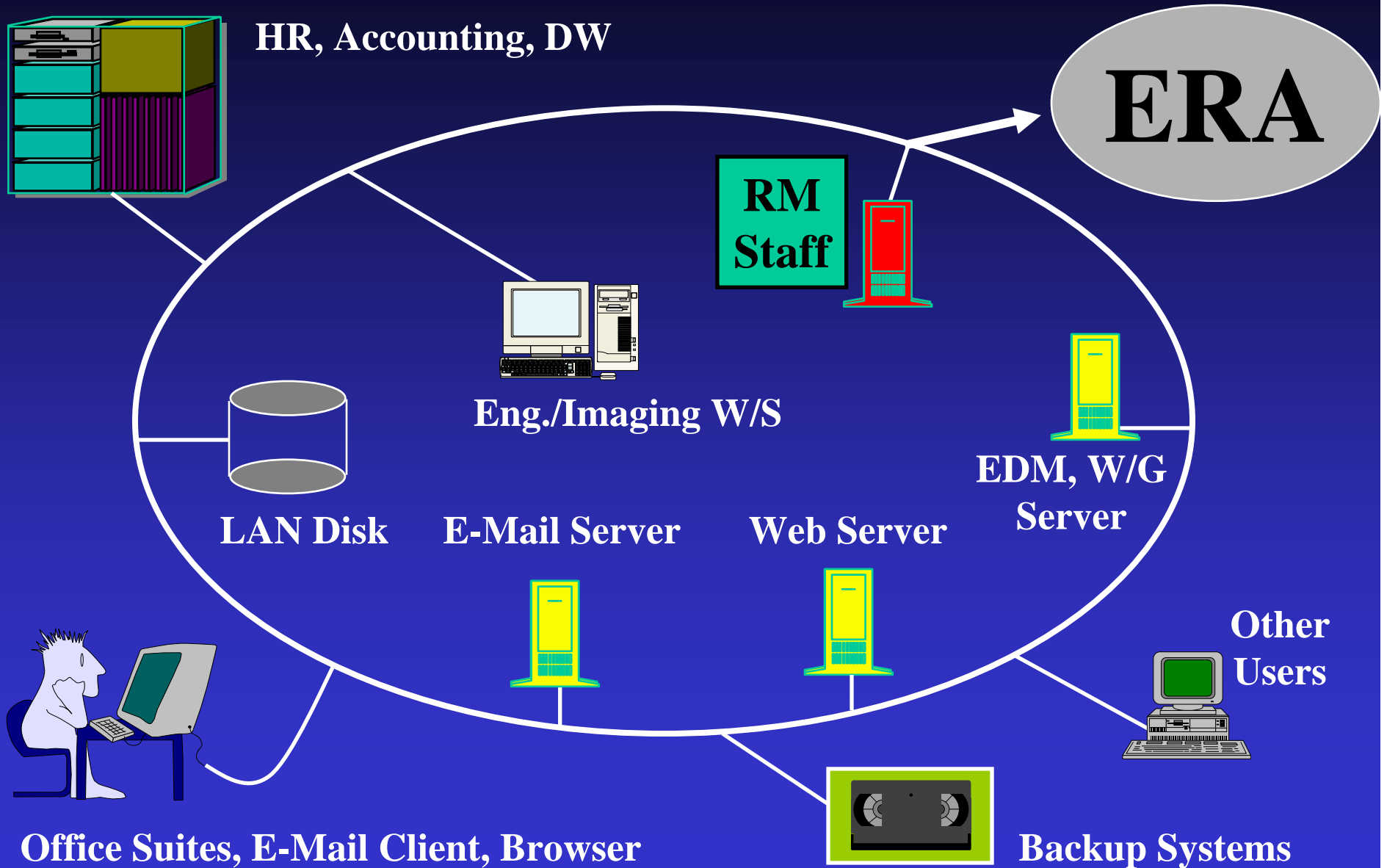
ERA Key Partnerships

- **Open Archival Information System (OAIS) Reference Model**
- **International Research on Permanent Authentic Records in Electronic Systems (InterPARES)**
- **Distributed Object Computation Testbed (DOCT)**
- **National Partnership for Advanced Computational Infrastructure (NPACI)**
- **Presidential Electronic Records Processing Operational System (PERPOS)**
- **NHPRC/San Diego Supercomputer Center**

ERA and E-Records Repositories

- **ERA is a Federal e-records system**
 - Private sector e-records repositories have similar design, implementation, and operational challenges
- **Factors in common with all repositories accepting diverse records from diverse sources:**
 - Are repository infrastructure requirements driven by both internal operational considerations and the need to meet the records transfer and preservation requirements of customers?
 - What are the system interface standards and business process expectations for customers to use to transfer e-records to a repository?
 - How many different electronic records data types can the system accept, retrieve, and produce in a usable format?
 - Does the customer have a Records Management business process to properly interface with the repository business processes?
 - How are technology obsolescence issues addressed with respect to aging hardware, software, and vendor-driven file formats?
 - What is the network infrastructure that repository users need?

Which Records Go To Repository?



E-Repository Technologies

- **Commercial e-Repositories are usually seen as ECM or e-Mail “archives”**
- **Typically built from software such as IBM Content Manager, Documentum, OpenText Livelink, or Oracle/Stellent**
 - **Serves as a repository for many data types**
 - **Holds metadata for information retrieval**
 - **Interfaces with individual contributors through a desktop computing interface – MS Office, Outlook**
 - **Relies on the individuals desktop for viewers**

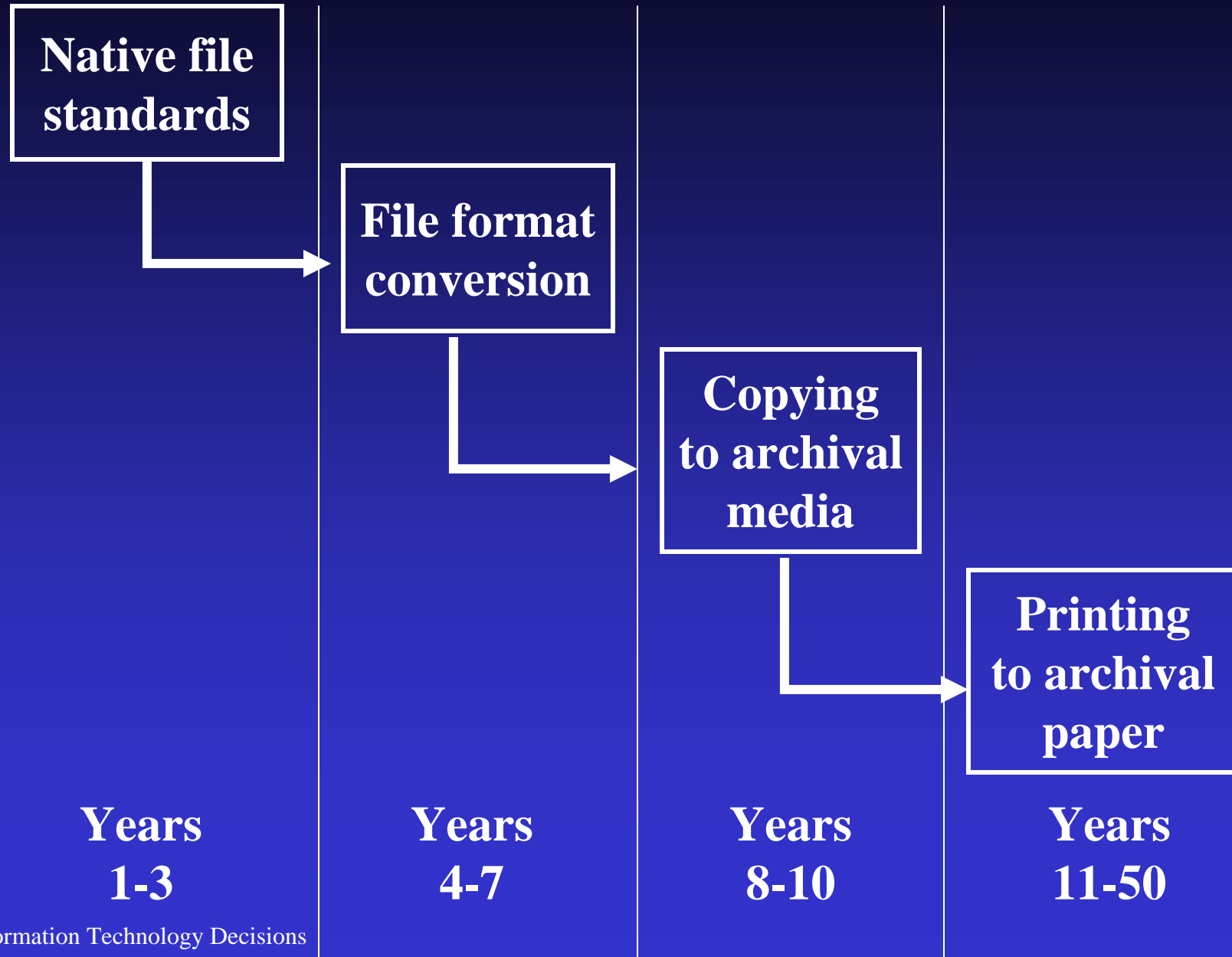
ERM System Challenges

- **Private sector ECM systems are known for becoming only partially implemented**
 - Requirements grow as the system is implemented
 - The number of users or processes needing an interface creates infrastructure scaling problems
 - The network bandwidth and storage capacity needed exceed the original requirements
 - System administrators serve in a limited capacity as records management “intermediaries”
- **ERA also works toward addressing such challenges**

ERA/ERM System Commonalities

- **ERA is very focused on “archival work” within the Federal government arena**
- **However, it also focuses on the independent preservation of “objects”**
 - **Lessons learned here can apply to private sector**
 - **The Persistent Object Preservation architecture allows preservation of any object**
 - **Current private sector preservation typically only extends to PDF/HTML “renditions”**
 - **Object related technologies are of use to all**

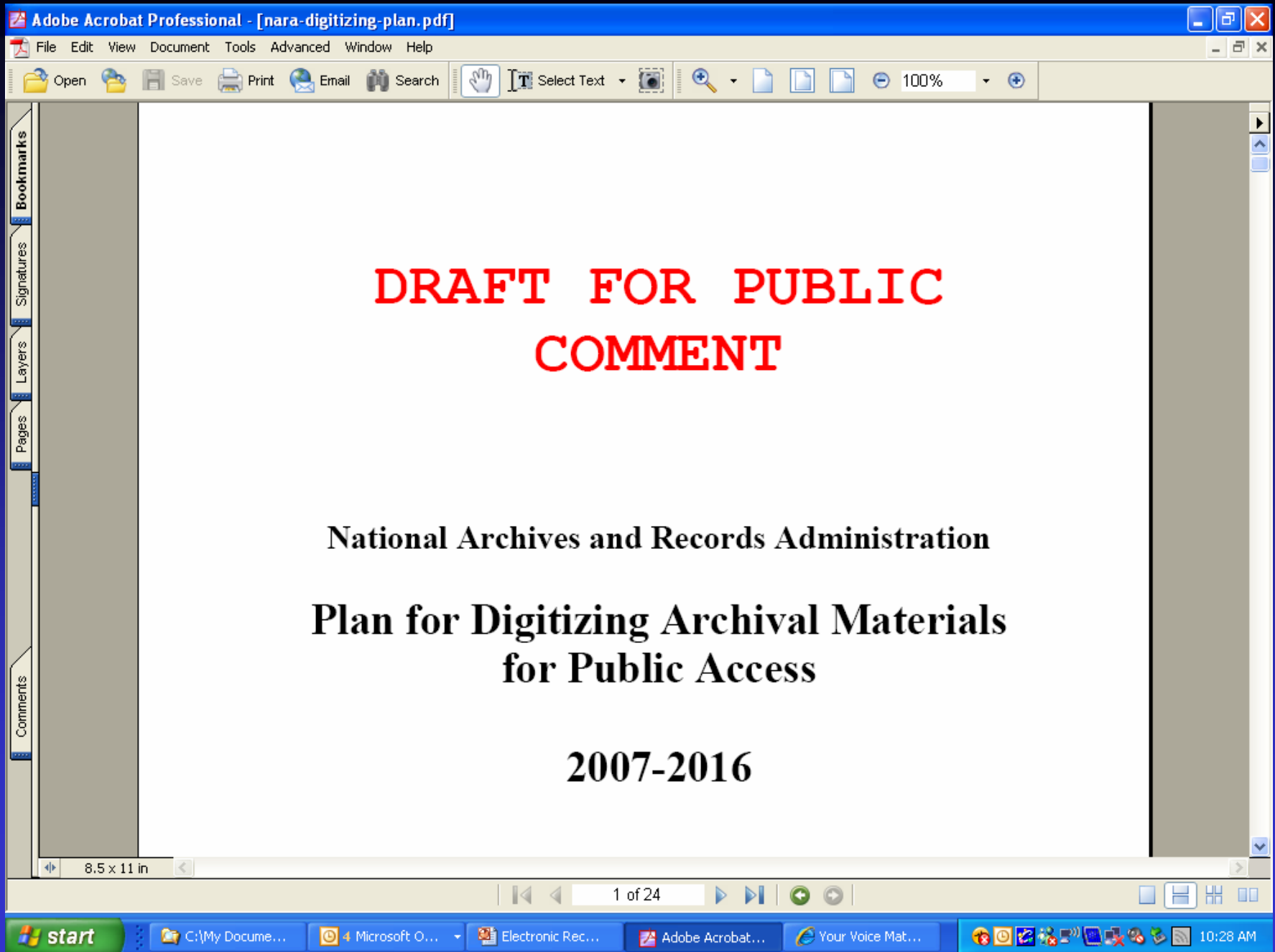
e-Records Preservation Life Cycle?



Draft NARA Digitizing Plan – Site

The screenshot shows a Windows Internet Explorer browser window displaying the NARA website. The address bar shows the URL: <http://www.archives.gov/comment/digitizing-plan.html>. The page title is "Your Voice Matters - Draft NARA Digitizing Plan Available for Public Comment". The website header includes the NARA logo and navigation links: Archives.gov Home, Subject Index, Contact Us, FAQs, and Español. A search bar is also present. The main content area features a large banner with the text "Your Voice Matters" and a photograph of a group of people. Below the banner, the page is titled "DRAFT NARA DIGITIZING PLAN AVAILABLE FOR PUBLIC COMMENT". It includes the following information: "Posted for comment: September 10, 2007", "Comments due: November 9, 2007", and "Send comments to: Vision@nara.gov or by fax to 301-837-0319". There is a link to "Draft Plan" with a PDF icon. The main text of the page states: "The National Archives and Records Administration (NARA) is seeking public comment on its draft **Plan for Digitizing Archival Materials for Public Access, 2007-2016**. This draft plan outlines our planned strategies to digitize and make more accessible the historic holdings from the National Archives of the United States." It further explains that the document is divided into several sections, including an introduction and background, a plan overview, and sections detailing current digitization activities and operating principles. A footer note indicates that PDF files require the free Adobe Reader. The Windows taskbar at the bottom shows the Start button, several open applications (C:\My Docume..., Microsoft O..., Electronic Rec..., Adobe Acrobat..., Your Voice Mat...), and the system tray with the time 10:24 AM.

www.archives.gov/comment/digitizing-plan.html



Adobe Acrobat Professional - [nara-digitizing-plan.pdf]

File Edit View Document Tools Advanced Window Help

Open Save Print Email Search Select Text 100%

TABLE OF CONTENTS

FOREWORD	1
SECTION I: INTRODUCTION AND BACKGROUND	3
SECTION II: OBJECTIVES AND STRATEGIES	6
Scope	6
Objectives of Digitizing	7
Definition of Digitizing for Public Access	7
Approach to Digitizing	8
Strategy One	8
Strategy Two	9
Strategy Three	9
Strategy Four	9
Prioritizing	10
Measurement	11
Process for Feedback	11
SECTION III: NARA MATERIALS TO BE “HARVESTED” AND MADE AVAILABLE ONLINE	13
SECTION V: NARA-LED DIGITIZING PROJECTS	19
APPENDIX A: NARA PRINCIPLES FOR PARTNERSHIPS TO DIGITIZE ARCHIVAL MATERIALS	20
APPENDIX B: RELEVANT NARA GUIDANCE	23

8.5 x 11 in 3 of 24

ERA Future

- **ERA will continue to be a model system and program with benefits for the gov't**
- **ERA architectures and processes will contain lessons learned for archivists, records managers, librarians, and any user or maintainer of an ECM repository**
- **Coordinating programmatic and technical plans with the insights gained from ERA is wise for both public & private sectors**

Information Resources

- A special thanks to Kenneth Thibodeau, Director, Electronic Records Archives Program for some of the ERA Program information in this presentation
- See www.archives.org/era for ongoing ERA related information

Contact Information



John T. Phillips

Information Technology Decisions

1803 Nantasket Road, Knoxville, TN 37922

Phone: 865-966-9413, Fax: 865-671-6667

Email: john@infotechdecisions.com

Web Site: www.infotechdecisions.com