

Building a Future for Our Times: The Electronic Records Archives of the National Archives and Records Administration, U.S.



Public Interest Declassification Board
26 January 2008

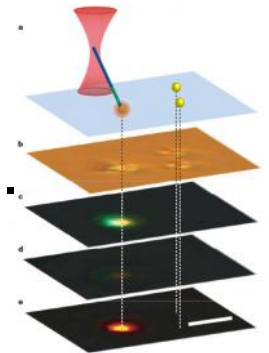
Dr. Kenneth Thibodeau
Director, Electronic Records Archives Program
National Archives & Records Administration

Challenges in Preserving the Records of Our Times

1. The volume of information in digital form is increasing exponentially.
2. The variety and complexity of digital information are increasing.
3. No one knows how to preserve and provide sustained access to authentic electronic records for most types of electronic records.
 - There are no proven methods for preserving most types of electronic records.
 - Archival science has not articulated adequate criteria for determining whether digital preservation or access methods preserve or present authentic records.
4. No one knows what information technology will be in the future .
 - People will want to use the best available technologies to discover and access records from the past
 - The best technologies will not be those that were used originally to create, process, and store the records.

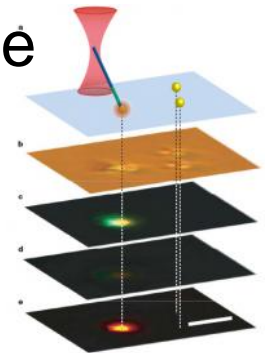
Building for the Future

1. Anticipate change in
 - The characteristics of electronic records,
 - The technologies used to preserve and provide access, and
 - The expectations and behaviors of researchers
2. Recognize those things that will not or should not change
 - Archival science provides stable principles, concepts, requirements and understanding.
 - NARA's mission and the functions that it must carry out in order to accomplish that mission.



Building for the Future - 2

3. Make reasonable assumptions about the future
 - Computers will continue to become more common in the activities of institutions, the lives of individuals, and the interactions of groups of people.
 - The power, speed, capacity and usefulness of information technology will continue to grow
 - Prices for Information Technology will continue to decline.
 - The Internet will continue to grow



NARA's Strategic Response: the Electronic Records Archives Program

- Construction of the Electronic Records Archives system for acquiring, preserving, managing and providing access to electronic records in the National Archives, Presidential Libraries and Federal Records Centers
- Support for the organizational and cultural changes that will enable NARA to maximize the potential of the ERA System
- Research on technologies that offer promise for addressing electronic records challenges.

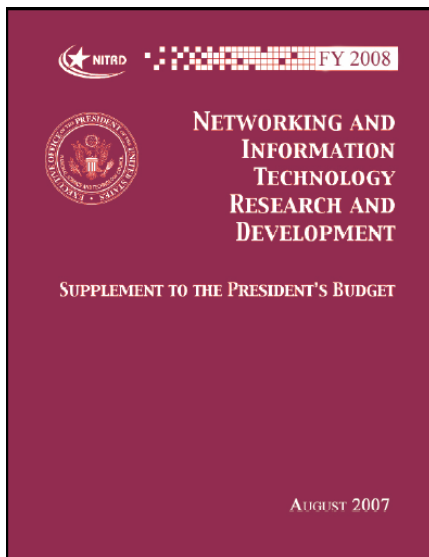


ERA Research Program



“Strategies to assure long-term preservation of digital records constitute another particularly pressing issue for research....”

FEBRUARY 2006



Human Computer Interaction and Information Management (HCI&IM)

NITRD Agencies: NSF, OSD and DoD Service research organizations, NIH, DARPA, NASA, NIST, AHRQ, NOAA, EPA, **NARA**

President's 2008 Request

Strategic Priorities Underlying This Request

Today's increasingly data-centric world requires the effective and strategic use of information assets. To advance the role of HCI&IM in providing strategic support for national priorities, R&D in this area focuses on:

Information integration: To support complex human thought, analysis, and timely decision-making, disparate forms of raw information must be managed, fused, and made accessible to the user in understandable formats.

Next-generation methods, technologies, and tools are needed to fully integrate and efficiently manage massive stores of distributed, heterogeneous information (e.g., science and engineering research data, [Federal records](#)).

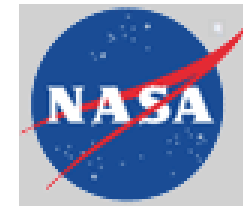
Key research issues include:

- **Information standards:** Data interoperability and integration of distributed data; usability; [provenance and integrity](#) (metadata); generalizable ontologies; accessibility
- **Decision support:** Timeliness of and access to data; user-oriented techniques and tools for summarization, synthesis, analysis, and visualization of information for decision-making; measurement and management of human responses to data
- **Information management (IM):** Efficient integration, maintenance, and access to complex, large-scale collections of heterogeneous data; scalable technologies; integration of policies ([differential sensitivity](#), security, user authentication) with data; integrated distributed data repositories; [long-term curation, data preservation](#); testbeds for evaluating approaches; sustainability and validation of complex models





Research *Partnerships*



National
Science
Foundation



San Diego
Supercomputer
Center



National Computational
Science Alliance



Jared#
Julg#
Iruxp



The Library of Congress



*Army Research
Laboratory*



DIGITAL LIBRARY
FEDERATION

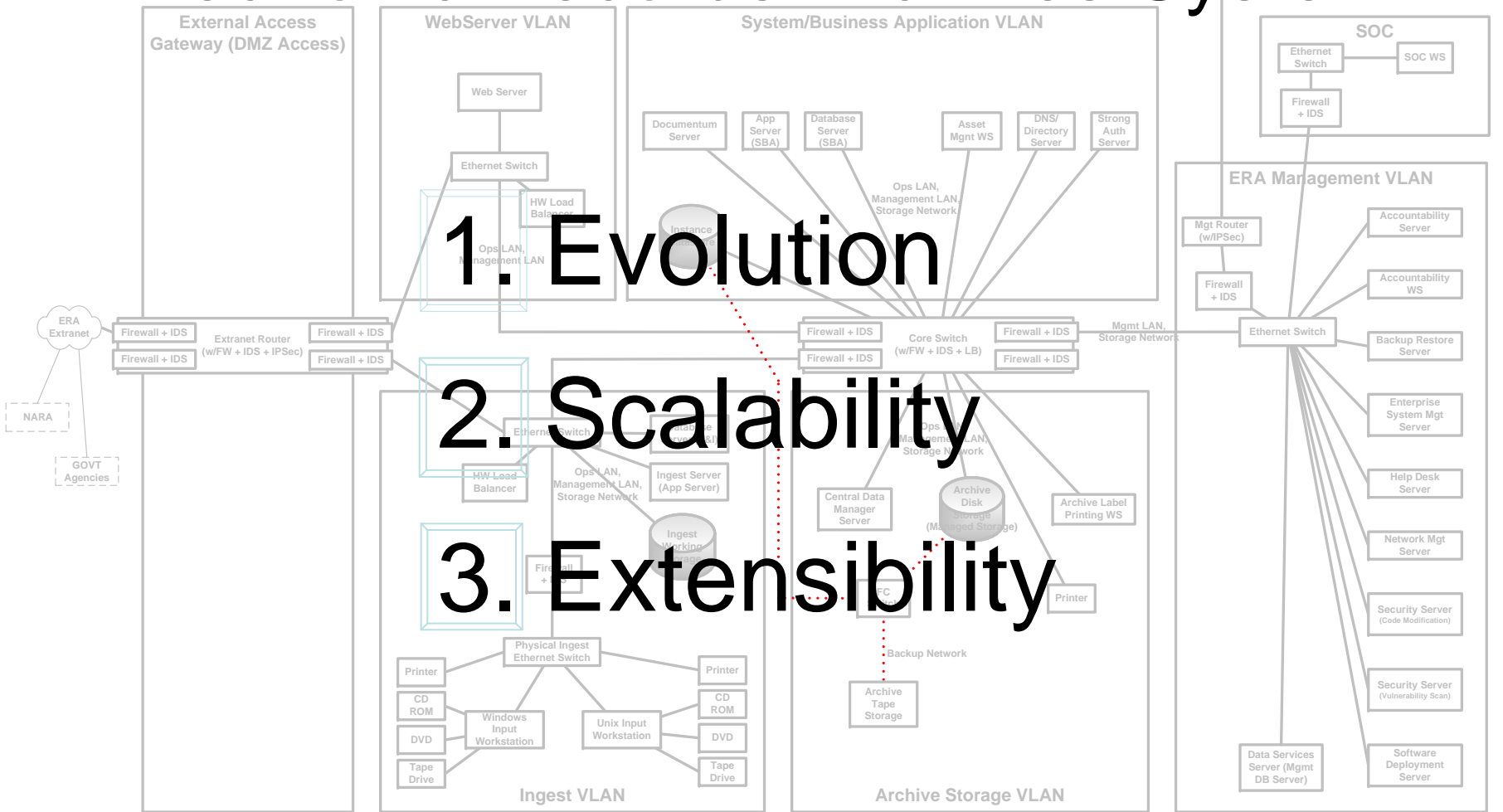


Fundamental Requirements for the Electronic Records Archives System

1. Evolution

2. Scalability

3. Extensibility

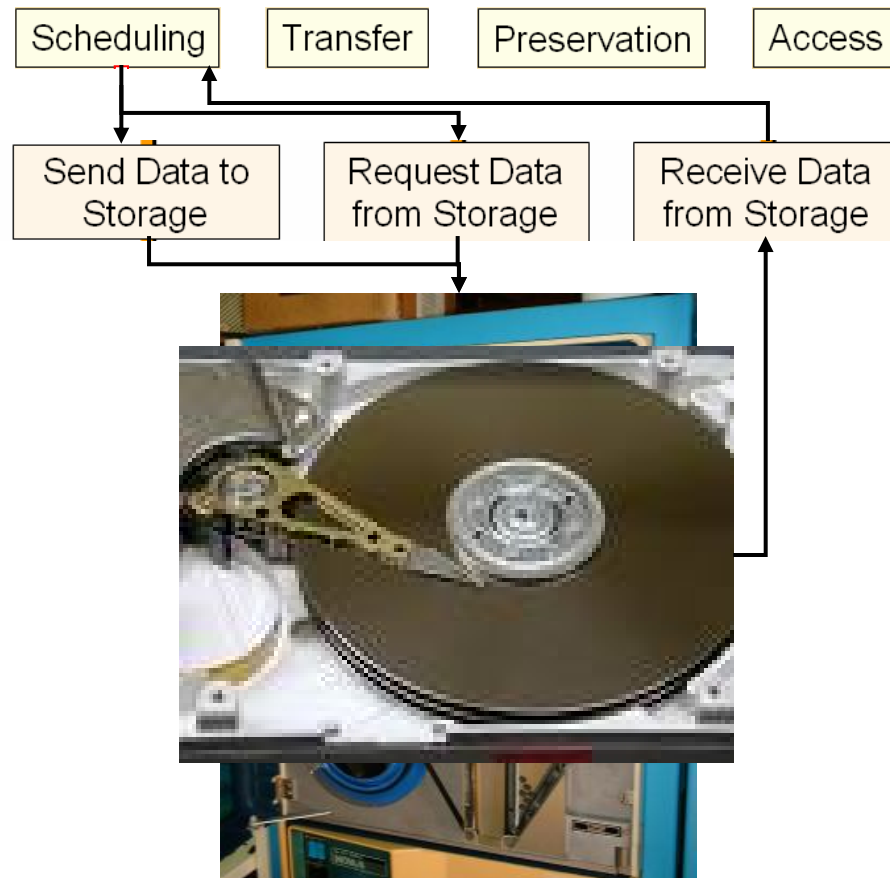


Derived from:
ERA Hardware Block Diagram -2007 0823
(Tab: 11R2 U/USBU Detailed Block)
Updated 24 Aug 2007

1Gb Ethernet ———
2/4Gb Fibre Channel ·····

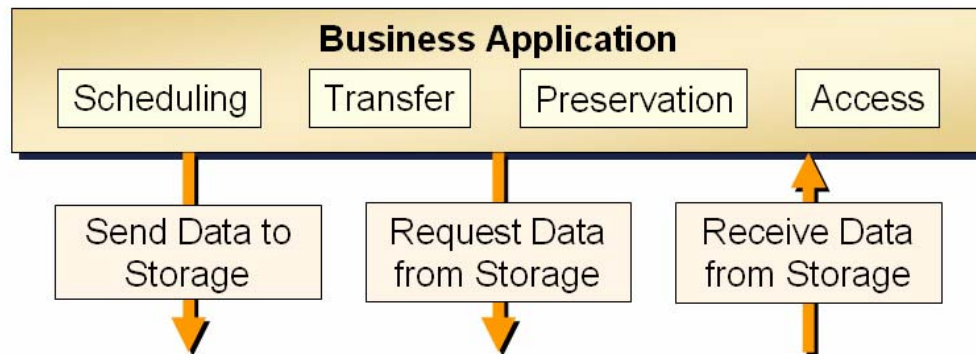


Evolution: Service Oriented Architecture



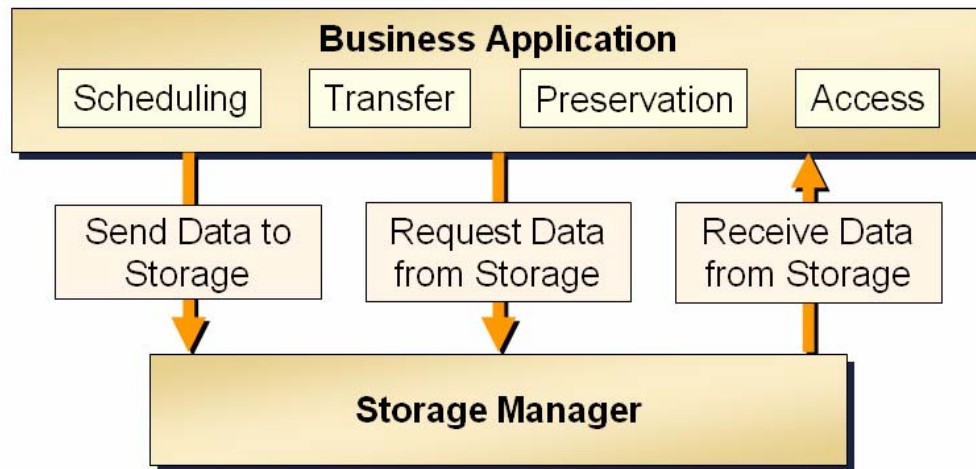
Service Example: Data Storage

Evolution: Service Oriented Architecture



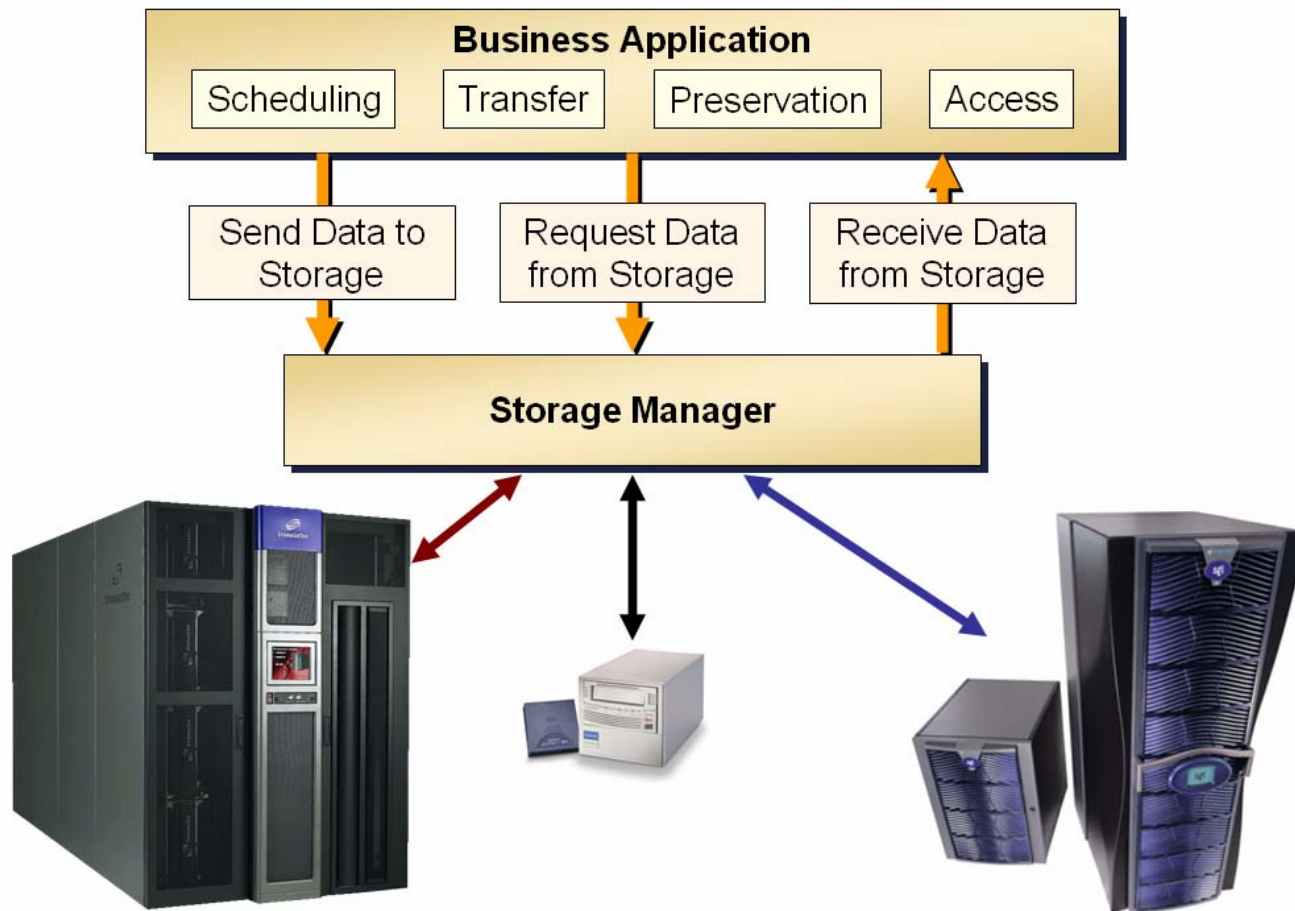
Service Example: Data Storage

Evolution: Service Oriented Architecture



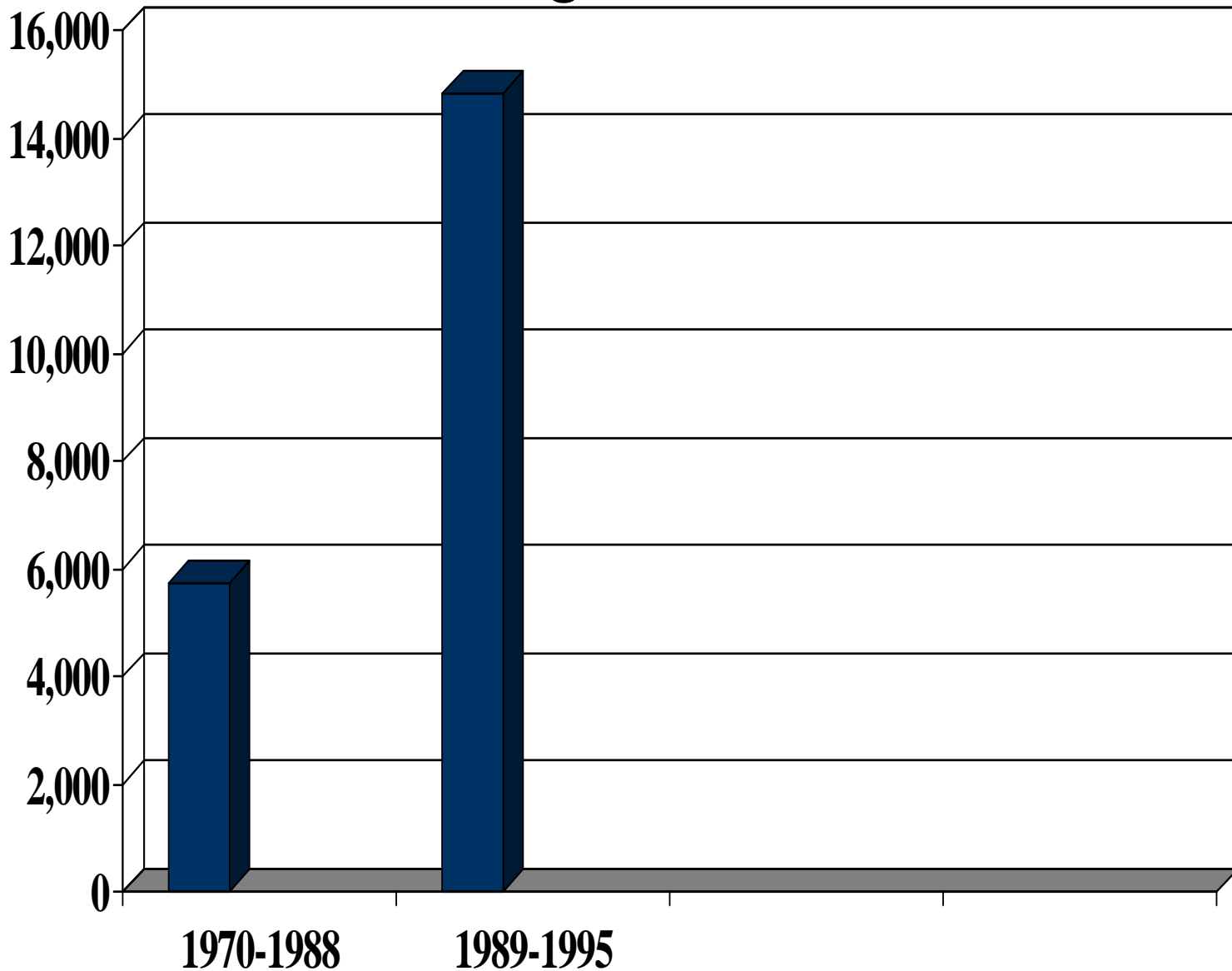
Service Example: Data Storage

Evolution: Service Oriented Architecture



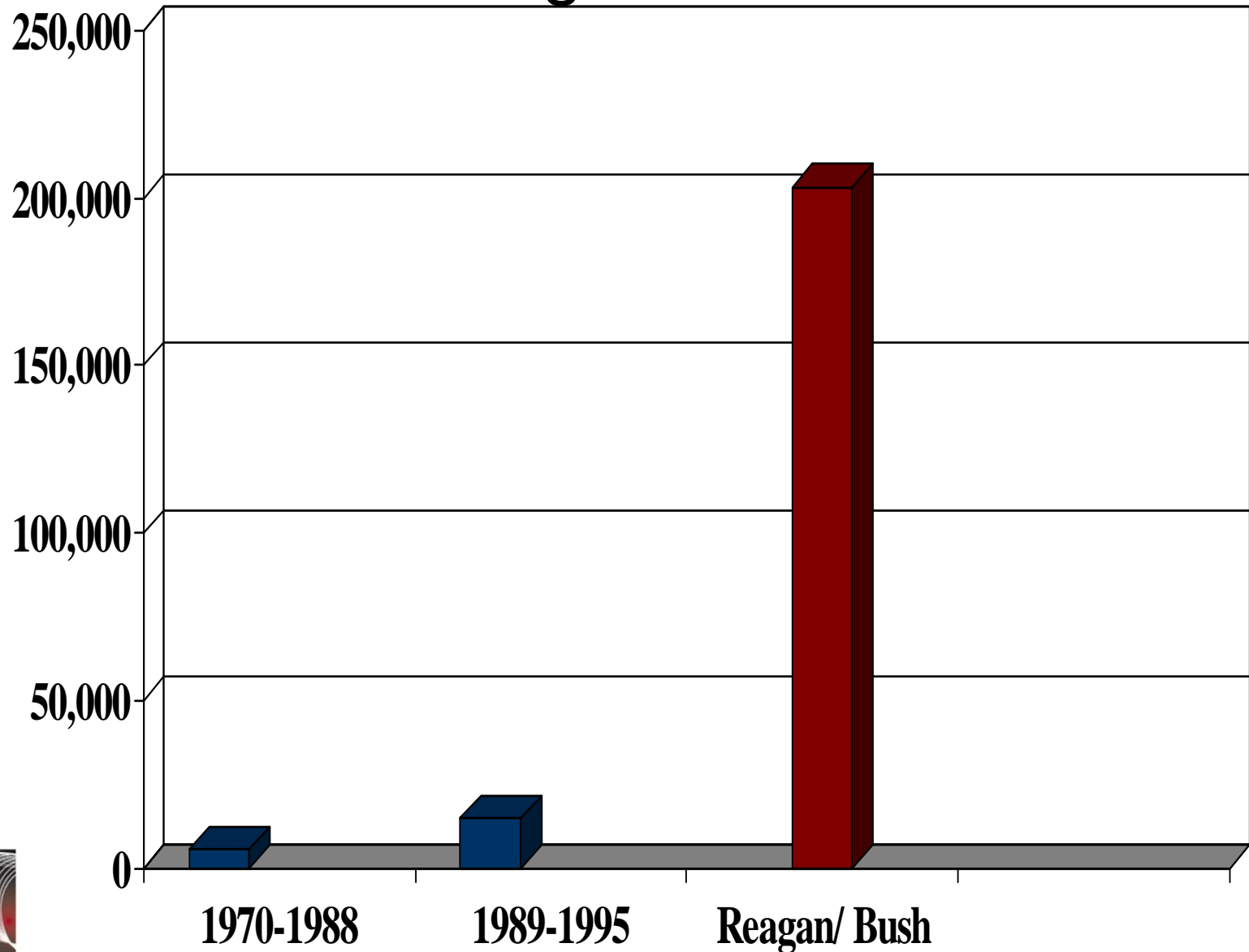
Service Example: Data Storage

Scalability: Transfers of Digital Files to NARA



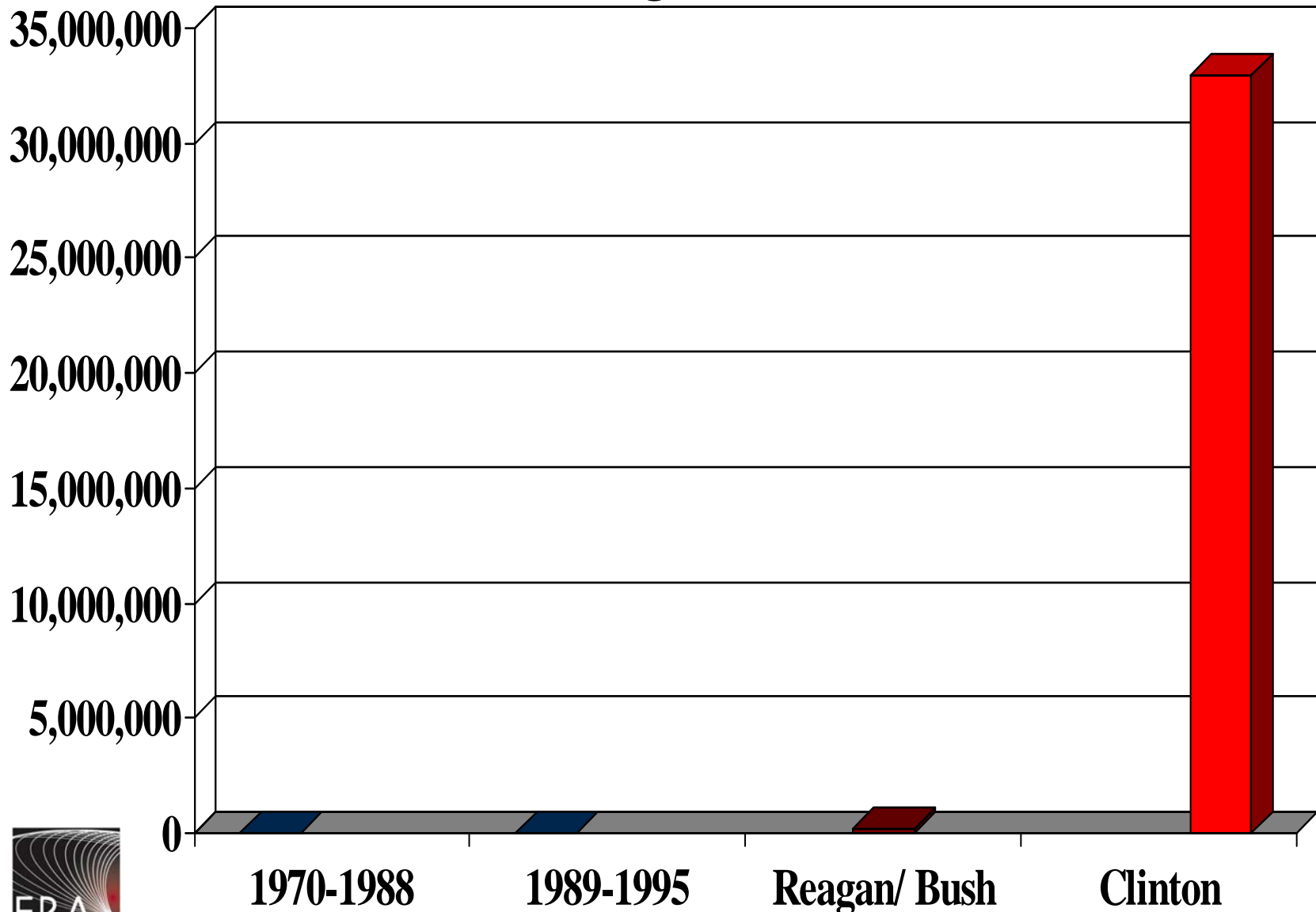
Scalability:

Transfers of Digital Files to NARA



Scalability:

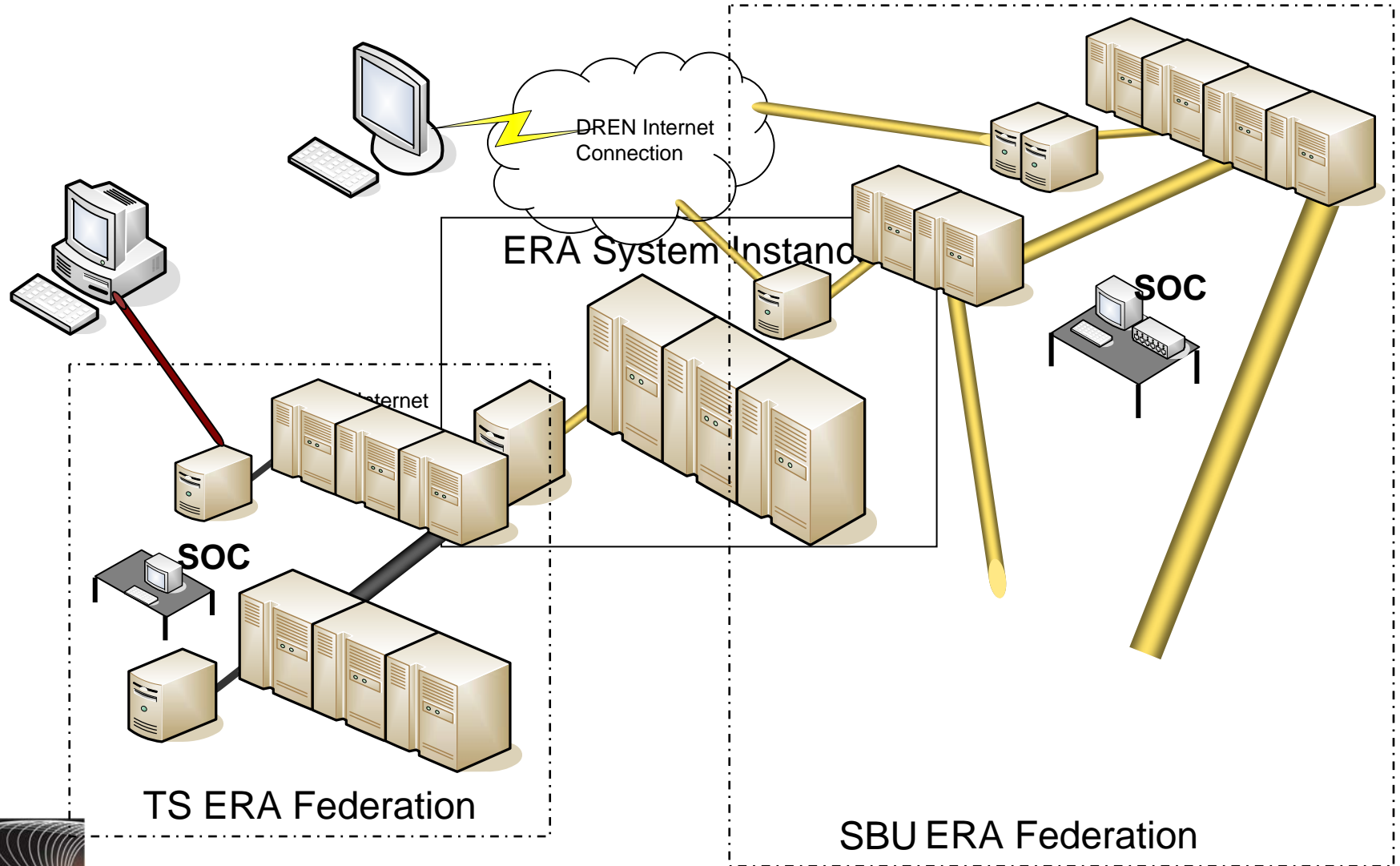
Transfers of Digital Files to NARA



Addressing Scalability

- Product evaluation
- ERA Architecture
 - Service Oriented Architecture
 - Instances and Federations

System Instances and Federations



Extensibility:

ERA: A Set of Nested Systems

Outer system:

- lifecycle management of records of all types

Inner Electronic Records System:

- Ingest, preservation, disposition, and access to electronic records

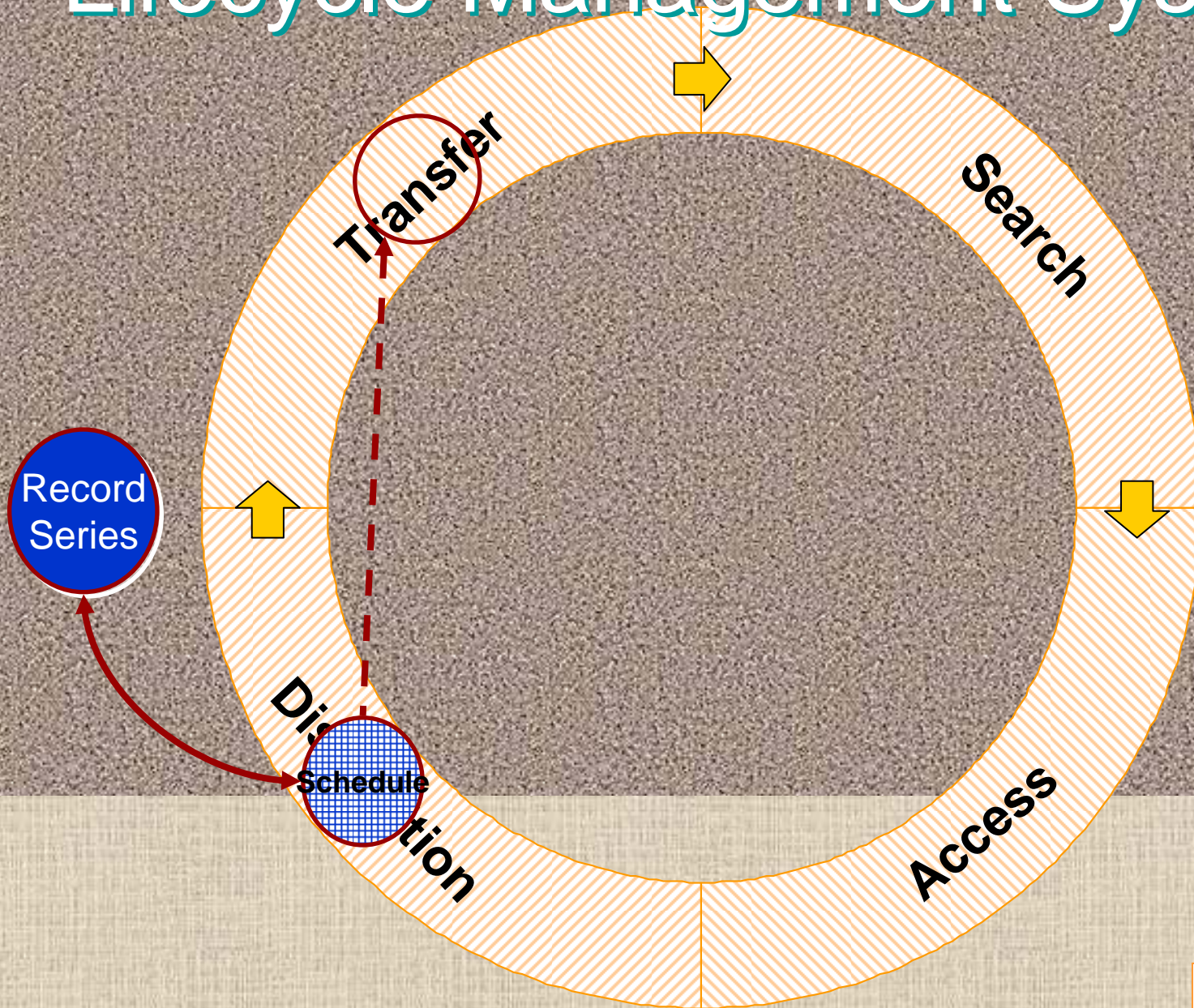
Search & Preservation Frameworks

- Support a variety of different approaches to different needs.

Archival “mini-systems”

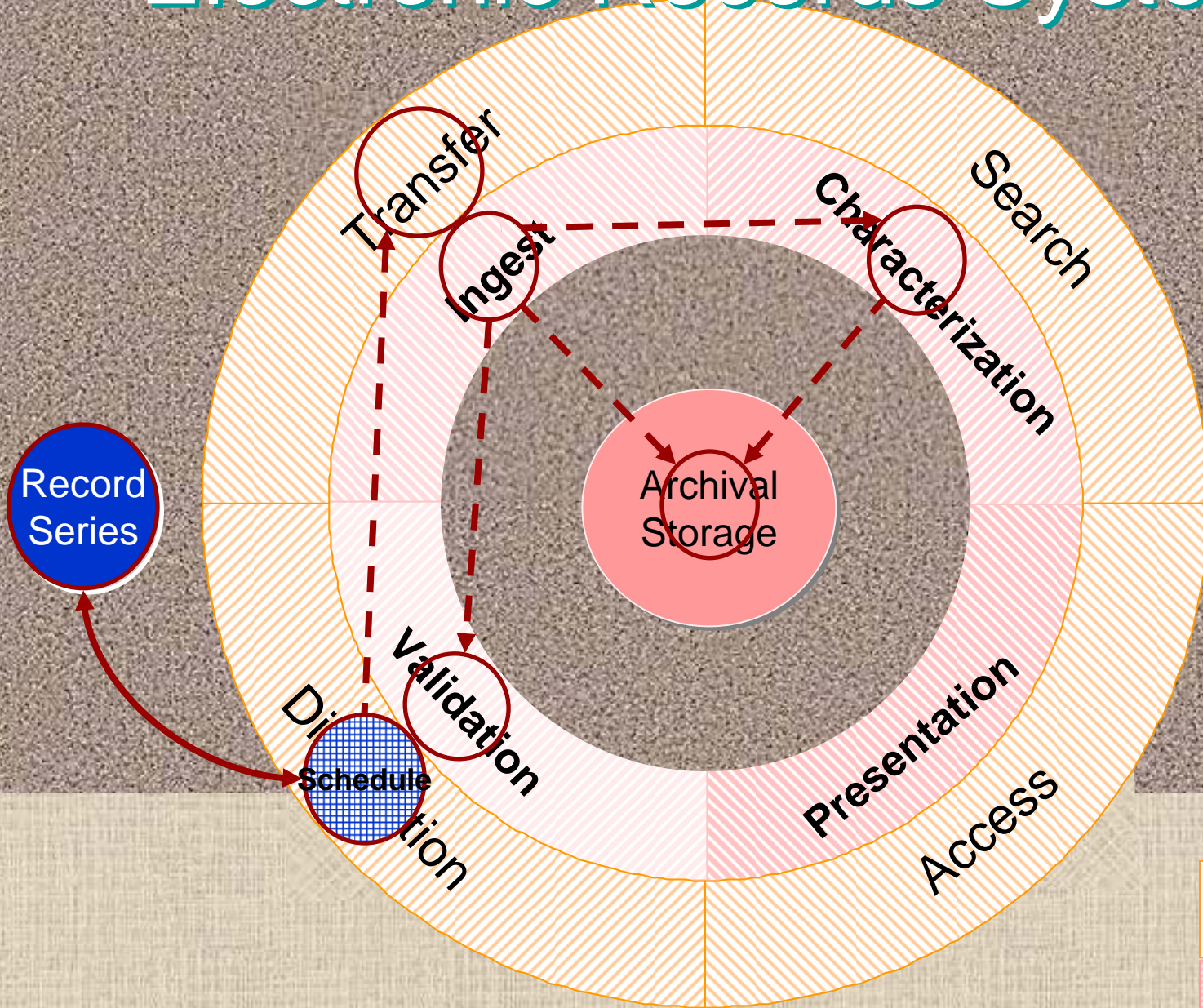
- Specific, systematic management for each series or aggregate of electronic records.

Lifecycle Management System



Lifecycle
Management

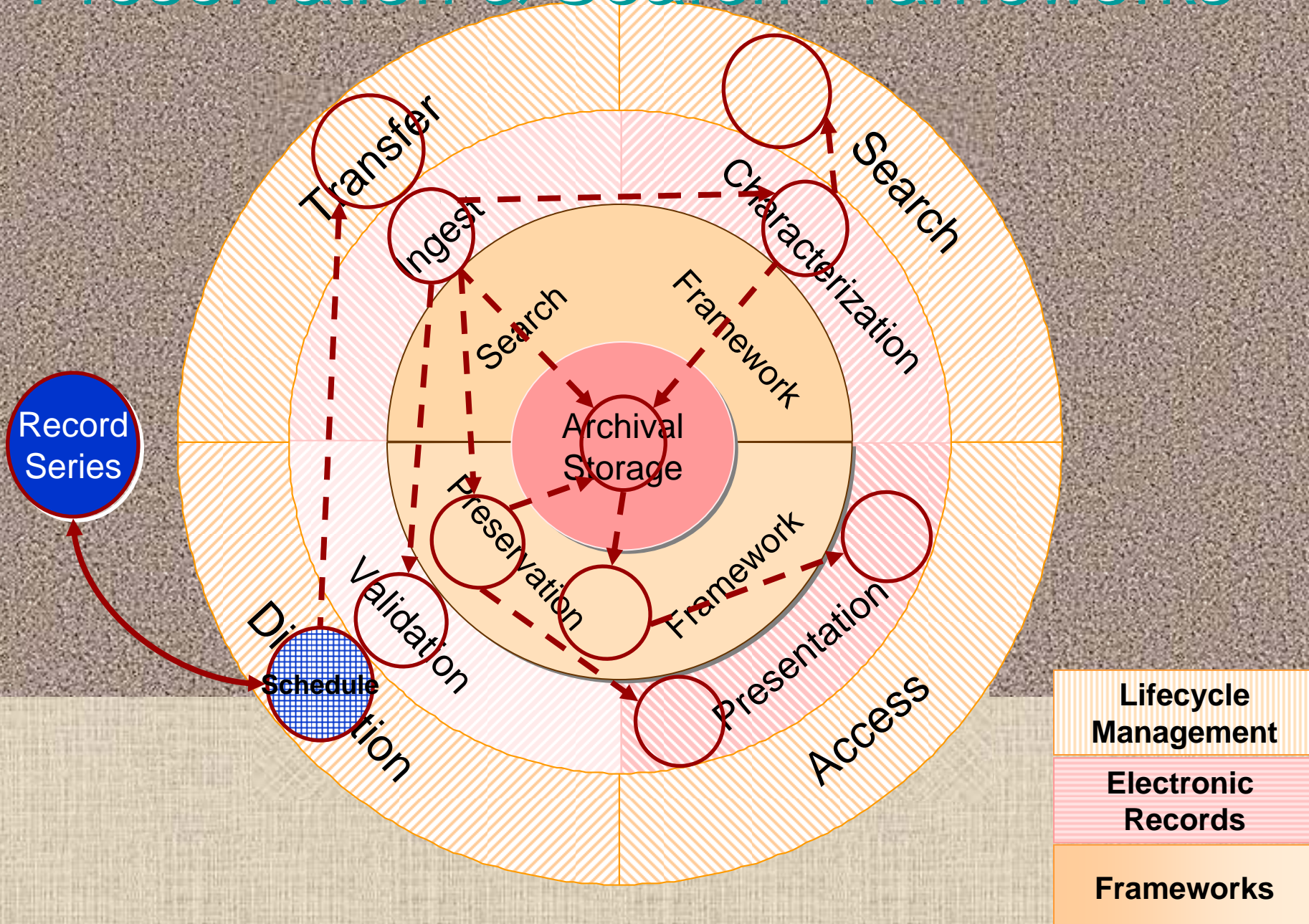
Electronic Records System



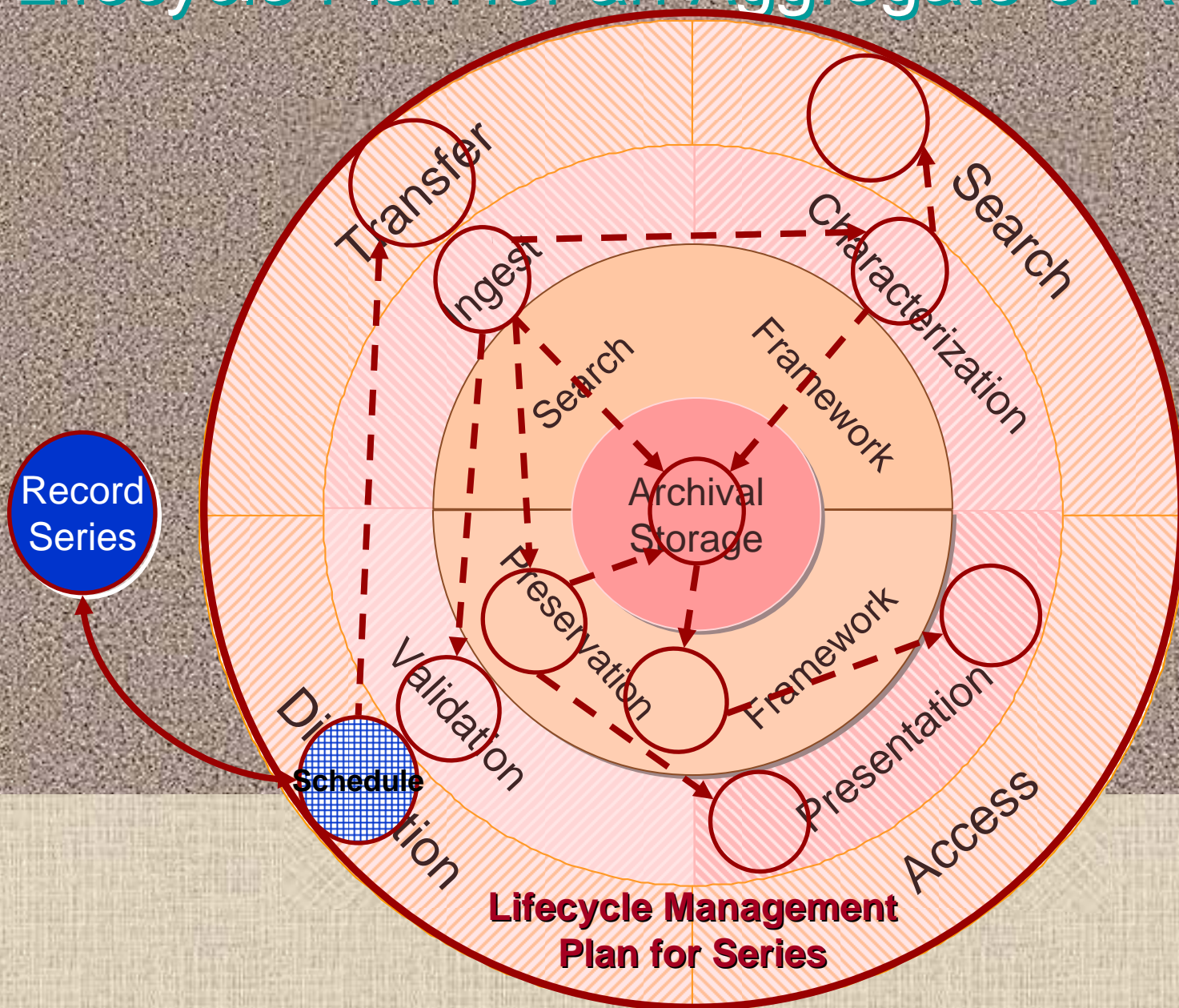
Lifecycle Management

Electronic Records

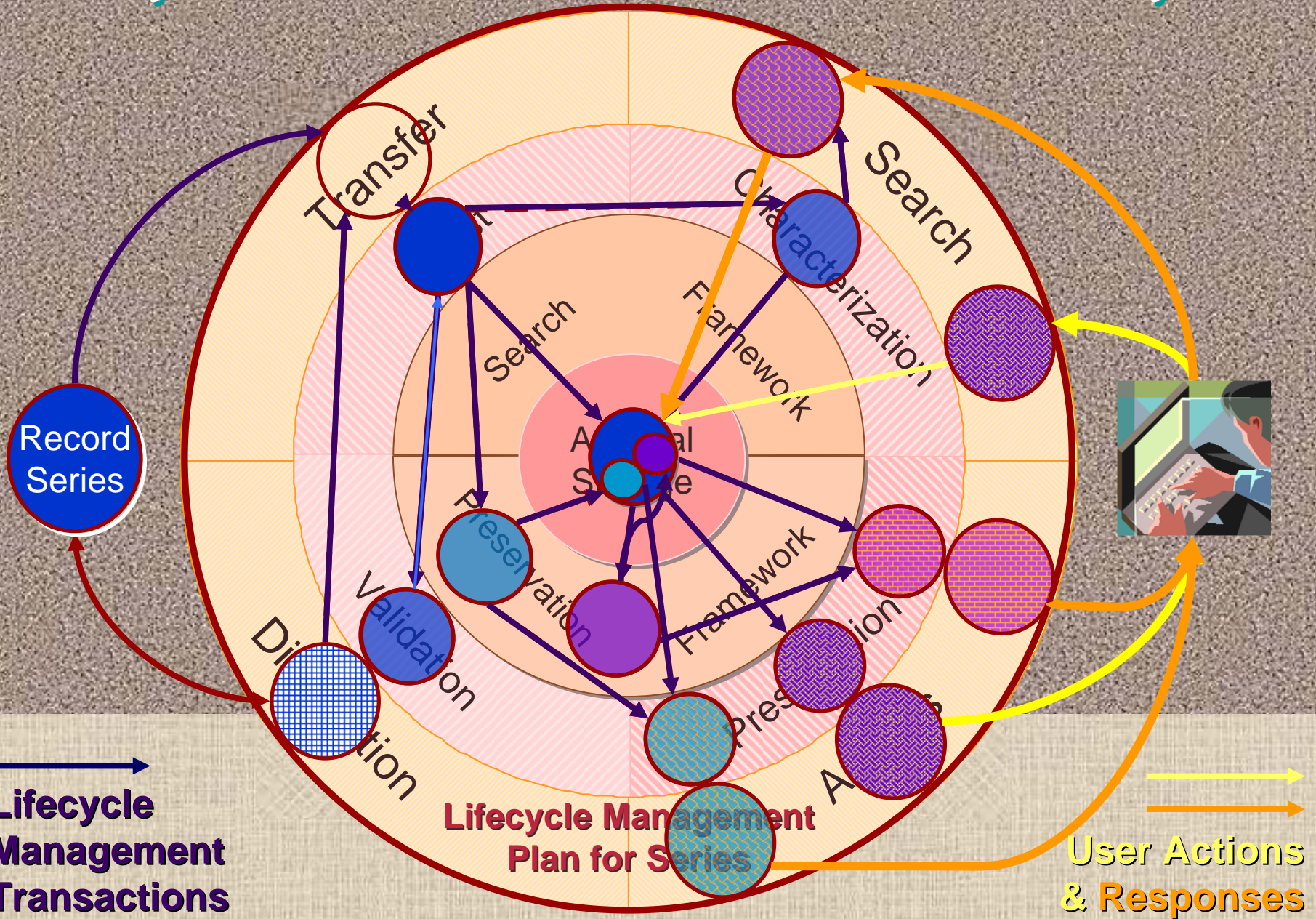
Preservation & Search Frameworks



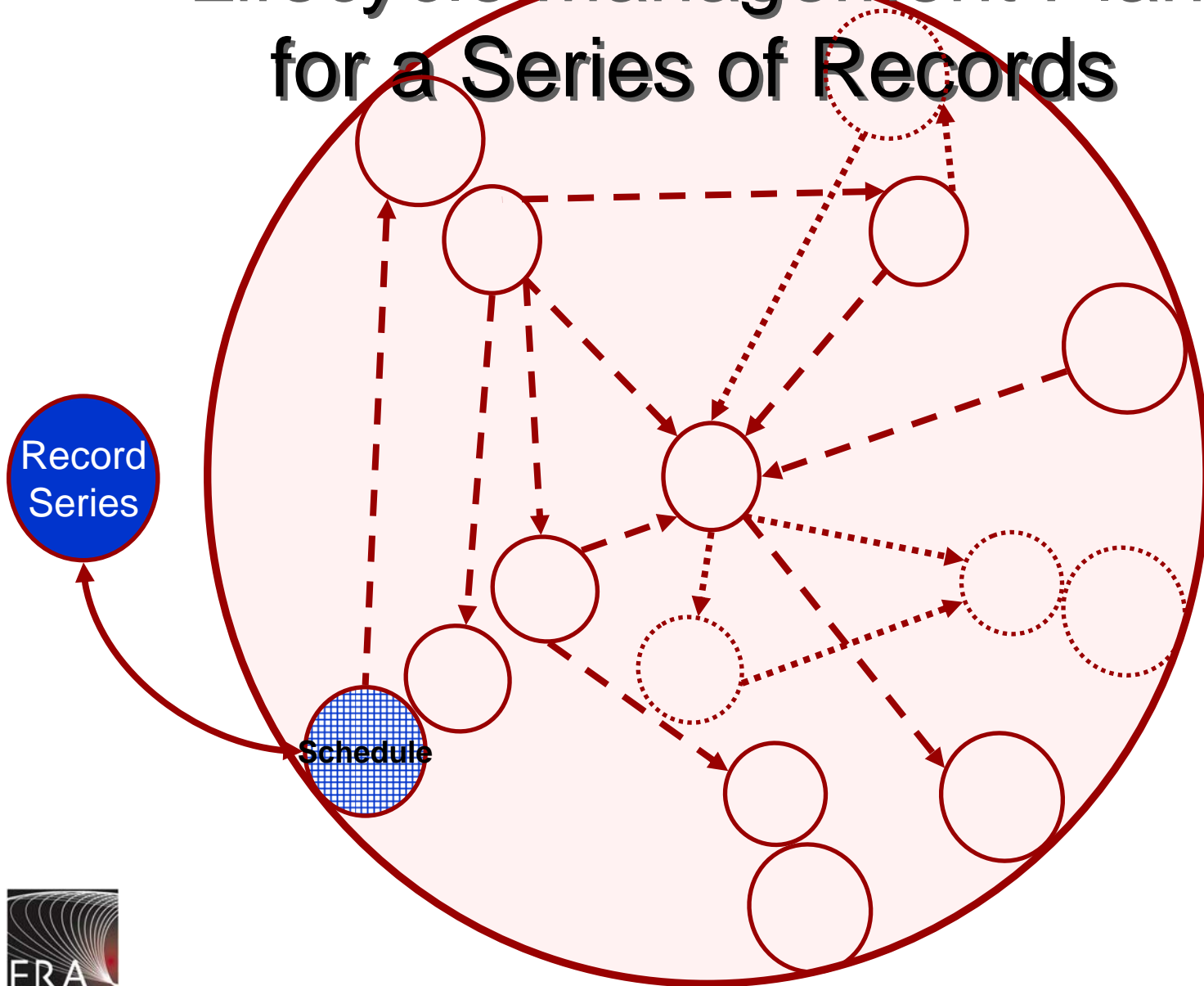
Lifecycle Plan for an Aggregate of Records



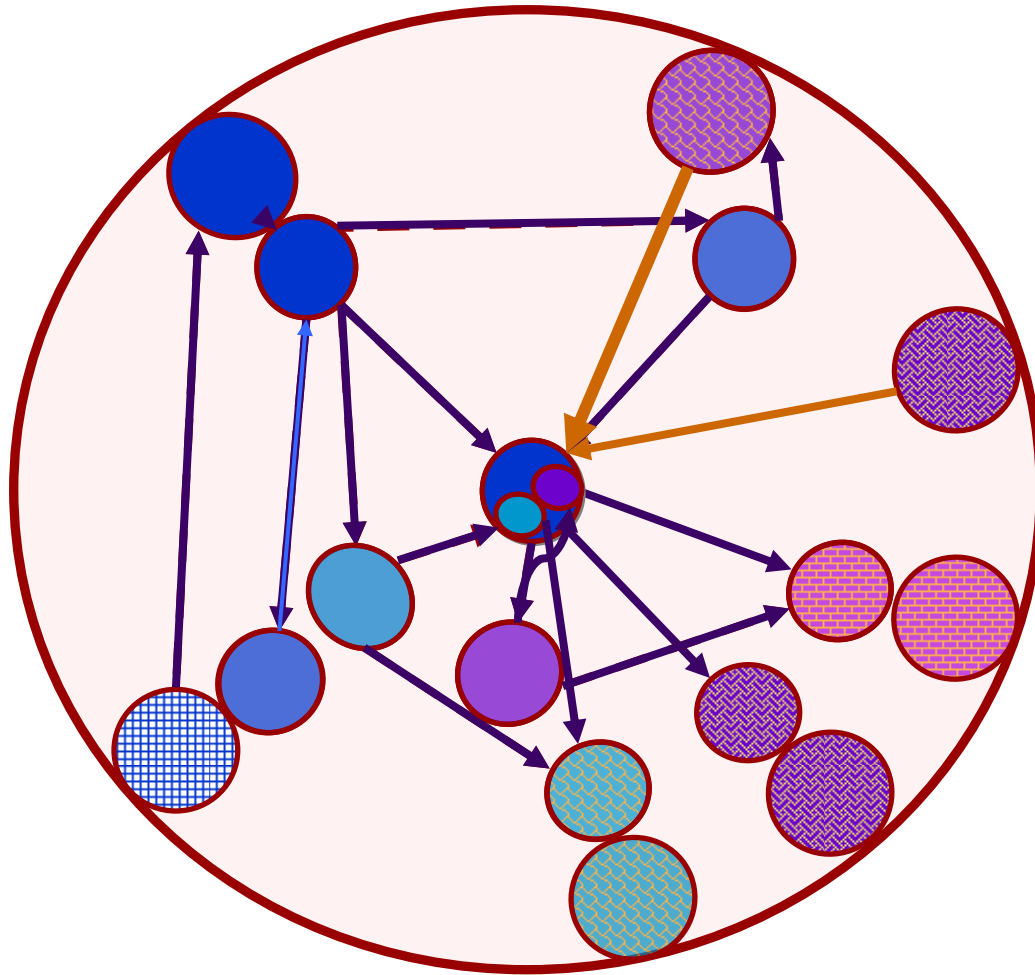
'Mini-System' Within the Electronic Records System



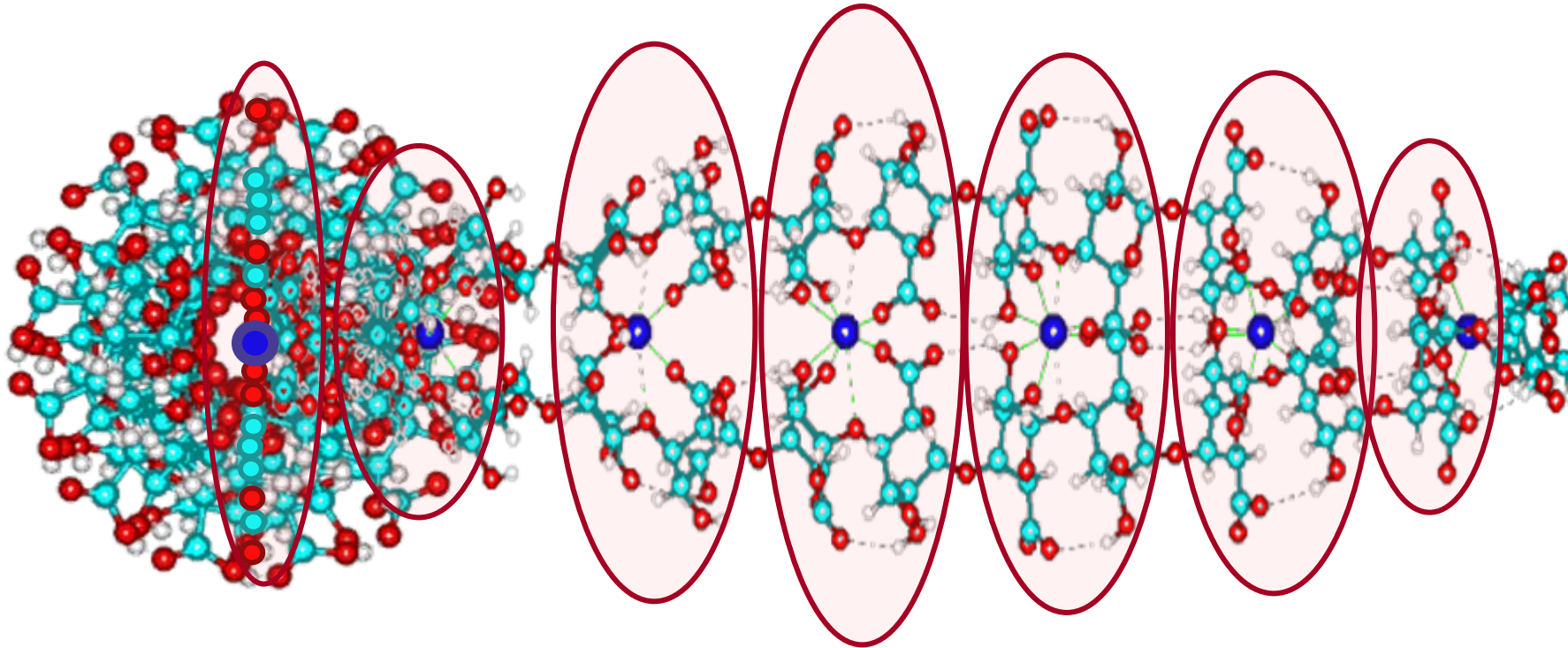
Lifecycle Management Plan for a Series of Records






The Lifecycle Management Plan: the basis for systematic Preservation and Access to a Body of Records



ERA as a Set of Mini-Systems

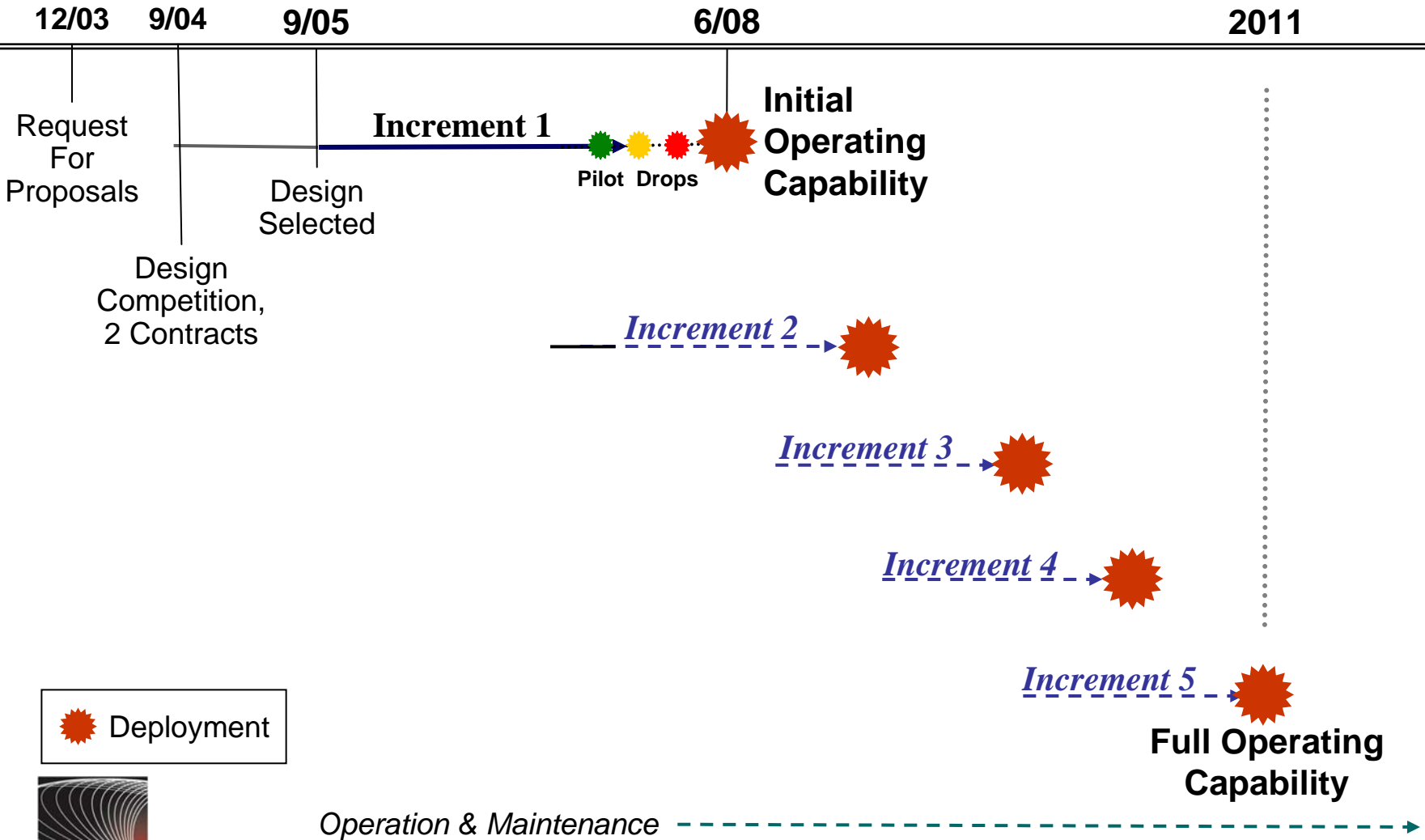


 Records Aggregate

 Lifecycle Management Plan


Each “Mini-system” is defined by the Lifecycle Management Plan for a Records Aggregate, such as a series or fonds

System Acquisition



 Deployment




Operation & Maintenance 

Incremental Development: 1 & 2

- Increment 1
 - SBU System for the National Archives (“Base System”)
 - Creation, review and approval of records schedules
 - Requests to transfer records, transfer of legal custody
 - Electronic records: transfer, inspection, archival storage
- Increment 2
 - SBU system for electronic records of the G W. Bush Administration
 - Rapid ingest, indexing, transformation to more accessible form.
 - Archival storage
 - Full content search
 - Basic case management for special requests, review results.
 - NSS system for presidential electronic records of the G W. Bush Administration
 - Same capabilities as SBU system
 - Enhancements to the National Archives SBU system
 - Appraisal of records



Incremental Development: 3 - 5

- *Increment 3*
 - *Public access to the Base System*
 - *Long-term preservation capability*
- *Increment 4*
 - *FOIA and other access case management*
 - *Review and redaction* 
- *Increment 5*
 - *Further enhancements, expansion*

Thank you.



For more information:

www.archives.gov/electronic_records_archives/index.html

