

# **Digital Preservation: Experiences and Thoughts of An American Archivist**

**Fynnette Eaton**

International Conference on Digital Preservation at the occasion of  
the retirement of J. Steenbakkers

2 November 2007

# Agenda

- Historical development of digital preservation at the U.S. National Archives
- Development of criteria for Electronic Records Archives (ERA) System
- Lockheed Martin's interpretation of system requirements for the ERA System

# Preservation of Digital Documents

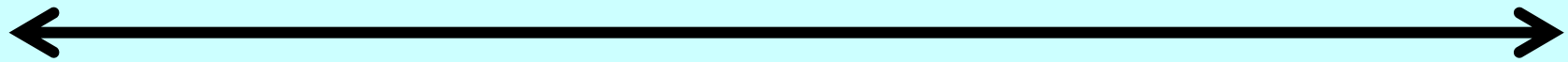
- It is impossible to preserve a digital document as a physical object.
  - To access a digital document, stored bit sequences must be interpreted as logical objects and presented as conceptual objects.
- It is only possible to preserve the ability to reproduce the document.
  - The process of digital preservation is complete only when the document is presented or rendered.

# Current Challenges

- **Diversity**
  - office automation, image, video and audio formats
- **Complexity**
  - decision support systems, GIS, applets and interactive WWW pages
- **Volume**, both in bytes and files
- **Rapidly changing nature of systems used to create records**

# Criteria for determining best option for preserving authentic digital information

- Feasibility
- Sustainability
- Practicality
- Appropriateness



## Preserve Technology

- Keep data in 'original' format
- Maintain original technology or equivalent
- Migrate data formats as technology changes
- Use state-of-the-art technology

## Preserve Objects

- Keep data in persistent form
- Export to state-of-the-art technology as technology changes

# Persistent Object Method



- Characterize significant properties of the things that are to be preserved.
- Express these properties in formal models
- Encapsulate objects in metadata defined in the models.
- Use software “mediators” to enable future technologies to interpret the models and metadata
  - to rebuild and repopulate collections
  - to re-present the records
  - support information discovery and delivery.

# *ERA Requirements*



- **Persistent**

- To manage and access the records over time.

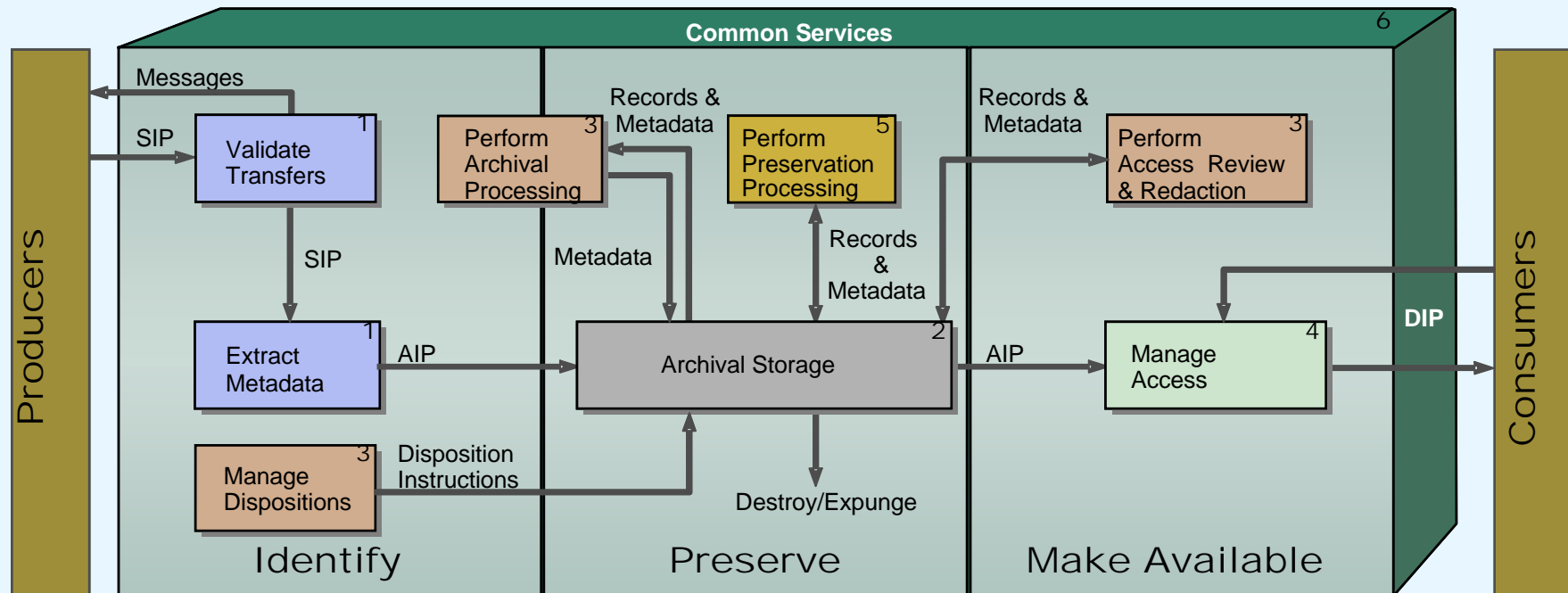
- **Authentic**







- To ensure that these are the original records
- Records that are created with attached documentary information

- **Scalable**

- To grow and adapt to increasing volumes and evolving types of electronic records
- To serve a variety of user groups

# Evolvability, Scalability, Extensibility: System Architecture

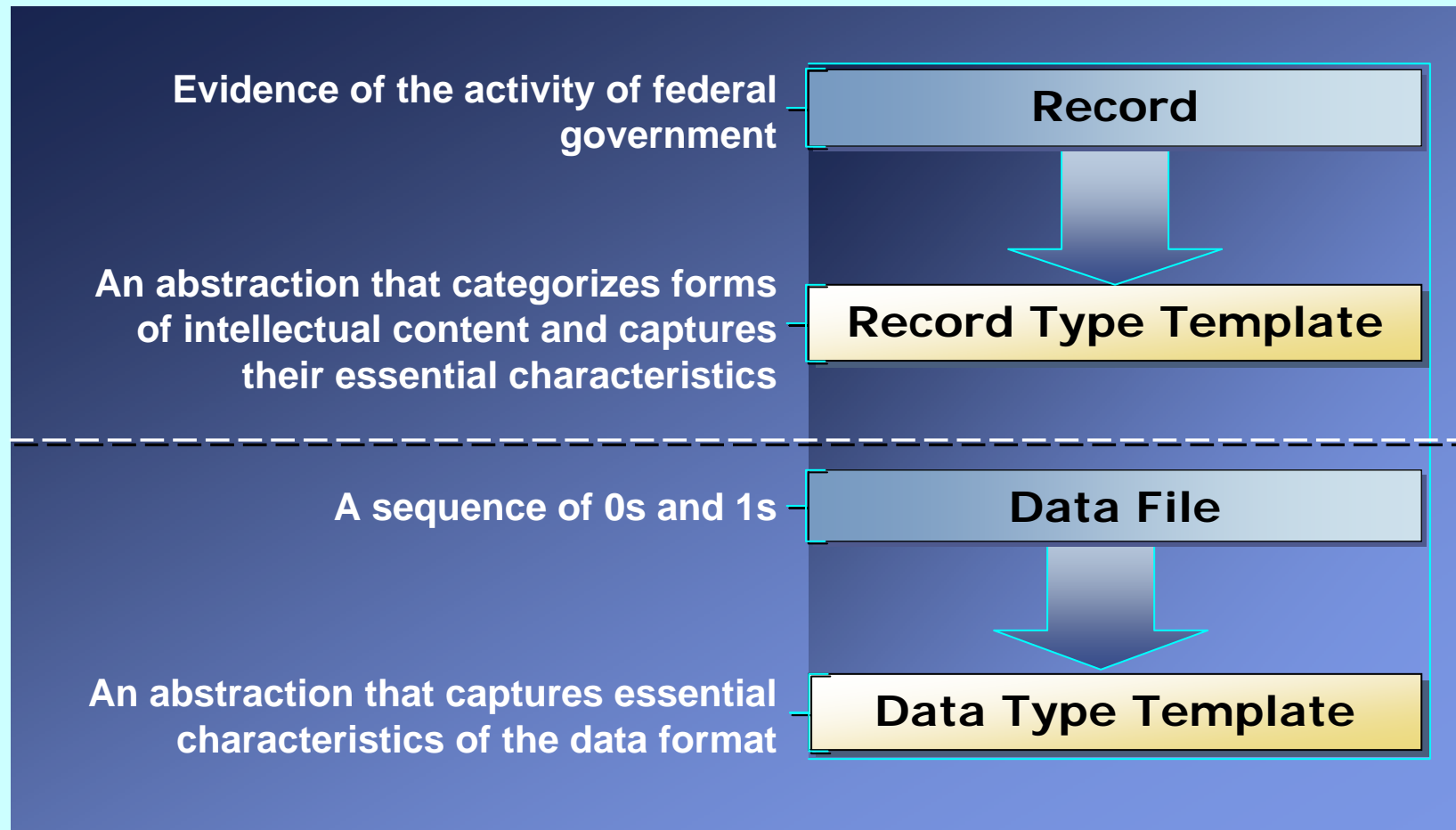


Legend:	OAIS Functions	ERA System -Level Packages	Service Oriented Architecture
	1 - Ingest	Ingest	Business Application Services
	2 - Archival Storage	Archival Storage	
	3 - Data Management	Records Management	Common Infrastructure Services
	4 - Access	Dissemination	
	5 - Preservation	Preservation	
	6 - Common Services	Local Services & Control, ERA Management	

# Principal Design Considerations

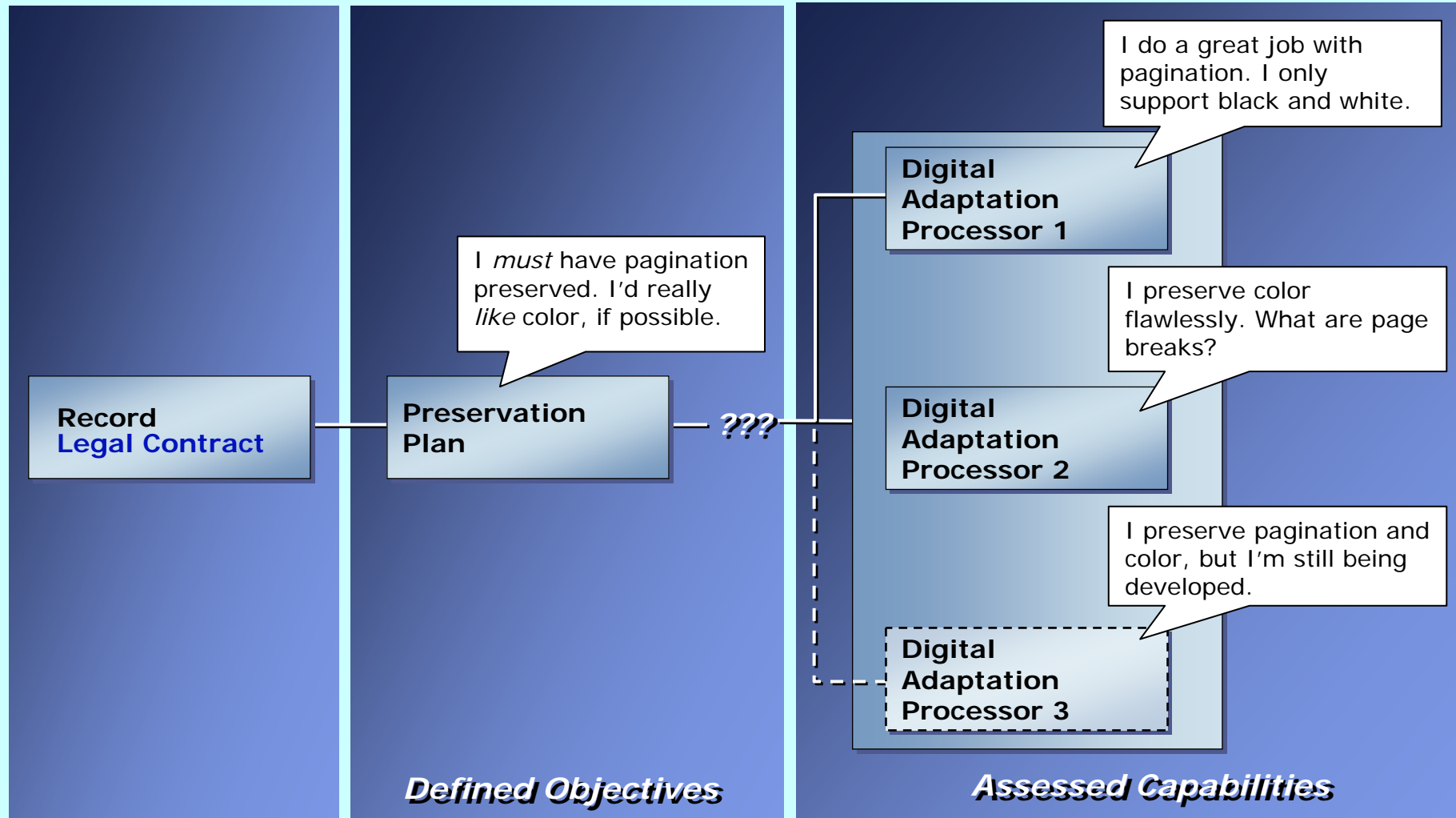
- **Evolvability**
  - Policy neutral
  - Allow ERA to change over time
- **Scalability**
  - Scale up or down, with volume and usage
  - Scale to other types of archives
- **Extensibility**
  - Easily add additional features in the future without major modification
- **Availability**
  - Have no single point of failure
  - Maximize "up time," but balance availability with cost

# Templates Drive Preservation



***Templates Capture Essential Characteristics***

# Preservation Planning Solution



# Dealing with Complexity

If it were only simple...



Real world is more complex...

