Identifying Requirements for an Electronic Records Processing Environment

Leslie Johnston, NARA
2014 LOC Preservation Storage Meeting

Optimized Ingest Framework (OIF)

A new model at NARA for managing the receipt and processing of electronic records of all formats for preservation and access

- Modular approach to NARA systems managing digital materials
- Departure from model of a single, monolithic system to manage the ingest, processing, preservation, and access to electronic records.
- Sustainable approach for necessary evolution of existing systems.
- Establishment of an environment to provide necessary processing flexibility and tools for a wide variety of digital files and metadata.
- More robust solution for digital preservation with reduced complexity.
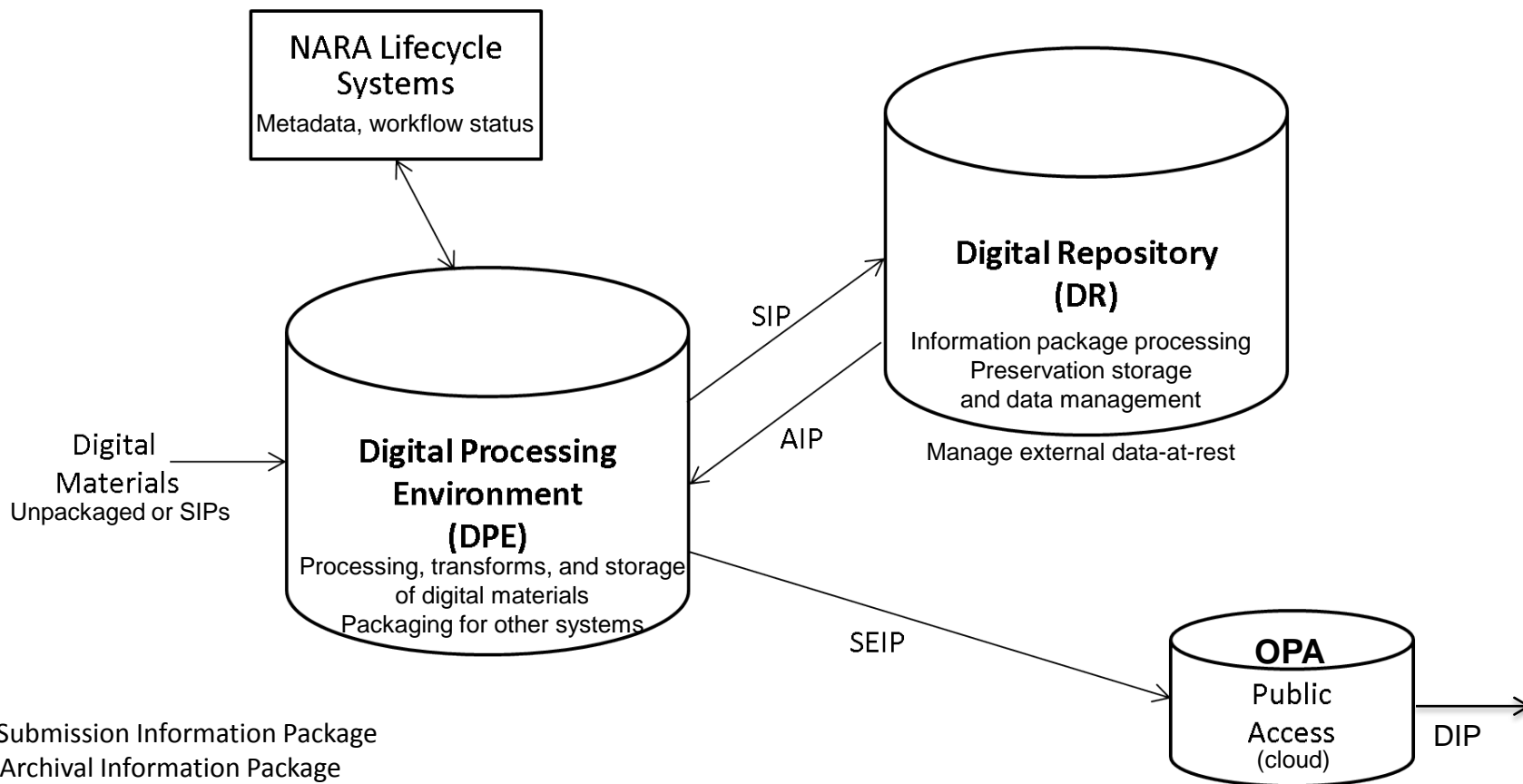- A stepping stone toward greater automation in managing digital files.

Optimized Ingest Framework (OIF)

Developed in a series of 2013 meetings with Stakeholders
- Developed OIF Vision
- Developed High-level Stakeholder Requirements for DPE

Concepts developed to address number of needs & factors:
- Need for a flexible processing environment with an expandable set of software tools to verify and process electronic records.
- Requirement to manage increasing volume and varieties of digital files.
- No current workflows for non-Federal digital materials (e.g., digital surrogate masters, Legislative, Donated, Supreme Court, etc.).
- Growing storage and access needs for digital surrogates.
- Need for a more efficient means to provide public access to ingested electronic records from agencies.
- Need to explore & develop capabilities and processes to address projected growth – Presidential Directive implications, Data-at-Rest solutions.

NARA Lifecycle Systems
Metadata, workflow status

Digital Repository (DR)
Information package processing
Preservation storage
and data management

Manage external data-at-rest

Digital Materials
Unpackaged or SIPs

Digital Processing Environment (DPE)
Processing, transforms, and storage
of digital materials
Packaging for other systems

SIP

AIP

SEIP

OPA
Public Access
(cloud)

DIP

SIP – Submission Information Package
AIP – Archival Information Package
SEIP – Search Engine Information Package
DIP – Dissemination Information Package

Opportunities in the Pilot

- Test assumptions made in the design of processing workflows and selection of tools for processing.

- Develop data packaging specifications and APIs for the exchange of metadata and files between multiple systems.

- Implement the pilot for both an on premise and a cloud environment to test system specifications, performance, and storage needs.

- Benchmark the use of a range of processing and format transformation tools for both a cloud and on premise environment for processor, memory, and storage needs.

- Benchmark network transfers between on premise and cloud environments and between cloud environments.

Leslie Johnston
NARA
leslie.johnston@nara.gov