

OAIS and It's Applicability for Libraries, Archives, and Digital Repositories

Robin Dale

RLG

Robin.Dale@notes.rlg.org

Characteristics of Digital Information in Libraries and Archives

- Diversity
 - of applications (commercial, “home-grown,” operational, etc.)
 - in the organization, structure and encoding of documents and data
- Complexity
 - varies greatly
- Volume
 - Previously low, but growing at exponential rates.

Characteristics and Diversity of Information to be Archived

- Data file collections
- Images
 - JPEG, GIF, TIFF, PS formats
- Text documents
 - PDF, TXT, HTML, XML formats
- Audio and video
 - WAV, AIFF, MPEG, etc.
- Other data types
- Links to other archives

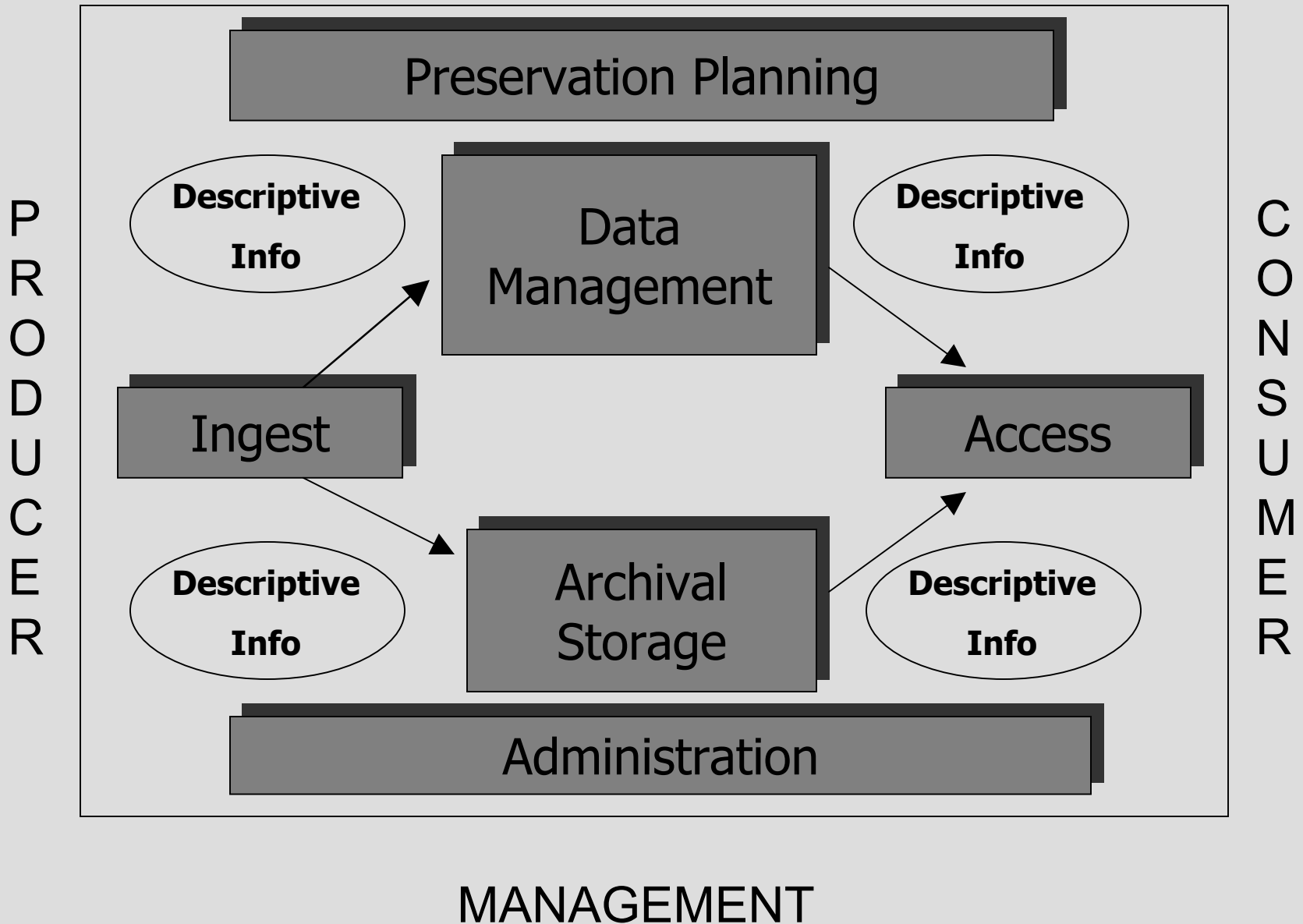
The Difficulties of Preserving that Information

- Technical factors
 - increasingly frequent technological changes
 - non-independence of archived information with regard to systems
 - control of migration process
- Organizational factors
 - organizational change
 - priority shift in tandem with leadership shift
- Scientific factors
 - the knowledge to interpret certain data/datatypes

Open Archival Information System (OAIS)

- Open Archival Information System Reference Model
 - Framework for understanding and applying concepts needed for long-term preservation of digital information
 - Adaptation and development to fit certain information types (digital images, space systems data, etc.) left to those who adopt the model
 - Facilitates description and comparison of archives (a common vocabulary of concepts)
 - Addresses a full range of archival functions
 - Basis for development of additional, related standards (OAIS Reference Model - April 2001 version - is a draft ISO standard)
 - Broad, international interest and work moving forward

Simplified OAIS Model



OAIS Functional Entities

- Ingest
- Archival Storage
- Data Management
- Administration
- Preservation Planning
- Access

From Framework to Service

- Implementation of OAIS framework refers to a set of *services*
 - functional requirements
 - agreements
 - production systems
 - education and training
 - business models for sustainability and growth
- Wide range of digital archiving services cannot be accomplished with off-the-shelf “solutions” for storage and digital asset management

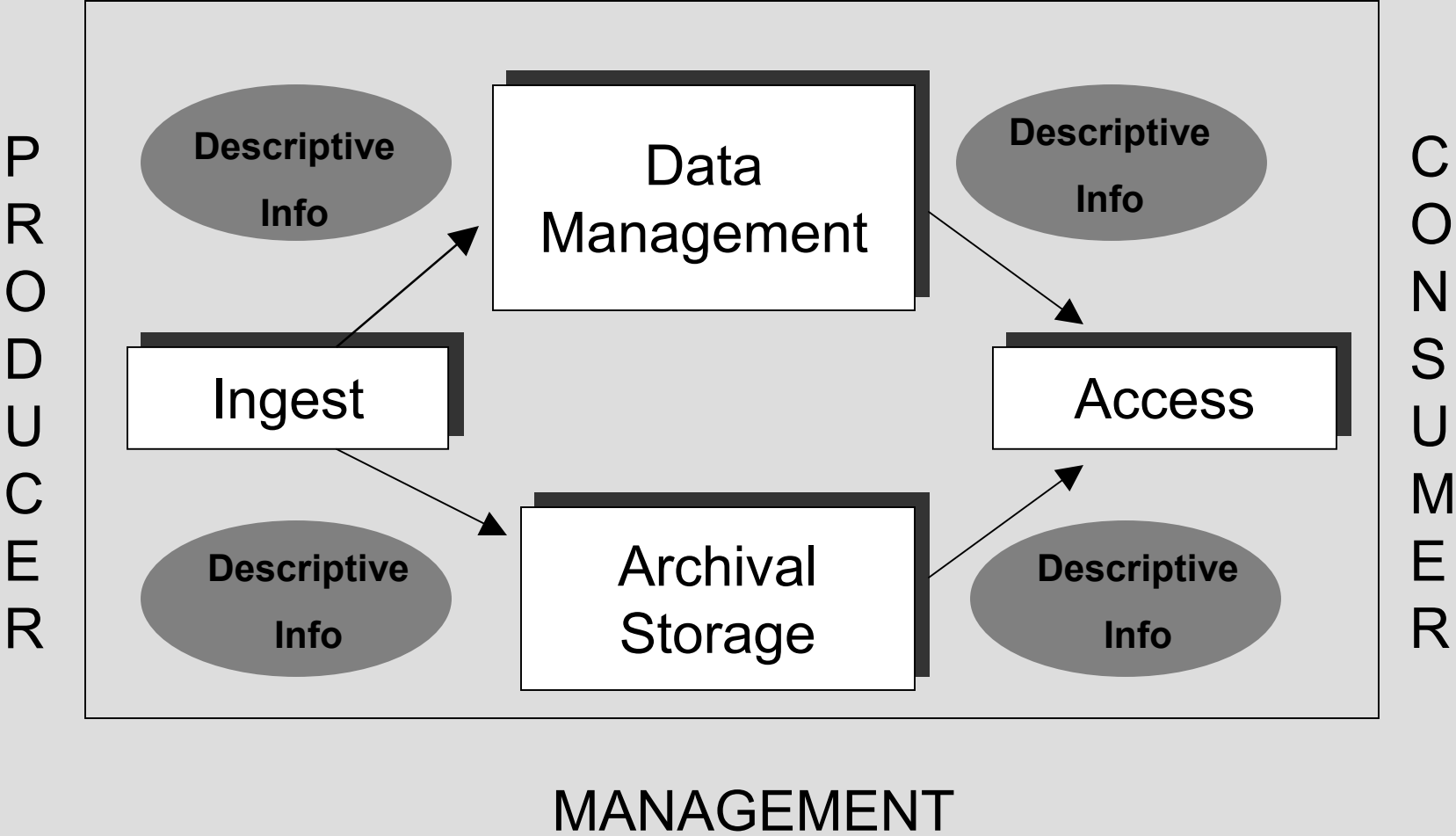
So Who is Using the OAIS?

- Research and development projects are underway; several have been completed
 - CEDARS (CURL Exemplars in Digital Archives)
 - NEDLIB(Networked European Deposit Library)
 - National Library of Australia
 - San Diego Supercomputer Center on behalf of NARA
- Libraries are in various stages of design, testing and implementation
 - Library of Congress
 - Harvard
 - Stanford

Community Collaboration: RLG & OCLC

- March 2000:
 - RLG and OCLC announced commitment to collaborate on identifying and supporting best practices for long-term retention of digital objects
 - facilitate consensus-building among stakeholders
- Collaboration in two areas:
 - Attributes of digital repository for research resources
 - Preservation metadata for long-term retention of digital objects

Preservation Metadata



OCLC/RLG Working Group on Preservation Metadata

- Characterize current practice/thinking on the use of metadata to support digital preservation
- Scope
 - Develop a comprehensive preservation metadata framework
 - Identify essential preservation metadata elements to support the framework
 - Identify and evaluate alternative implementation approaches
 - Develop testbed/pilot applications
 - Recommend best practices/common approaches for the use of metadata to support digital preservation

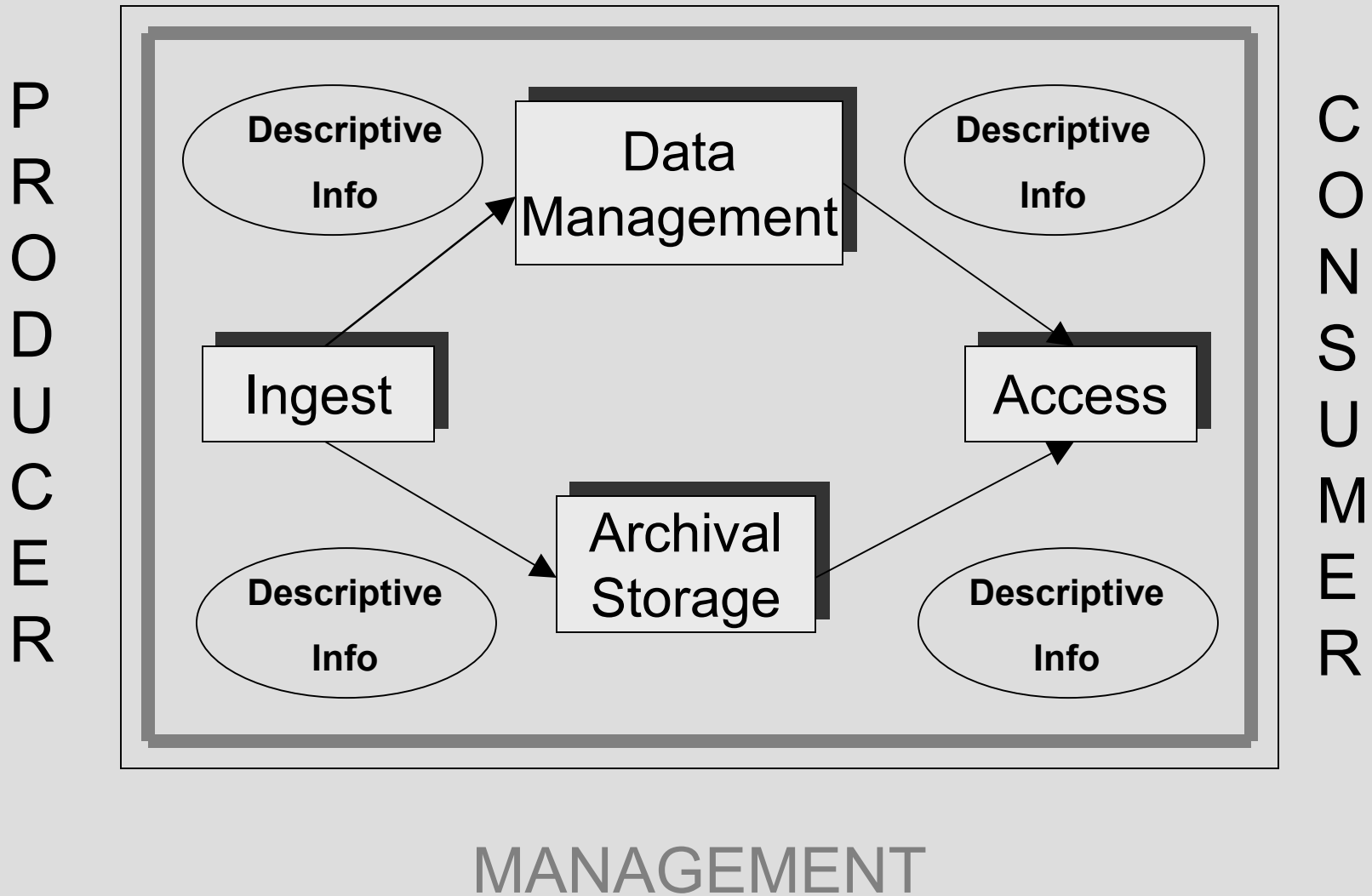
Metadata Strategy

- Convened working group of leading experts in preservation metadata
- White paper characterizing “state of the art” in preservation metadata:
 - reviewed existing preservation metadata element sets
 - identified convergence/divergence of these sets
 - identified potential areas for consensus building
- Build on existing work to:
 - identify “essential” preservation metadata elements
 - best practices/guidance for implementation
 - testbed/pilot applications

Metadata Findings

- White Paper: *Preservation Metadata for Digital Objects: A Review of the State of the Art*
www.oclc.org/digitalpreservation/presmeta_wp.pdf
- Currently engaged in addressing the issues raised in the paper; revising to include working group feedback
 - Content information is current topic
- Group's experience has underscored both the importance and the challenges of reaching consensus on preservation metadata.

Defining Attributes



RLG/OCLC Working Group on Digital Archive Attributes

- Scope:
 - Create a definitional document describing the characteristics of reliable, digital archives capable of handling heterogeneous research collections
 - Identify tools to support institutions in seeking or building archiving services
 - Present recommendations to RLG and OCLC suggesting opportunities and solutions for developing preferred tool sets
 - Identify any follow-on activities required

Attributes Strategy

- Draft report written by Kelly Russell (former Project Manager, Cedars Project, UK)
- Draft reviewed by working group members and invited experts
- Document revised and strengthened based on comment
- Document to be released for multi-month public review and comment period
- Final report and recommendations released in 4th quarter, 2001

Attributes Findings

- Two distinct types of “attributes” identified
 - Qualities
 - Responsibilities
 - As outlined by OAIS Reference Model and applied project experience (Cedars Project)

Further work

- Preservation Metadata
 - Set of “essential” preservation metadata elements
 - Best practices/guidance for implementation
 - Testbed/pilot applications
- Digital Repository Attributes
 - Public draft to be released within weeks
 - Available at: www.rlg.org/longterm/
 - Working group incorporates comments and issues final report in 4th quarter (possibly October) 2001

Summary

- OAIS provides:
 - a model
 - a common vocabulary
 - a useable framework
 - is adaptable/scalable to community and size of collection
 - an emerging international standard