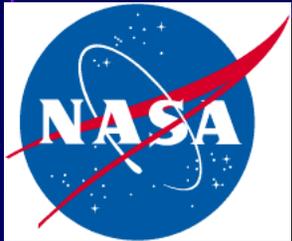


# Reference Model for an Open Archival Information System (OAIS)

*Partnerships in Innovation: Serving a  
Networked Nation*



*Donald Sawyer/NASA/GSFC*

*Lou Reich/NASA/CSC*

*15-November-2004*

# Purpose and Scope: 1

- Framework for understanding and applying concepts needed for long-term digital information preservation
  - Long-term is long enough to be concerned about changing technologies
- Also can be starting point for model addressing non-digital information

## Purpose and Scope: 2

- Provides set of minimal responsibilities to distinguish an OAIS from other uses of 'archive'
- Framework for comparing architectures and operations of existing and future archives

## Purpose and Scope: 3

- Basis for development of additional related standards
- Addresses a full range of archival functions
  - Ingest, Archival Storage, Data Management, Access, Preservation Planning, Administration

# Applicability

- Applicable to all long-term archives and those organizations and individuals dealing with information that may need long-term preservation
- Does NOT specify an implementation

# Conformance

- How does an archive conform?
  - It discharges the set of minimal responsibilities
  - It supports the basic information concepts that address a definition of information and types of information packages
- How do other documents conform?
  - By using OAIS terms and concepts

# Who wants to conform to OAIS?

- All organizations that need to preserve digital information for extended periods
  - To demonstrate a level of awareness of digital preservation needs
- Other standards and documents
  - For effective communication and integration

# Open Archival Information System (OAIS)

- Information
  - Any type of knowledge that can be exchanged
  - Data are the representation forms of information
- Archival Information System
  - Hardware, software, and people who are responsible for the acquisition, preservation and dissemination of the information

# View of an OAIS Environment

- **Producer** provides the information to be preserved
- **Management** sets overall OAIS policy
- **Consumer** seeks and acquires preserved information of interest



# OAIS Responsibilities: 1

- Negotiates and accepts information from information producers
- Obtains sufficient control to ensure long-term preservation
- Determines which communities (designated) need to be able to understand the preserved information

## OAIS Responsibilities: 2

- Ensures the information to be preserved is independently understandable to the Designated Communities
- Follows documented policies and procedures that ensure the information is preserved against all reasonable contingencies
- Makes the preserved information available to the Designated Communities in forms understandable to those communities

# OAIS Information Definition

- Information is always expressed (i.e., represented) by some type of data
- Data interpreted using its Representation Information yields Information



# Content Information

- The information that is the original target of preservation
- Deciding what is the Content Information may not be obvious and may need to be negotiated with the Producer
- The Content Data Object in the Content Information may be either a Digital Object or a Physical Object (e.g., microfilm, a physical sample)

# Preservation Description Information (PDI): 1

- Reference Information
  - Provides one or more identifiers, or systems of identifiers, by which the Content Information may be uniquely identified
- Provenance Information
  - Describes the source of Content Information, who has had custody of it, what is its history

# Preservation Description Information (PDI): 2

- Context Information
  - Describes how the Content Information relates to other information outside the Information Package
- Fixity Information
  - Protects the Content Information from undocumented alteration

# Examples of PDI

- Reference
  - Bibliographic description; Persistent Ids
- Provenance
  - Metadata on preservation process
- Context
  - Pointers to related collections
- Fixity
  - Digital signatures, checksums

# Information Package Definition

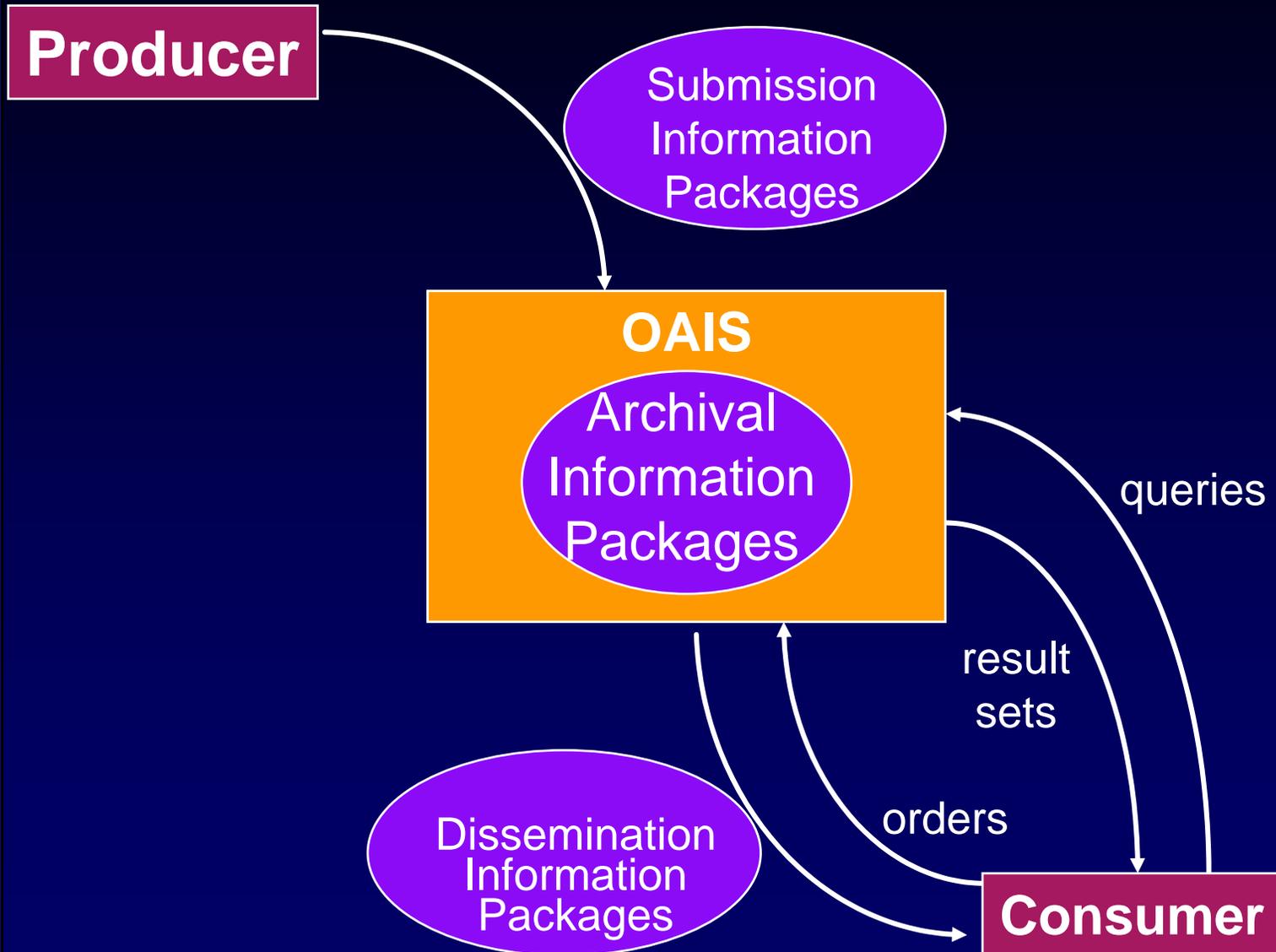


- An **Information Package** is a conceptual container holding two types of information
  - **Content Information**
  - **Preservation Description Information (PDI)**

# Information Package Variants

- **Submission Information Package**
  - Negotiated between Producer and OAIS
  - Sent to OAIS by a Producer
- **Archival Information Package**
  - Information Package used for preservation
  - Holds complete set of Preservation Description Information for the Content Information
- **Dissemination Information Package**
  - Includes part or all of one or more Archival Information Packages
  - Sent to a Consumer by the OAIS

# External Data Flow View



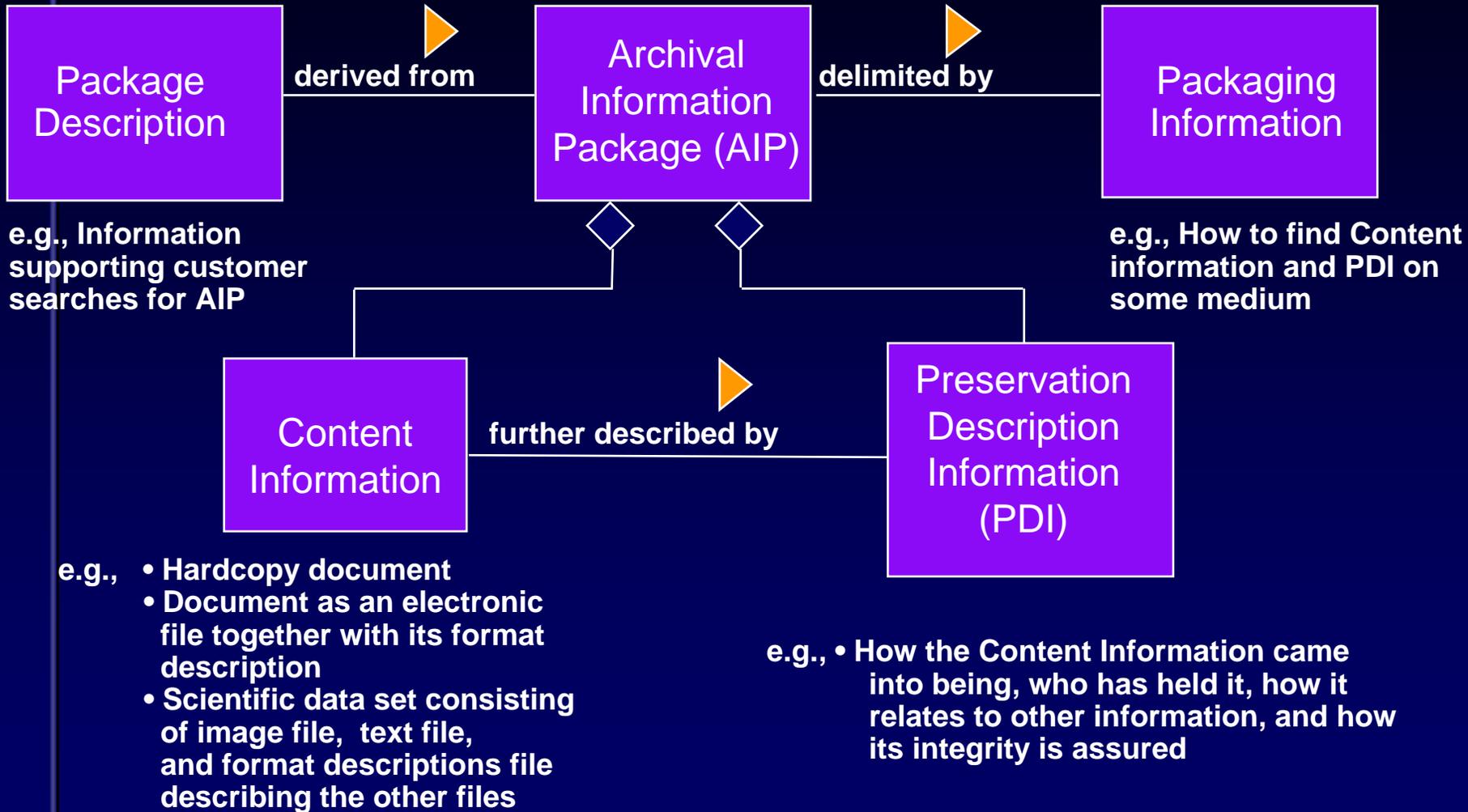
# Packaging Information

- Information which, either actually or logically, binds and relates the components of the package into an identifiable entity on specific media
- Examples of Packaging Information include directory structures, filenames, and tape marks

## Package Description

- Contains the data that serves as the input to documents or applications called Access Aids.
- Access Aids can be used by a Consumer to locate, analyze, retrieve, or order information from the OAIS.

# OAIS Archival Information Package



# Functional Entities In An OAIS: 1

- **Ingest:** This entity provides the services and functions to accept Submission Information Packages (SIPs) from Producers and prepare the contents for storage and management within the archive
- **Archival Storage:** This entity provides the services and functions for the storage, maintenance and retrieval of Archival Information Packages

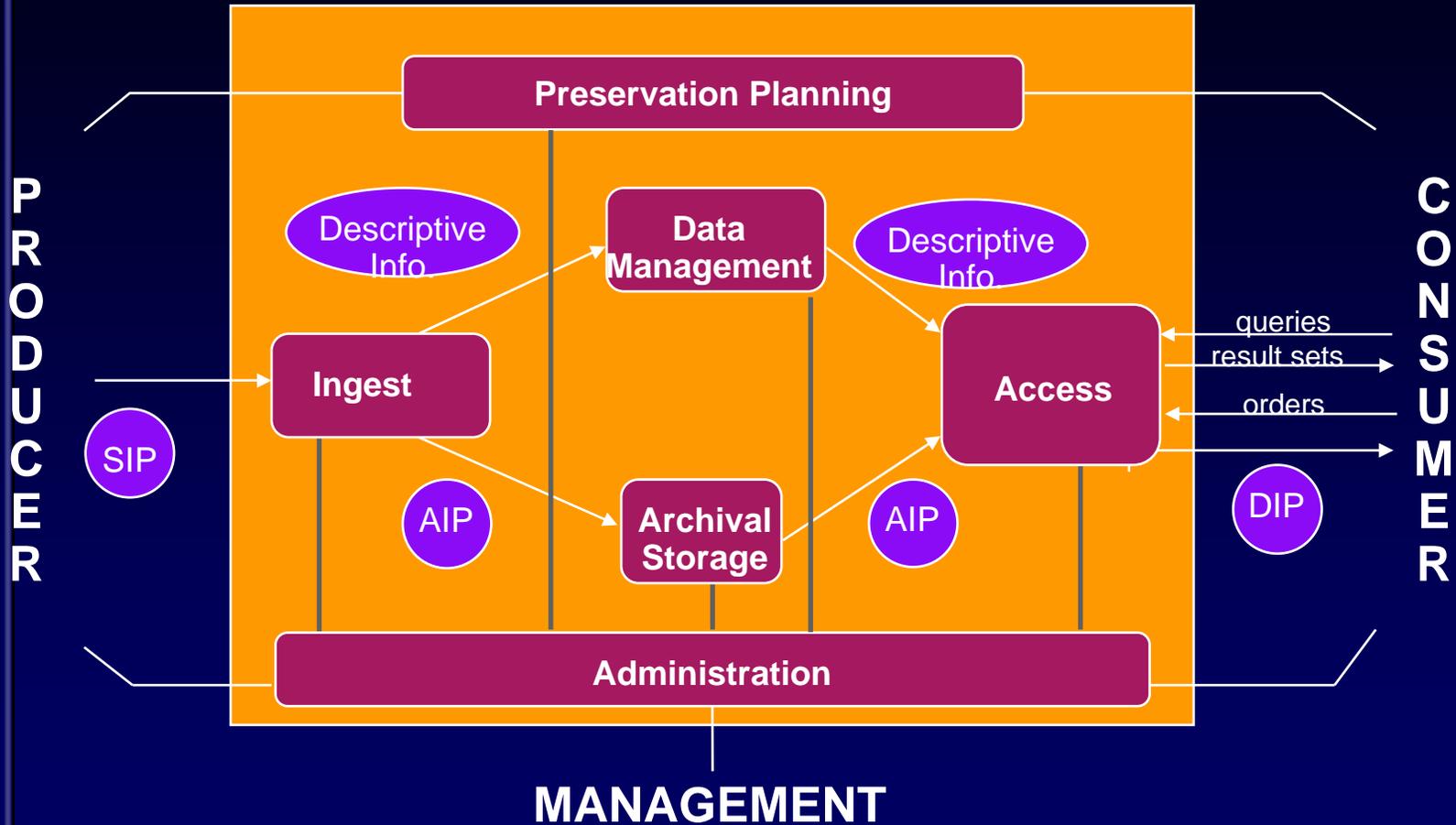
## Functional Entities In An OAIS: 2

- **Data Management:** This entity provides the services and functions for populating, maintaining, and accessing both descriptive information that identifies and documents archive holdings and internal archive administrative data.
- **Administration:** This entity manages the overall operation of the archive system

# Functional Entities In An OAIS: 3

- **Preservation Planning:** This entity monitors the environment of the OAIS and provides recommendations to ensure that the information stored in the OAIS remain accessible to the Designated Community over the long term.
- **Access:** This entity supports Consumers in determining the existence, description, location and availability of information stored in the OAIS and allows Consumers to request and receive information products

# OAIS Functional Entities



SIP = Submission Information Package

AIP = Archival Information Package

DIP = Dissemination Information Package

# Reference Model Summary

- Reference model is applicable to all digital archives, and their Producers and Consumers
- Establishes common terms and concepts for comparing implementations, but does not specify an implementation
- Identifies a minimum set of responsibilities for an archive to claim it is an OAIS
- Provides detailed models of both archival functions and archival information
- Discusses OAIS information migration and interoperability among OAISs

# Reference Model Status

- Widely adopted as starting point in digital preservation efforts
  - Digital libraries (e.g., Netherlands National Library)
  - Traditional archives (e.g., US National Archives)
  - Scientific data centers (e.g., National Space Science Data Center)
  - Commercial Organizations (e.g., Aerospace Industries Association preservation working team)
- Published as final CCSDS standard (Blue Book) available from:  
<http://www.ccsds.org/documents/650x0b1.pdf>
- Published as a final ISO standard: ISO 14721: 2003

# ISO Archive Follow-on Activities: 1

- Producer-Archive Interface Methodology Abstract Standard
  - Provides framework for Producer/Archive interactions
  - Identifies steps and types of information exchanged during the 'negotiation'
  - May be used as a checklist by archives
  - Published as final CCSDS standard (<http://www.ccsds.org/documents/651x0b1.pdf>)
  - Currently undergoing ISO review

## ISO Archive Follow-on Activities: 2

- XML Formatted Data Unit (XFDU) Structure and Construction Rules
  - Draft standard addressing packaging of science data and metadata for exchange and preservation
- Specification for the Transfer Phase of the Producer-Archive Interface
  - Draft standard addressing formal mechanisms to define a Submission Information Package for transfer and validation

# Related Standards Efforts

- METS (Metadata Encoding and Transmission Standard)
  - XML based encoding of metadata regarding objects within a digital library (<http://www.loc.gov/standards/mets/>)
- OCLC/RLG Working Group on Preservation Metadata
  - Developed metadata framework to support preservation of digital objects (<http://www.oclc.org/research/projects/pmwg/>)
- RLG/NARA Task Force on Digital Repository Certification
  - Producing certification requirements for establishing and selecting reliable digital information repositories (<http://www.rlg.org/longterm/certification.html>)