

Standards for Building Trust in Digital Repositories

NARA/UMD

**Partnerships in Innovation II:
From Vision to Reality and Beyond**

7 October 2008

Donald Sawyer - VIE Inc.

Overview

- Why is there the issue of 'trust in digital repositories'?
- Some Standards for Building Trust
- Development of the OAIS reference model
- Development of the TRAC
- Development of PAIS
- Conclusion

Why is There a Trustworthy Issue?

- Preserving information in digital forms is much more difficult than in forms such as paper and film.
 - Not only a problem for traditional archives, but also for many organizations that have never thought of themselves as performing an archival function
- Digital information preservation does not have the long historical experience that exists for traditional archival science
 - Digital preservation is an active area of research
 - Academic courses addressing digital preservation are few but growing

A Few Relevant Standards

- Standards play a key role in engendering trust in any preservation effort
 - Adherence constrains processes and helps ensure consistency
 - Use of widely adopted standards helps build a consensus on trust
- Some noteworthy digital preservation related standards and efforts include:
 - Reference Model for an Open Archival Information System (OAIS)
 - Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC)
 - Producer-Archive Interface Standard (PAIS)
 - Producer-Archive Interface Methodology Abstract Standard (PAIMAS)
 - <http://public.ccsds.org/publications/archive/651x0b1.pdf>
 - PREMIS Data Dictionary for Preservation Metadata, v2
 - <http://www.loc.gov/standards/premis/v2/premis-2-0.pdf>
 - CCSDS/ISO Digital Repository Audit and Certification Working Group
 - <http://wiki.digitalrepositoryauditandcertification.org/bin/view/Main/WebHome>

OAIS Reference Model: Motivation

- Reference Model for an Open Archival Information System (OAIS)
 - By establishing minimum requirements for an OAIS archive along with a set of archival concepts
 - Provide a common framework from which to view digital archival challenges,
 - Enable more organizations to understand the issues and take the proper steps to ensure Long Term information preservation,
 - Provide a basis for more standardization and, therefore, a larger market that vendors can support.

OAIS Reference Model: Process

- Developed under Consultative Committee for Space Data Systems and ISO processes
 - <http://public.ccsds.org/publications/archive/650x0b1.pdf>
- Organized US contribution under a framework with NASA lead and NARA support
- Conducted an “open” process to stimulate broad, international, dialogue
- Investigated other reference models

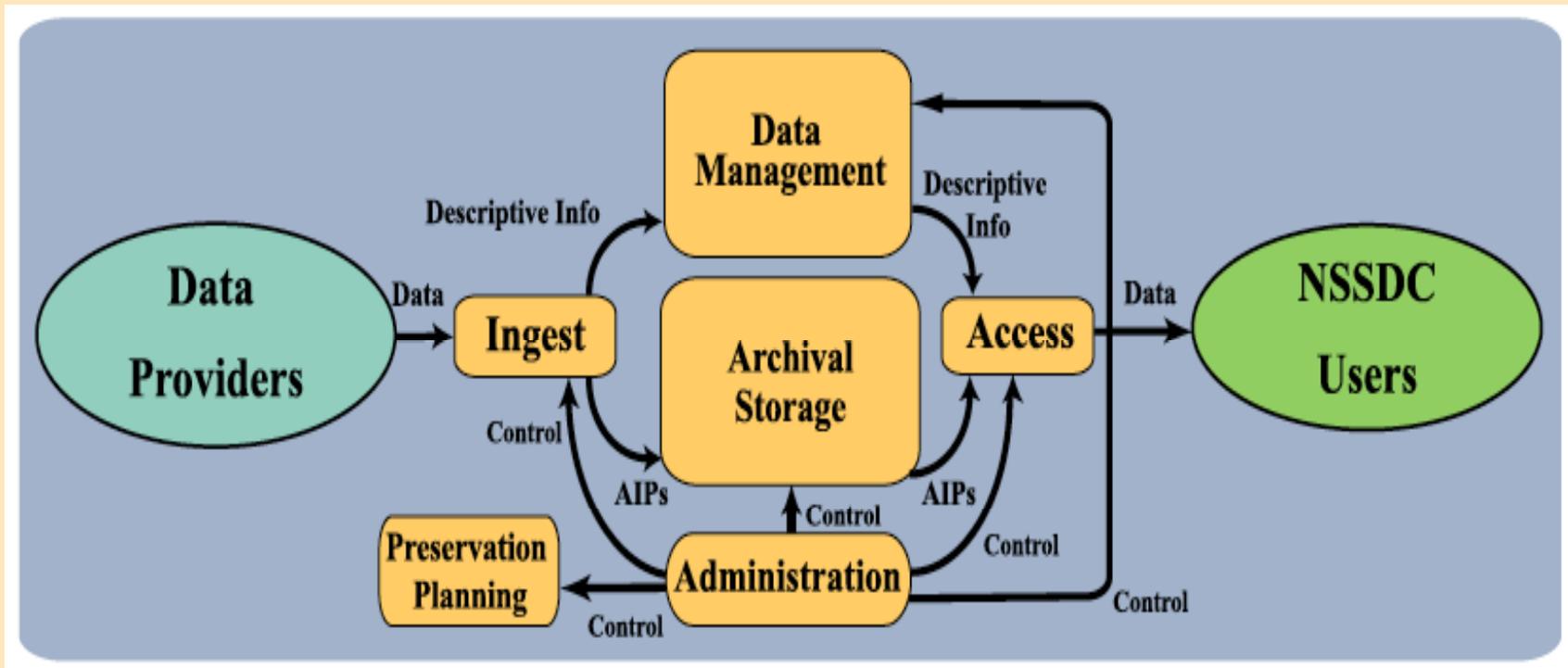
OAIS Reference Model: Result

- Reference Model targeted to several categories of reader
 - Archive designers
 - Archive users
 - Archive managers, to clarify digital preservation issues and assist in securing appropriate resources
 - Standards developers
- Adopted terminology that crosses various disciplines
 - Traditional archivists
 - Scientific data centers
 - Digital libraries
- 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)

OAIS Reference Model: NSSDC Usage

- National Space Science Data Center at NASA's Goddard Space Flight Center
 - Adopted OAIS functional model
 - Reporting to NASA Headquarters
 - Internal analysis of processes
 - Adopted OAIS information model
 - Terminology
 - Re-engineering archival storage processes, including explicit Archival Information Package specifications
 - Re-engineering ingest processes using Submission Information Packages

NSSDC Uses OAIS Concepts



Archive Information Packages
Archival Information Collections
Archive Information Units
Submission Information Packages
Content Information
Representation Information

Provenance Information
Context Information
Reference Information
Fixity Information
Designated Community

TRAC: Motivation

- Trustworthy Repositories Audit & Certification: Criteria and Checklist
 - OAIS provides terms and concepts, but is not an implementation prescription
 - How can anyone know if an archive/repository is following practices that will ensure long term digital preservation?
 - No community consensus
 - Still a research issue
 - Need to build consensus was frequently heard

TRAC: Process

- Developed under joint leadership of Research Libraries Group and NARA
 - <http://www.oclc.org/research/announcements/2007-03-12.htm>
- International ‘experts’, typically associated with existing archives/repositories, were invited to participate
- Past work was reviewed, particularly:
 - Trusted Digital Repositories: Attributes and Responsibilities, RLG-OCLC, May 2002
 - Many OAIS terms and concepts also adopted

TRAC: Results

- Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC), released May 2007
 - <http://www.crl.edu/PDF/trac.pdf>
- Provides criteria
 - requirements to be met by repositories,
 - example evidence that may be used to meet each requirement
- Provides an audit checklist
 - Forms to be filled in addressing evidence examined, findings and observations, and result
- Particularly convenient for an archive/repository to examine its processes and look for deficiencies
 - My view: no repository likely to meet all requirements

NSSDC TRAC Usage

- NSSDC reviewed its processes
 - Most requirements were met
 - Several were not, including:
 - Need better provenance documentation on internal processing and transformations
 - Insufficient disaster recovery planning
- NSSDC pared the requirements to a reduced set (about 80%) thought to be particularly relevant to NASA Space Sciences repositories
 - Provided these to Astrophysics community for consideration
 - Further paring being done by NASA HQ and may become a recommendation to Space Physics repositories

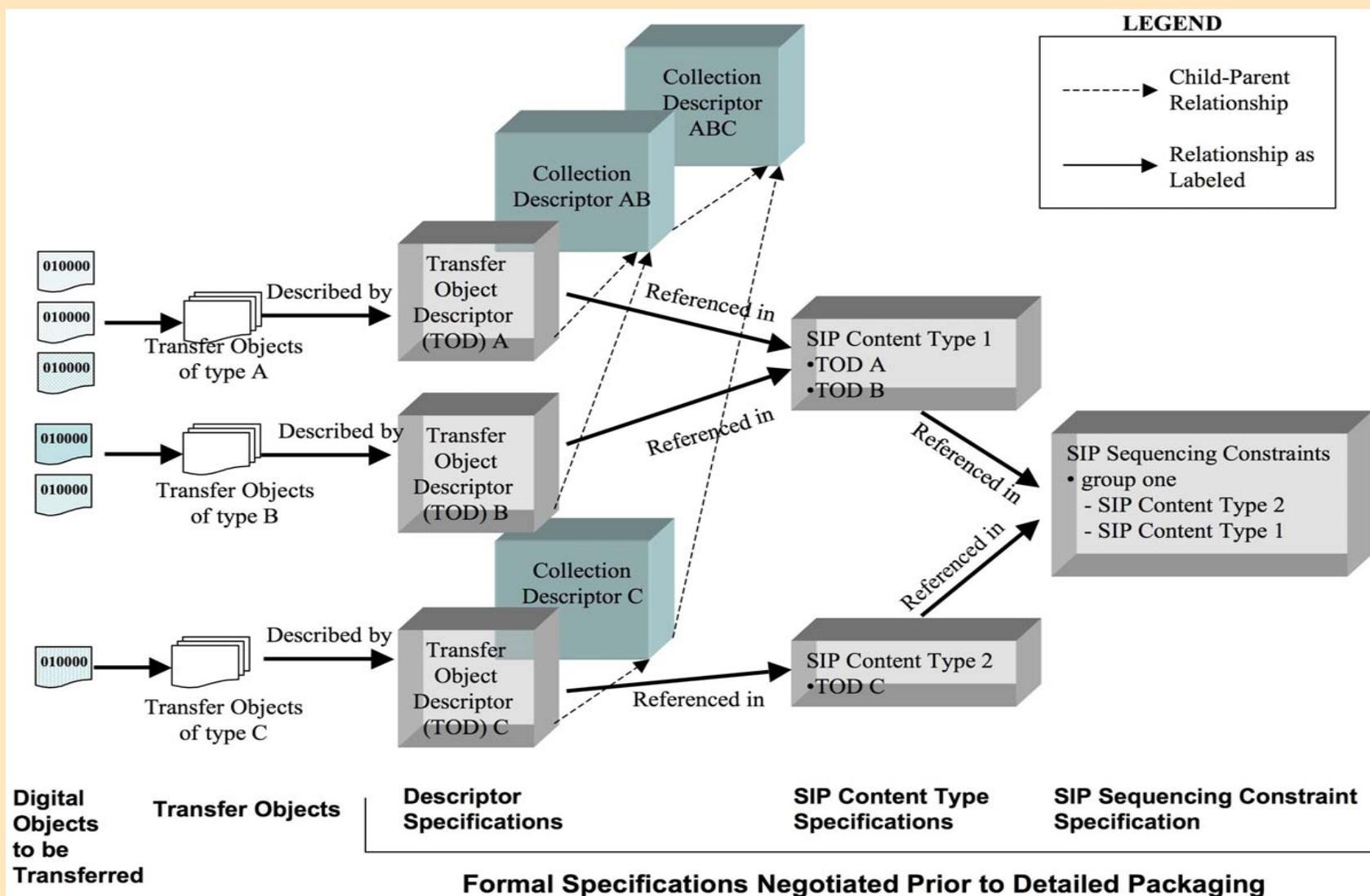
Producer-Archive Interface Standard: Motivation

- Many repositories expressed a desire to improve automation of the Producer-Archive interaction
 - Producer-Archive Interface Methodology Abstract Standard (PAIMAS) developed
 - Expanded OAIS Producer-Archive Interface
 - First step, but not directly implementable
 - Still need a standard implementation for at least part of the Methodology

Producer-Archive Interface Standard: Process

- Established CCSDS/ISO working group
 - Participants have been primarily Space Agencies
 - Make working documents available from Web
- Narrowed scope to address:
 - Modeling of information to be submitted
 - Modeling of how information will be packaged and exchanged
 - Provide a concrete information packaging technique
 - CCSDS/ISO XFDU standard consisting of ZIP container with standardized XML based manifest also meeting PAIS requirements

PAIS Formal Descriptions



Producer-Archive Interface

Standard: Status

- Expect outstanding-issue resolution at Berlin meeting, Oct. 13-17, 2008
- Draft standard expected to be distributed for formal CCSDS review by end of year
- Prototype software currently under development

Conclusion

- OAIS reference model established a major baseline for further digital preservation related standards development
- Many related standards have followed or are in process
- All are serving to improve trust in repositories that follow them
- Consensus on very highly trusted digital repositories providing long term preservation is likely years in the future