

Detailed Functional Model of the OAIS Reference Model: Local implementation of functions

Report to DISC, October 1, 2008

By Amy McCrory, Tschera Connell and Beth Black

This report begins with definitions of a few terms integral to the OAIS Reference Model, and proceeds with an analysis of how closely OSU Libraries meet the requirements of the model.

I. OAIS terms

OAIS—Open Archival Information System

A model showing how information is acquired, preserved, and disseminated. The OAIS DISC Subgroup report of October 1, 2008 evaluates OSUL's activities in the context of the OAIS model.

Producer

Provides the information to be preserved and disseminated.

Consumer

Persons or client systems that access the information in the repository. (Their action is not limited to simple consumption of resources; their behavior may help to shape OAIS actions such as decisions about the types of objects to collect, metadata creation, and method of delivery of content.)

Management

Sets OAIS policy.

SIP—Submission Information Package

An Information Package that is delivered by the Producer to the OAIS for use in the construction of one or more AIPs. Includes data (bitstream) and whatever metadata the Producer supplies.

AIP—Archival Information Package

Consists of data, metadata that accompanies it on submission, and preservation metadata, packaged as a single entity for transmission and storage in the OAIS. Parts of the AIP can be extracted for presentation, data manipulation, etc. A single object may have multiple AIPs, if the OAIS managers want to do that (for instance, to have one with confidential information for staff only, another with public information only).

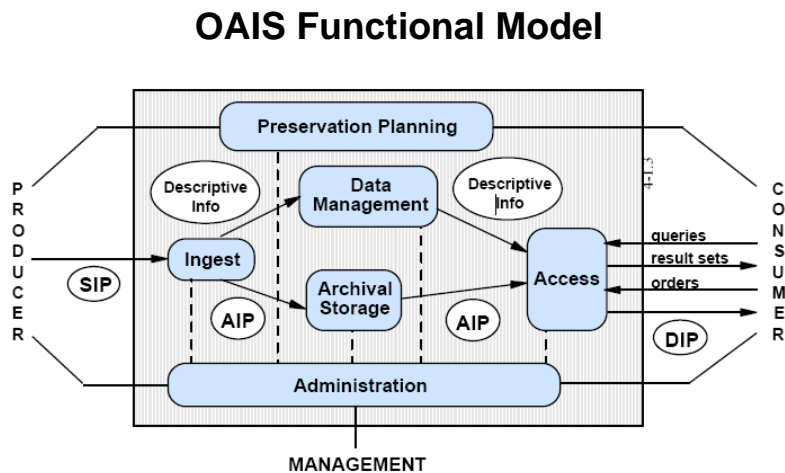
DIP—Dissemination Information Package

The Information Package, derived from one or more AIPs, received by the Consumer in response to a request to the OAIS. An archive works with [Consumers](#) over time to ensure that DIPs remain useful.

Preservation Metadata

May include Descriptive, Structural, Technical, and Rights metadata.

II. OSU Libraries compliance with the OAIS Functional Model



Note: Our digital production and preservation activities do not fall under one “umbrella.” This makes it difficult to put some of these activities into this model—for example, OhioLINK ebooks, ETDs, EJC, DRC, EAD. The following notes describe the Knowledge Bank only.

All images included in this report are taken from the CCSDS Blue Book “Reference Model for an Open Archival Information System.”

Common Services

*The model for Common Services is the **IEEE POSIX OSE Reference Model**.*

All Common Services are handled by OIT, in coordination with the Systems Librarian in the Libraries WIT Department.

Operating System services

- Architecture and software

Network services

- Protocols and APIs for data transfer

Security services

- Security is handled at the Kinnear Road Center, which recently received federal certification

Ingest

This entity provides the services and functions to accept Submission Information Packages (SIPs) from Producers (or from internal elements under Administration control) and prepare the contents for storage and management within the archive.

Many of these functions are automated in DSpace. Personnel, who provide human intervention, maintenance if the system fails, or assessment of system-generated data, are identified.

Receive Submission

- Transfer of data
- Legal transfer of custody
- Confirmation of receipt (by e-mail or DSpace notification)

WIT—Systems Engineer.

SRI—Head; Metadata Librarian; Coordinator for Input and Quality Control.

Quality Assurance

- DSpace automatically generates system log files

WIT—Systems Engineer

Generate AIP

- Dspace automates the transformation of the SIP into an AIP and sends the AIP to the Archival Storage function. This includes assigning a handle to the item, assigning an unique URL based on the handle to each bitstream, creating the preservation and technical metadata information (e.g. provenance, date/time submitted, checksum) and transforming the metadata into proper form for inclusion in the database.
- Dspace does not create a mechanical audit report & send it to Administration but does have logs of system activity.
- We do not currently perform reviews of the submissions as described in the Audit Submission portion of the Administration function so we do not have audit reports.
- DSpace prevents submission and sends requests for additional information to the submitter if required information is not provided during the submission process.

Required items include a title, a license, or a file; other items/metadata elements are possible at the community level.

WIT—Systems Engineer

Generate Descriptive Information

- DSpace automatically creates an extracted text index for searching, generates metadata from AIP, and generates thumbnails where appropriate.

WIT—Systems Engineer

Coordinate Updates

- DSpace automatically transfers AIPs to archival storage and descriptive information to the POSTGRES database.
- DSpace automatically stores the data that associates the bitstream with the descriptive metadata.

WIT—Systems Engineer

Archival Storage

This entity provides the services and functions for the storage, maintenance and retrieval of AIPs. Archival Storage functions include receiving AIPs from Ingest and adding them to permanent storage, managing the storage hierarchy, refreshing the media on which archive holdings are stored, performing routine and special error checking, providing disaster recovery capabilities, and providing AIPs to Access to fulfill orders.

Receive Data

- Receives transfer requests from Ingest
- DSpace makes entries in appropriate database tables
- DSpace puts files in appropriate directories, file structure
- DSpace does not state anticipated frequency of use
- DSpace creates handles; becomes part of the database key, unique ID
- DSpace sends a storage confirmation message

WIT—Systems Engineer

Manage Storage Hierarchy

- Most is covered in “Receive Data,” above
- DSpace automatically generates some operational statistic, such number of items/month, and item review. We wrote our own script that shows how much space we have left.

WIT—Systems Engineer

OIT—Systems Engineer

Replace/Refresh Media

- Evaluated and selected storage medium
- “Refreshment”: copying between media without reformatting the bitstream
- “Replication”: making copies of the media for redundancy
- “Repackaging”: Migration to a new medium under a new operating system and file system (Windows to Linux, for example?) See section 5 of model
- When moving to new hardware, OIT and Library IT staff have refreshed the media. As part of the back up requirements of the SLA with OIT, there are replicated copies of the bitstreams on tape back up. We have not performed repackaging yet at the time of this writing.

OIT—Systems Engineers. They inform **WIT**—Systems Engineer of actions

Error Checking

- Checking for corruption, fixity; we use DSpace checksum batch function, which generates reports of how many bitstreams have errors, and which ones. It runs each night on a group of bitstreams.

WIT—Systems Engineer

Disaster Recovery

- SLA with OIT includes offsite backups.
- We have used this service before, for other Virtual Machines, with satisfactory results.

OIT—Systems Engineers

Provide Data

- DSpace automatically responds to calls for data (for example, someone clicking on a handle in the KB).
- DSpace automatically writes a log of the transaction

WIT—Systems Engineer

Data Management

This entity provides the services and functions for populating, maintaining, and accessing both Descriptive Information which identifies and documents archive holdings and administrative data used to manage the archive.

Administer Database

- The database structure is determined by DSpace; structured data enables interaction with user access function.

WIT—Systems Engineer

Perform Queries

- DSpace executes queries.

WIT—Systems Engineer

Generate Report

- DSpace generates reports for usage statistics, lists of people with access to data and their various levels of authorizations.
- Staff can write custom queries to generate reports as needed.

WIT—Systems Engineer

Receive Database Updates

- Dspace adds, modifies, or deletes data in the database after actions by Ingest or Administration occur.

WIT—Systems Engineer

Administration

The entity that contains the services and functions needed to control the operation of the OAIS on a day-to-day basis.

Negotiate Submission Agreement

- Evaluates the quantity and characteristics of material that the community wishes to submit to the Knowledge Bank
- Advises and/or refers the community to *Copyright Help Center @ OSU* to address rights issues, if applicable
- Mocks up data model of community material on staging server, following community guidelines,
- Develops custom metadata standards for data model, if applicable
- Negotiates priorities, method (community submitted or Libraries submitted) and time table for submission including training for community submitters

KB Management Team (Administration)—KB Program Coordinator;
SRI—Head; **WIT**—Head)

Manage System Configuration

- Dspace creates logs of system performance
- Dspace or WIT creates reports as requested for inventory, performance, and operational statistics
- WIT and SRI use these reports for monitoring the operation of the system

WIT—Head, Systems Engineer

Archival Information Update

Updates include changes to bitstream (digital object) and / or metadata. Changes to metadata are community controlled, while ability to change bitstreams is restricted.

- KB Management Team provides policies for change requests to content.
- SRI and WIT establish procedures for implementing changes.

KB Management Team determines policies
SRI and **WIT** establish procedures and implement changes

Physical Access Control

- Determine and implement policies for top administrative rights to system. Implement procedures for temporarily restricting access to content, i.e., embargoes and moving walls.

KB Management Team determines policies.
SRI and **WIT** implement policies and procedures.

Establish Standards and Policies

- Provide updates and progress reports to the Digital Initiatives Steering Committee (DISC).
- Receive recommendations for archive system enhancement.
- Receive proposals for archive data standards, which are vetted through DISC.
- Develops standards and policies for operation of the KB.

KB Management Team

Audit Submission

- OSUL is not doing formal auditing at this time. SRI is working on formal procedures.

Activate Requests

- Dspace generates responses to dissemination requests (searchers, handles, and browses).

WIT—Systems Engineer

Customer Service

- KB Help listserv is monitored by OSUL staff, including Program Coordinator, members of SRI and WIT departments.
- SRI creates customized documentation for communities.
- WIT responds to trouble calls, pages, e-mails.

Administration—Program Coordinator
SRI—Head, Metadata Librarians
WIT—Head, Systems Engineer

Preservation Planning

This entity provides the services and functions for monitoring the environment of the OAIS and providing recommendations to ensure that the information stored in the OAIS remains accessible to the Designated User Community over the long term, even if the original computing environment becomes obsolete. (Preservation Planning was not a component of the original OAIS recommendation produced by CCSDS.)

Additional comments: Preservation Planning includes not only designing and adopting a preservation plan, but monitoring the preservation environment over time.

Monitor Designated Community

Includes monitoring the producers and consumers to track changes in requirements and mechanisms.

- KB Project Manager has a follow-up meeting once a year with each community.
- KB Management Team monitors active projects through Project Load.

Administration—KB Project Manager
KB Management Team

Monitor Technology

This means keeping aware of new technologies that can improve the preservation environment, and avoid the potential for obsolescence and incompatibility.

- Not currently a formal OSUL activity. Staff in various departments tend to monitor technological developments as the need arises relevant to their own work.

Develop Preservation Strategies and Standards

Choose preservation standards and formats and write guidelines and/or processes for implementation. Continue to monitor developments in standards, and determine when local practices need updating.

- Currently, this takes the form of periodic reports and updates as requested by DISC, followed by documents being added to Digital Projects page. Not a formal activity of any particular department, staff member, or cross-departmental group other than DISC.

Develop Packaging Designs and Migration Plans

- Developing plans for implementing recommended digital preservation strategies and policies.
- WIT has migrated to different hardware platforms twice in the last year, and has written plans in place. These could be used as a model for higher-level planning.
- PDF/A subgroup has outlined options for migrating from PDF to PDF/A. This activity can be extended to other problems.

Access

This entity provides the services and functions that support Consumers in determining the existence, description, location and availability of information stored in the OAIS, and allowing Consumers to request and receive information products.

Coordinate Access Activities

- DSpace software provides a user interface for customers to submit query requests, browse requests or orders for specific items in the repository.
- DSpace software provides DIP via web user interface and download mechanism for the bitstreams included in the items.

- The Libraries provide feedback web forms through which customers may send assistance requests.

WIT – Head; Systems Engineer

SRI – Head, Metadata Librarians

Generate DIP

- DSpace software retrieves items from the database and file structure in response to requests.

WIT – Systems Engineer

Deliver Response

- DSpace software handles on-line delivery responses via the web user interface. This includes identification of access to determine access rights and transmission of the requested materials via web protocols.
- Off-line requests would be directed by KB personnel to the appropriate community.

WIT – Head, Systems Engineer

SRI – Head, Metadata Librarians

Administration – Program Coordinator (for off-line requests)