

ANNEX B. FEDERATION OF ARCHIVES

Users of multiple OAIS archives may have reasons to wish for some uniformity or cooperation among them. For example, Consumers of several OAIS archives may wish to have

- a master catalog to aid in locating information over several OAIS archives,
- a common schema for access and dissemination, or
- a single access site.

Producers may wish to have

- a common schema for submission to different archives, or
- a single depository for all their products.

Managers may wish to have means for

- enforcing submissions, or
- increasing the uniformity and quality of user interactions with the OAIS.

Therefore, it may be advantageous for OAIS archives to cooperate to meet these wishes. The motivation might come from the archives themselves, or it may be imposed by an authority that has some influence over them. In the former case, the archive might be motivated by the desire to keep consumers happy with their products, or to keep users happy with their quality of service, or simply by the need to compete with other archives in order to survive or grow. Situations like this can and have motivated agreements without the need for any explicit federation establishing an external authority.

In cases where explicit federation is established, the external authority is represented in this Reference Model by Management. It is not the purpose of this section to discuss the detailed organizational architecture (or intrigue) of the Management interactions, since these are outside the model, but rather to describe the OAIS external interconnections that might allow the above wishes to be met.

At a rudimentary level of federation, **Figure B-1** represents a simple mutual information exchange agreement between archives. (Note: In this and the following figures, the OAIS is represented as a “five-port device” following the arrangement of Figure 4-1. In each case, a two-archive federation is shown for simplicity, although the concept can be extended indefinitely.) The essential requirement for this federation is a set of mutual Submission Agreements, subscriptions, and user interface standards to allow DIPs from one archive to be ingested as SIPs by another. Therefore, it assumes that some pair-wise compatibility has been established between the archives. This does not necessarily require common access, dissemination and submission methods for all participants, although that may be expected encourage more exchange.

///I view this interaction as each archive having a submission and dissemination agreement with the other archives. In general the dissemination requests would be subscriptions so no one had to be aware of activity at other archives. If many archives are involved in this relationship, there would be pressure for standards so an archive did not have to produce DIPs in several different formats///

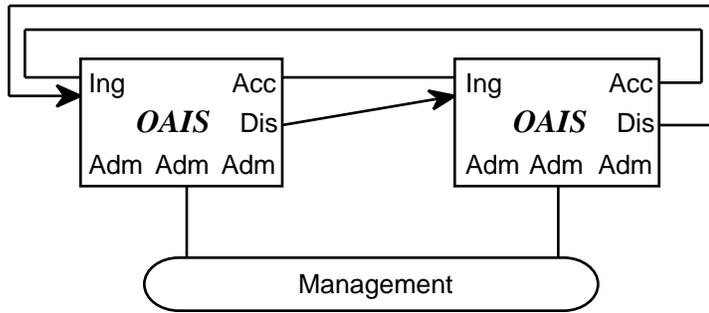


Figure B-1. A Simple OAIS Federation for Mutual Data Exchange

Figure B-2 is an example of OAIS archives that have standardized their submission and dissemination methods for the benefit of users. No special external element, other than management, is needed for this. Its disadvantage is that there is no formal mechanism for exchange of catalog information. Where does the consumer look for the desired information? Perhaps the archives agree to exchange context and catalog information, or perhaps one of the archives agrees to take on the role for both.

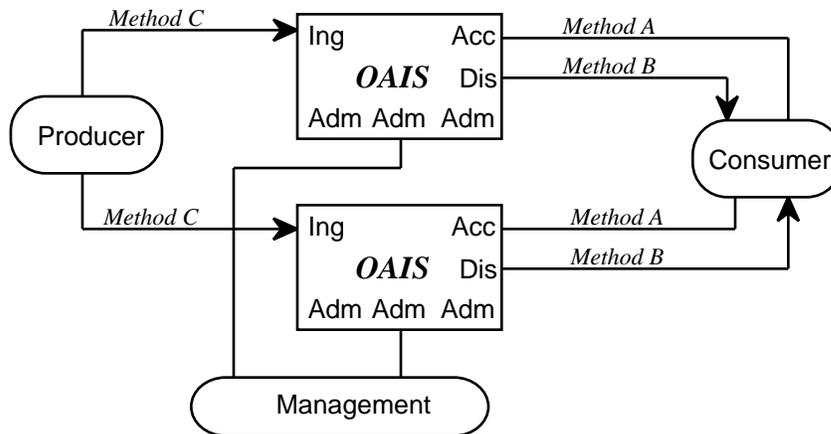


Figure B-2 An OAIS Federation with Standard Ingest and Dissemination Methods
///I don't think of this as a form of federation but rather as an interarchive standardization based on external standards///

Figure B-3 shows a way to solve the catalog problem using an entity eternal to the OAIS. Here, a pair of Producer-Consumers, each of whom maintains his own OAIS archive, have joined together to share information. The Common Catalog & Manager is the external binding element that serves as a common access point for the information in both archives. The Common Catalog may limit its activity to being a finding aid, or it may provide full combined Access service as shown in the Figure. Optionally, it may include common Dissemination of products from either or both archives.

///I think that this architecture is the basic concept of federated archives and needs to be expanded based on the finding aids available to the consumers///

This architecture closely matches that of NASA's TIMED (Thermosphere, Mesosphere, Ionosphere Energetics and Dynamics) space mission, in which four individual archives are operated by Principal Investigators. As Producer, each principal Investigator is responsible for archiving his own products. However, there are instances where the Producer incorporates the products of other Principal investigators into his product. In this application, the mission operator is the external entity. The Common Catalog component provides a Dissemination service which can assemble special products obtained from multiple archives, and also accept subscriptions for them. Additionally, the Manager component exploits its position to enforce the submission of data products by each Principal Investigator to his own archive, according to the agreement established for the mission.

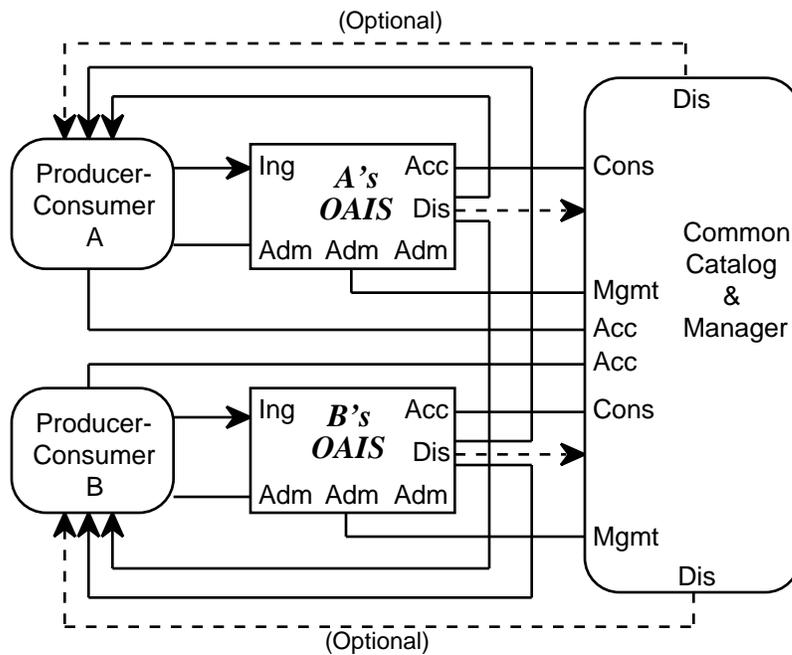


Figure B-3. An OAIS Federation Employing a Common Catalog

Figure B-4 shows a federation with an external entity on the Ingest side. As with the example of the Common Catalog, this entity permits enforcement of broad submission agreements involving several producers and archives. Here, the Common Ingest Staging entity can also take the responsibility to route submissions to the appropriate archive.

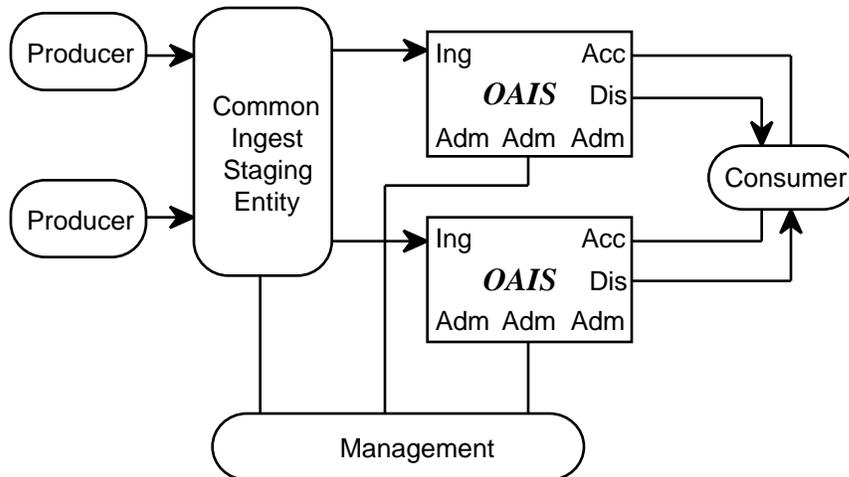


Figure B-4. An OAIS Federation with Common Ingest Staging

Of course, the arrangement of Figure B-4 is little help for the consumer who does not know where to look for information. A more elaborate arrangement than any of those shown above would include both a Common Catalog and Common Ingest Staging, to provide an external appearance similar to that of a single OAIS. Such a complete embedding of OAIS archives may be thought to be less efficient than a single distributed archive under one administration. However, it does allow for a negotiated degree of autonomy for each archive.

/// I think that a combination of b3 and b4 is a dividing line between a federation of archives and a distributed archive///

It should be evident from the above examples that the OAIS model is consistent with federation to accomplish specific objectives. However, it should also be considered that some of these objectives may be accomplished through voluntary action.

[Contributor's Note: In discussing this, we may discover that implementation of federations may be made easier by adding some internal functions (perhaps in administration and access) to support them. We may even discover that some previously unrecognized means of inter-OAIS exchange among entities, such as Access-to-Access exchange, should be explicitly recognized in the model -PG]

///Unfortunately we need a separate type of drawing to address Claude Huc's case of several independent archives sharing a storage management system///

An Analysis of levels of Interoperability

Lou Reich 1/27/98

A way of visualizing federated information systems is by categorizing their producer and consumer communities. Another way of visualizing federation of distributed systems is by investigating the level of integration of the functional components as presented in Section 4.2.1 of the participating archives and common implementations of the information models presented in Section 4.2.2 of the Reference models.

The following table defines four categories of archives:

- Independent - Local community only, no management interaction or interest in Standards at other archives
- Cooperating - Local community with potential Global producers, no common access, joint submission and dissemination standards, Subscription requests for key data at the other archive
- Federated - Local and Global communities, has both Local and Global access, the Local community has priority over the Global community. Optionally Global Dissemination and Ingest can be added.
- Distributed - Global Community only, single management

Each row of the table represents a different analysis case. All cases are not considered due to time and space constraints (table would be way too large) and intuition on their importance. They have not been included explicitly but are included in discussions:

Name	Ing	Stor	DM	Admin	Access	Dis	AIP	SIP	DIP	Descrip	MAN
IND	L	L	L	L	L	L	L	L	L	L	L
coop1	L	L	L	L	L	L	L	L+1G	L+1G	L	L
coop2	L+G?	L	L	L	L	L+1G	L	G	L+1G	L	L+G
Fed1	L	L	L	L	L	L	L	L	L	L+1G	L
Fed2	L	L	L+G	L+G	L+1G	L	L	L	L+G	L+G	L
Fed3	L	L	L+G	L	L+G	L+G	L	L	L+G	L+G	L+G
Fed4	L+G	L	L+G	L	L+G	L+G	L	G	L+G	L+G	L+G
Fed5	?	G	?	L+G	?	?	?	?	?	?	L+G
Dist1	G	L	G	L+G	G	G	G	G	G	G	G
Dist2	G	G	G	L+G	G	G	G	G	G	G	G

Table 1 -Analysis of Archive Interoperability

Comment on Levels of Interoperability described in Table 1:

- Independent - Fairly obvious, though it may have common AIP with other archives it does not assume to have any knowledge about other Archives
- Cooperating 1 - The only requirement for this architecture is that an archive support at least one common SIP and DIP format for interarchive requests. The control mechanism for this sort of interoperation can be subscription requests at each archive

- Cooperating 2 - Common ingest and dissemination for global users have been added based on common SIPs and DIPs similar to those defined by interarchive actions described above
- Federated 1 - Global access is accomplished by the export of a standard format associated description to a global site which independently manages a set of descriptors from many archive and has finding aids to locate which archive is the owner of collections of interest. An example of this architecture is the Master Directory/GCMD work . It is a close call whether this is a federated or a cooperating archive case.
- Federated 2 - Global access is accomplished by having a global node which can distribute a query to multiple local archives. This means the local Data Management must store an additional associated description in the global format or have a translator from the global queries to local queries. An option in this case is to establish a common DIP format to ease the load on consumers who may order products from many archive. An example of this is early versions of IMS V0 or CEOS IMS.
- Federated 3 - In addition to the functionality of Federated 2 a standard ordering and dissemination mechanism is available through the global node. This is a fully functional, modern federated system. The global system will have some effect on the schema designs in local archive. In this situation it is optimal to build new local archives based on the global schemas and finding aids to ensure high degrees of interoperability. An example of this level of interoperability ECS and ICS in the earth observation domain
- Federation 4 - In addition to the functionality of Federated 3 a common ingest facility is defined. This type of functionality may be present in the NASA EOSDIS but a common ingest facility tends to push towards a fully distributed system.
- Federation 5 - This level of interoperable system is a common storage management function among the participating archives. The access, dissemination and ingest facilities can be at any of the previous described levels of interoperability. The CNES SPOT/IMAGE archive is an example of this level of Federation
- Distributed 1 - This level of interoperable system has integrated functional areas of Ingest, Data Management, Access and Distribution. The only assumption is the each site has its own local storage and some local administration. The distribution of functional areas among sites may be very specialized(i.e., there may be an ingest site, a storage site and a separate access site). The difference between this model and Federated 3 or 4 is there is no local community with different requirements from the global requirements.
- Distributed 2 - This level of interoperable system adds common storage to the features of Federated 1. At this point the architecture is similar to the Independent Archive except the community is dispersed and distributed technologies are used.

*SUBSCRIBER @EditionMgr @EditionClient @0449E884 \a SUBSCRIBER
@EditionMgr @EditionClient @0449F59C \a*