# ARCHIVE INTEROPERABILITY

An OAIS may be geographically distributed but with all parts under the same Management, for example the Archival Storage Functional Entity could be divided over several separate locations to increase resilience against disaster. In other cases OAIS Archives with separate Managements may wish to co-operate as described below.

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

* common Finding Aids to aid in locating information across several Archives;
* a common Package Description schema for access;
* a common DIP schema for dissemination; or
* a single global access site.

Producers may wish to have:

* a common SIP schema for submission to different Archives; or
* a single depository for all their products.

Managers may wish to have means for

* cost reduction through sharing of expensive hardware, software, and preservation efforts; and
* increasing the uniformity and quality of interactions with several Archives.

Therefore, it may be advantageous for Archives to cooperate to meet these wishes. The motivation might come from the Archives themselves, or an authority that has some influence over them may impose it. In the former case, the Archive might be motivated by:

* desire to reduce costs;
* desire to keep Consumers happy with their products;
* desire to keep users happy with their quality of service; or
* 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. However, in cases where explicit federation is established, the external authority is represented in this Reference Model by Management.

The purpose of this section is to explore the degree of interaction and cooperation among Archives. Subsection 6.1 focuses on the technical levels of interaction, while subsection 6.2 discusses management issues concerning the tension between cooperation and autonomy.

## Technical Levels of Interaction between OAIS Archives

In general one OAIS is not interoperable with another; however, there are a number of reasons that some level of interoperability may be desirable, motivated for example by Users, Producers or Management; even interoperable Archives may have different Designated Communities—even for the same digital objects—and hence different requirements for Representation Information and/or Descriptive Information.

OAIS associations can be categorized technically by both external and internal factors. External factors include characteristics of the Producer and Consumer communities. Internal factors could include common implementations of the information models presented in 4.2, or multi-Archive sharing of one or more of the functional entities presented in 4.1.

This subsection describes four categories (not an exhaustive list) of Archive associations. The first three categories have successively higher degrees of interaction:

* *Independent*: Archives motivated by local concerns with no management or technical interaction among them.
* *Cooperating*: Archives with potential common Producers, common submission standards, and common dissemination standards, but no common Finding Aids.
* *Federated*: Archives with both a Local Community (i.e., the original Designated Community served by the Archive) and a Global Community (i.e., an extended Designated Community) which has interests in the holdings of several OAIS Archives and has influenced those Archives to provide access to their holdings via one or more common Finding Aids. The Access needs of the Local Community usually have priority over those of the Global Community. Global dissemination and Ingest are optional features.
* *Shared resources*: Archives that have entered into agreements with other Archives to share resources, perhaps to reduce cost. This requires various standards internal to the Archive (such as ingest-storage and access-storage interface standards) but does not alter the user community’s view of the Archive.
* *Distributed Digital Preservation:* Archives that have entered into agreements with other Archives to manage functions as part of conducting distributed digital preservation (the use of replication, independence, and coordination to address the known threats to digital content through time to ensure their accessibility resources) in a way that does not alter the user community’s view of the Archive.

The remainder of this subsection gives a more detailed view of these categories of association.

### Independent Archives

An independent Archive is assumed to serve only a single Designated Community. The Archive and the Designated Community must agree on the design of DIPs and Finding Aids. An independent Archive may choose to design these structures based on formal or de-facto standards, which would allow cooperation with other Archives that implement the same standards. However, the design decisions to use these standards are not based on the possibility of inter-operation with other Archives, but rather on local requirements and cost savings.

The classification of an Archive as independent is not based on its size or distributed functionality. An independent Archive may occupy one site, or may be physically distributed over many sites. It may use many standards for a given internal element. However, if there is no interaction with other Archives, the Archive is independent.

### Cooperating Archives

Cooperating Archives are based on standards agreements among two or more Archives. The simplest form of cooperation between Archives is when one Archive acts as a Consumer of material from another Archive. In this case the consuming Archive must support the DIP format of the producing Archive as a SIP format. Cooperating Archives have related communities of interest, so they order and ingest data from other cooperating Archives and possibly have common data Producers. No common access, submission or dissemination standards are assumed. The only requirement for this architecture is that the cooperating groups support at least one common SIP and DIP format for inter-Archive requests. The control mechanism for this sort of inter-operation can be Event Based Order requests at each Archive.

Figures 6‑1 and 6‑2 illustrate the concept of cooperating Archives.

At a rudimentary level of Archive interaction, figure 6‑1 represents a simple mutual information exchange agreement between Archives.

NOTE – In this and the following figures, the OAIS is represented as a multi-port device following the arrangement of figure 6‑1. 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, Event Based Orders, 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 between established between the Archives. This does not necessarily require common access, dissemination and submission methods for all participants, although that might encourage more exchange. This level of agreement would also be useful when the holdings of one Archive were consolidated/transferred into another Archive because of Management issues.

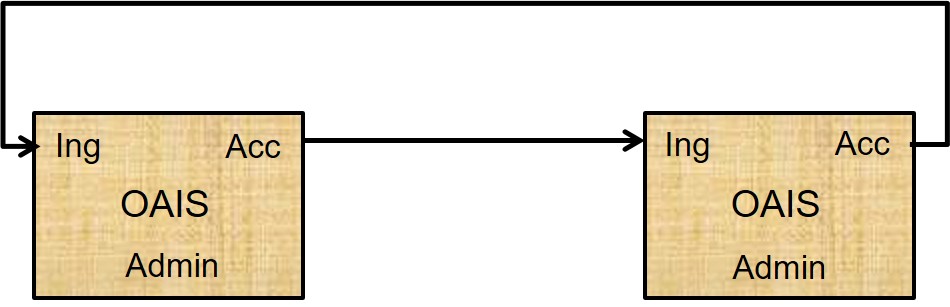


Figure 6‑1 : Cooperating Archives with Mutual Exchange Agreement

Figure 6‑2 is an example of OAIS Archives that have standardized their submission and dissemination methods for the benefit of users. No special external element is needed for this. Its disadvantage is that there is no formal mechanism for exchange of Description Information so Consumers must have separate Search Sessions to locate AIPs of interest.

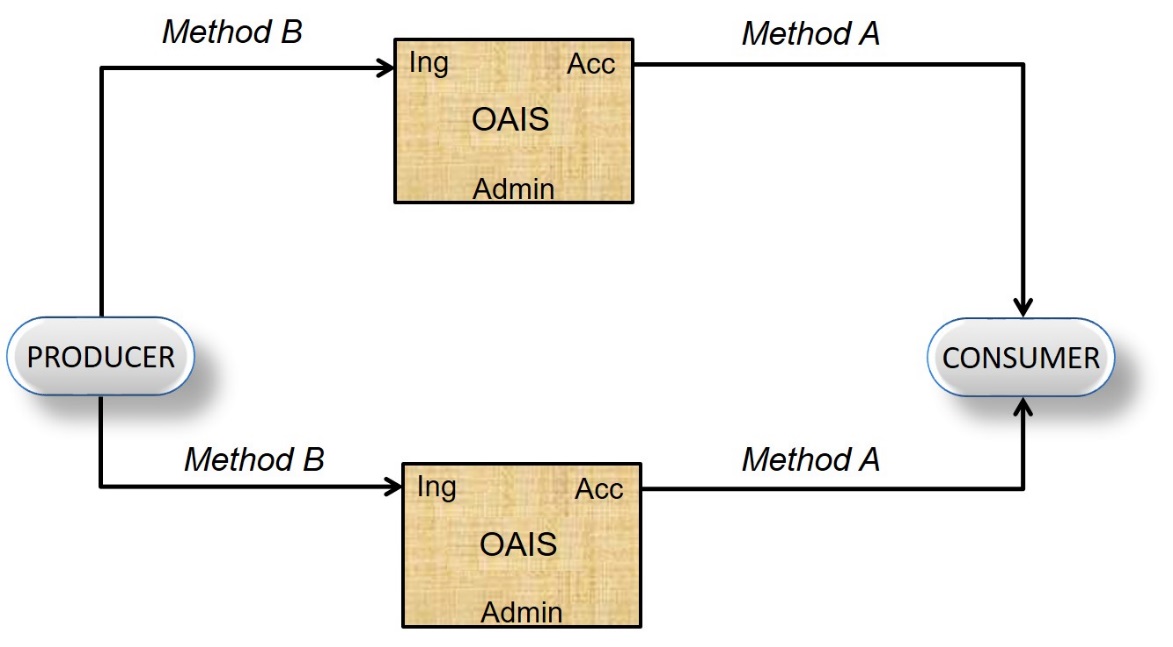


Figure 6‑2 : Cooperating Archives with Standard Ingest and Access Methods

### Federated Archives

Federated Archives are conceptually Consumer-oriented. In addition to the Local Community (i.e., the original Designated Community served by the Archive), there exists a Global Community (i.e., an extended Designated Community) which has interests in the holdings of several OAIS Archives and has influenced those Archives to provide access to their holdings via one or more common Finding Aids. However, the Local Consumers are likely to have access priority over the global Consumers.

At the federated level of association, external elements can be introduced to improve inter-operability. For example, figure 6‑3 illustrates a functional architecture to solve the Access problem described in 6.1.2, using an entity external to the Federated OAISes. Here, two OAIS Archives that have similar Designated Communities have decided to federate to allow Consumers to locate Information Packages of interest from either OAIS with a single Search Session. The Common Catalog & Manager is the external (global) binding element that serves as a common access point for the information in both Archives. DIPs containing the Finding Aids from each OAIS are ingested into the Common Catalog as is shown by the dotted lines in figure 6‑3. The Common Catalog may limit its activity to being a Finding Aid or it may include common dissemination of products from either or both OAISes as shown by the dashed lines in the figure.

NOTE – Extra access ports are added to the diagram to illustrate the potentially differing views of each consumer community.

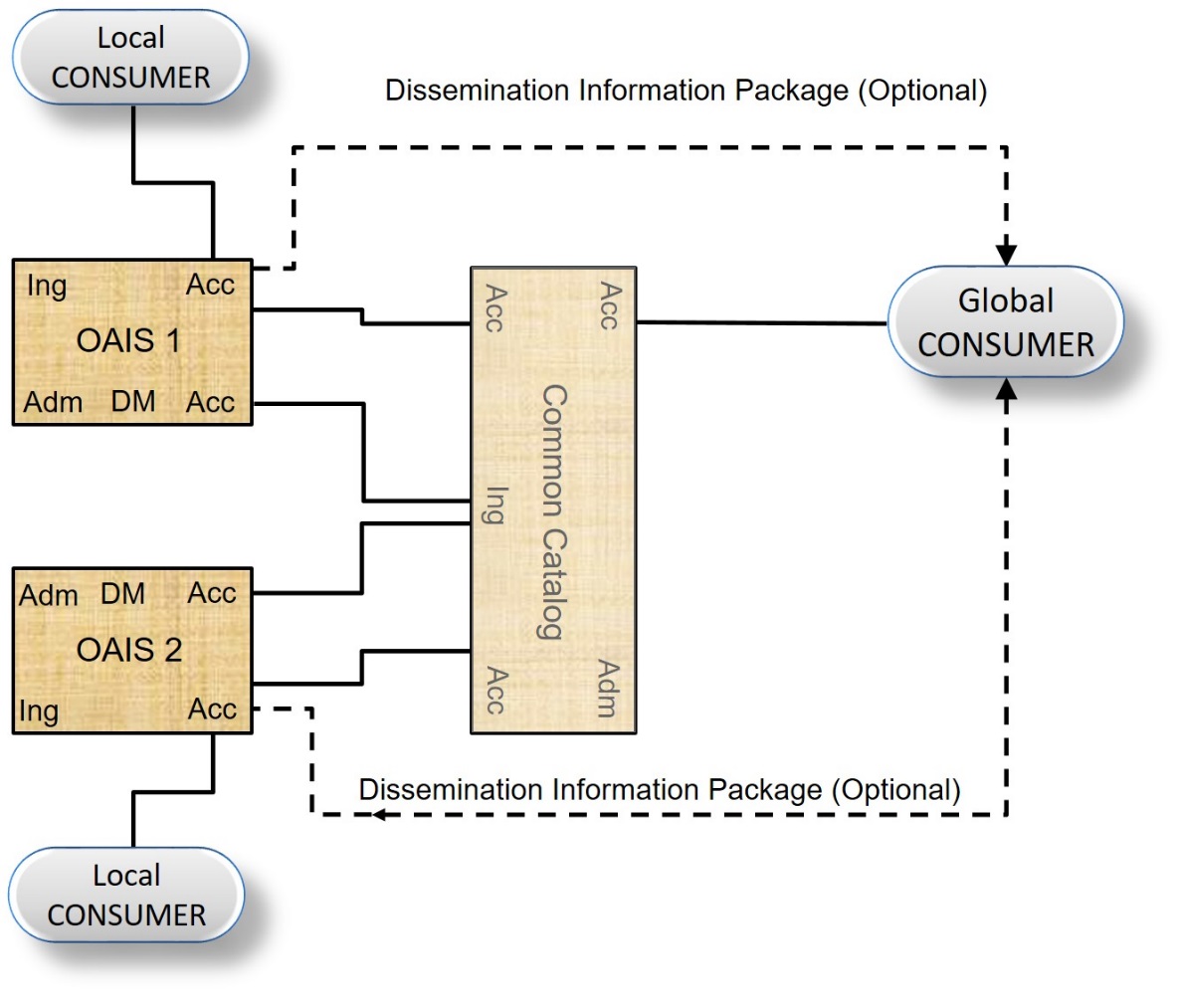


Figure 6‑3 : An OAIS Federation Employing a Common Catalog

Federated Archives may be further classified into three levels of functionality.

* *Central Site*: Global access is accomplished by the export of a standard-format Associated Description to a global site. The global site independently manages a set of descriptors from many Archives and has Finding Aids to locate which Archive owns a collection of interest. The Consumer is given a combined view of the holdings of multiple sites, which is maintained centrally. To view details of the documents, the user must access the site that contains the actual document. This is made easier when sites and clients support a standard set of protocols.
* *Distributed Finding Aid*: Global access is accomplished by having a global node that can distribute a query to multiple local Archives. This means that the local Data Management entity 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 be ordering products from many Archives.
* *Distributed Access Aid*: Global access is accomplished through addition of a standard ordering and dissemination mechanism, available through the global nodes to the functionality of the Distributed Finding Aids discussed above. This is a fully functional, federated system. Here, the global system may influence the Associated Description schema designs in each local Archive; it would be optimal to build new local Archives based on the global schemas and Finding Aids to ensure high degrees of inter-operability.

There are several major policy/technology issues that must be addressed when an OAIS joins a Federation or several independent OAISes decide to create a Federation. These issues include:

* **Unique AIP Identifiers** for each AIP in the Federation; it would be advantageous for these identifiers to be Unique Persistent Identifiers within some specified naming system. An OAIS has the responsibility to provide each of its AIPs with a unique identity. When an OAIS joins a Federation, there is no assurance that some of its current OAIS AIP identifiers are not already used by other members of the Federation. An example of a general solution to this problem is to form the AIP identifiers in the Federation by assigning a Unique ID for each OAIS in the Federation and concatenating it to each AIP preserved by that OAIS. This OAIS name could be formatted according to a standard that gives the Consumer or other Federation members the information needed to establish a connection to the OAIS that contains the AIP of interest. An example of a standard that accomplishes this is the ISO X.500 Directory Services Naming.
* **Duplicate AIPs** in several different OAISes with different AIP names. This problem is caused by the fact that prior to Federation, some OAISeswill have duplicated Content Information from AIPs in other OAISes to enable local user access. In this case a Global Consumer will see all the copies as separate, uniquely identified AIPs. Detailed examination of the PDI associated with the Content Information should allow the Consumer to locate the original, authoritative copy, but the search process could be very frustrating to the user. A practical way to handle this is to have a field in the Associated Description of all AIPs that identifies whether they are the original or a copy. This technique is not effective if, prior to federation, two or more Archives received the content information from the Producer to Archive. In this case the federated Archive would view these duplicate AIPs as unique, original AIPs.
* **The Preservation of Federation Access to AIPs** when an OAIS terminates operations. Unfortunately, many Archives will close while their holdings still are of value to the Federation community. The federation should have an agreement for each member OAIS, which states which OAIS has the responsibility to take over the preservation of a closed OAIS’s holdings.
* **User Authentication and Access Management** for global users. If an OAIS has a policy that restricts the access to some of its AIPs or charges for the dissemination of some Information Packages, there is a problem of how to identify and authenticate the user who is making requests through the central node. Each OAIS will have implemented an Authentication and Access Management system for its Local Users and the infrastructure for this function will have to be extended to include Global Users. Some examples of techniques used for this in current systems are:
* Default priorities where all members from the Global Node share a common set of access constraints and the Global Node handles all the authentication to verify the Consumer as a legitimate user of the Global node. The authentication at member OAIS is done assuming that all requests from the global node are from a single user.
* User Credential passing where the specific remote user can be authenticated by any of the Federation OAISes and the global node simply acts as an intermediary to carry the authentication dialog.

There are many factors influencing the decision of which of these techniques should be used by a specific Federation. The major criterion is the granularity of the Access Constraints in the Federation. If there is little private data and no charge for data dissemination, a policy that determines a user’s access constraints by the source he uses to discover and order AIPs is very reasonable. This involves little modification to the OAIS Authentication system, simply adding the Global Node as a Consumer. The Global Node will have to include mechanisms to identify Global Users to each of the Federated OAIS Authentication mechanisms.

If there are charges for disseminating preserved information or significant private data that needs individual user authentication, the proxy techniques will not be sufficient, and User Credential passing techniques such as passwords and Certificates must be applied. The technologies to enable these supporting mechanisms are still evolving.

### Archives with Shared Functional Entities

In an association involving Archives with shared functional entities, Management has entered into agreements with Archives to share or integrate functional entities. The motive for this may be to share expensive resources such as hierarchical file management system for Archival Storage, peripheral device for Ingest or dissemination of Information Packages or supercomputers for complicated transformations between SIPs, AIPs or DIPs. This association is fundamentally different from the previous examples, in that it is no longer possible to ignore the internal architecture of the Archive.

Figure 6‑4 illustrates the sharing of a common storage function, consisting of an Archival Storage entity and a Data Management entity, between two Archives, OAIS 1 and OAIS 2. The Access and Ingest facilities can be at any of the previously described levels of inter-operability. In fact, each Archive can serve totally independent communities as implied in this figure. However, for the common storage element to succeed, standards are needed at the internal Ingest-storage and Access-storage interfaces.

Additional potential shared services include registries of Representation Information and name resolvers such as the DNS. In the former case a registry of Representation Information should also be an OAIS and the Representation Information it holds should be part of the Content Information it holds. The Representation Information it holds might, for example, be part of the Representation Information Network for the Content Information within an AIP in another OAIS. In such a case the OAIS holding that AIP may cache copies of the Representation Information Network held in the registry. Whether it does so or instead relies on the registry to maintain the Representation Information Network, the ultimate responsibility for the understandability of the Content Information remains with the OAIS which holds the AIP.

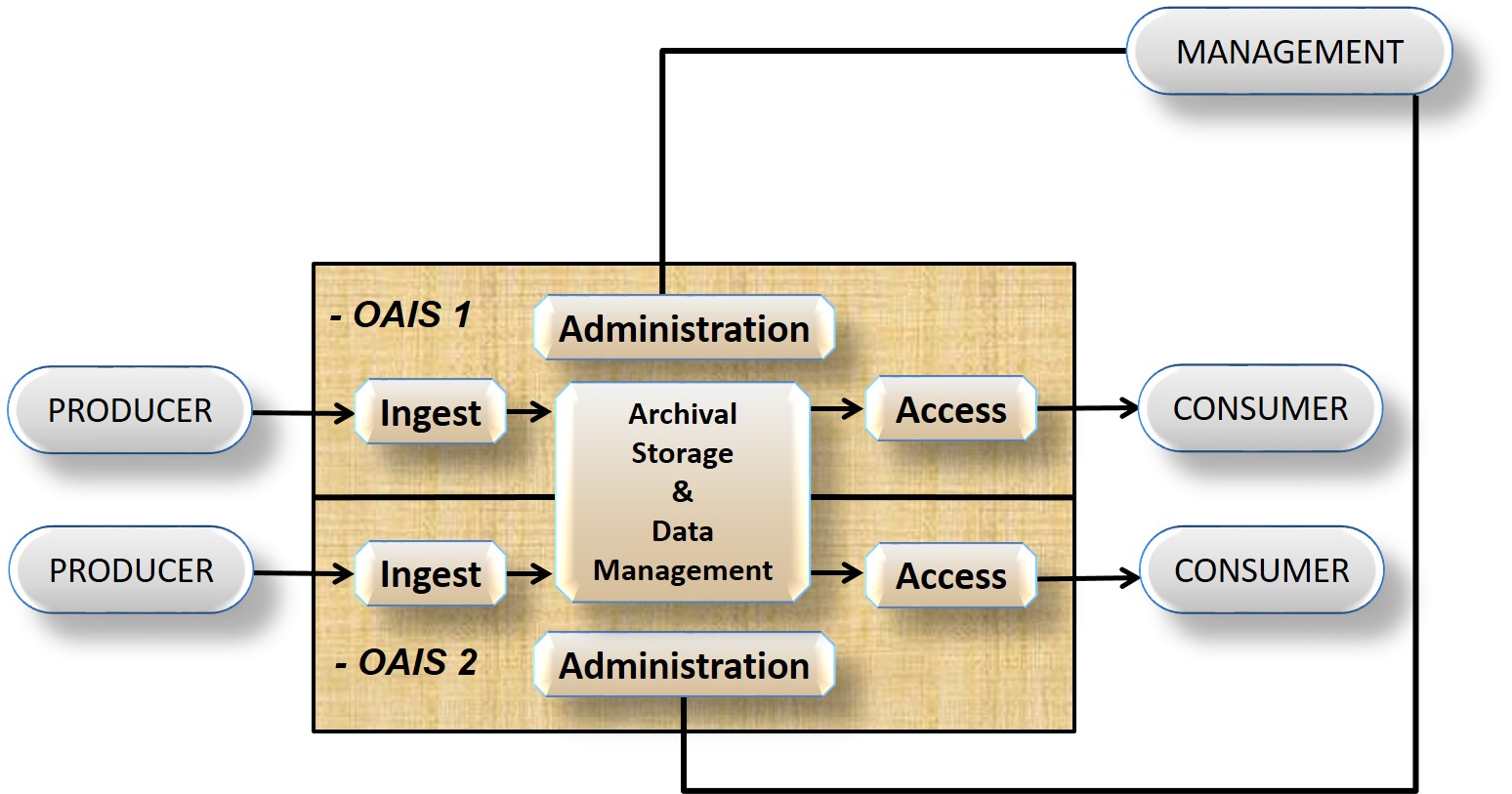


Figure 6‑4 : Archives with Shared Storage

### Distributed Archives with Distributed Functional Entities

A distributed OAIS with distributed functional entities can appear in the form of, e.g., an OAIS that has entered into agreements with other OAISes or organisations to link or integrate their distributed functionalities or/and services as distributed functional entities or distributed functional areas in a distributed complementary and collaboratively way.

Such a distribution of functional entities or functional areas of an OAIS, as well as their collaborative composition into distributed OAIS, can be of a physical, organisational, and/or administrative nature. The motivation for such a physical, organisational, or administrative distribution of functional eEntities and functional areas of distributed OAIS may be to distribute resources to achieve the complete set of functional entities and functional areas to establish a distributed OAIS in a complementary and collaborative way. Other motivation factors can be to mitigate risks for loss of data or outsource parts of the OAIS in a way that enables the OAIS to remain an OAIS.

Distribution can be achieved on different levels, either by having an OAIS that depends on a number of contributing OAISes, or by having an OAIS that depends on distribution of functional entities and functional areas. Combination of these is also possible. Here the distribution can be in terms of the physical location, organisation, or logical functions

This type of association of functional nntities and functional areas of an OAIS is different from the examples in previous sections, in that it does not only federate, share or cooperate w.r.t. functional entities and functional areas but physically, logically and/or organizationally distributes them in accordance with competences, capacities and characteristics of contributing archives, e.g. locations, organizations or administrations. The involved functional entities and functional areas of each of the participating OAISes can be at any of the previously described levels of interoperability.

The denotations ‘outer OAIS’ and ‘inner OAIS’ will be used in the following examples. These denotations are used for precision of a distributed (outer) OAIS in focus and its dependencies to distributed (inner) components in focus, i.e. functional entities, functional areas or other OAISes. In general, an outer OAIS depends on the existence of its inner OAIS, inner functional entities and inner functional areas to become a complete OAIS. The denotation ‘inner’ is only used to specify inner parts of an outer OAIS. The difference between an inner OAIS and an inner functional entity/inner functional area is that the inner functional entity/inner functional area does not constitute a full OAIS.

Figure 6-5 is illustrating an example where an outer OAIS has three other inners OAISes as part of physically, organizationally, and administratively distributed Archival Storage.

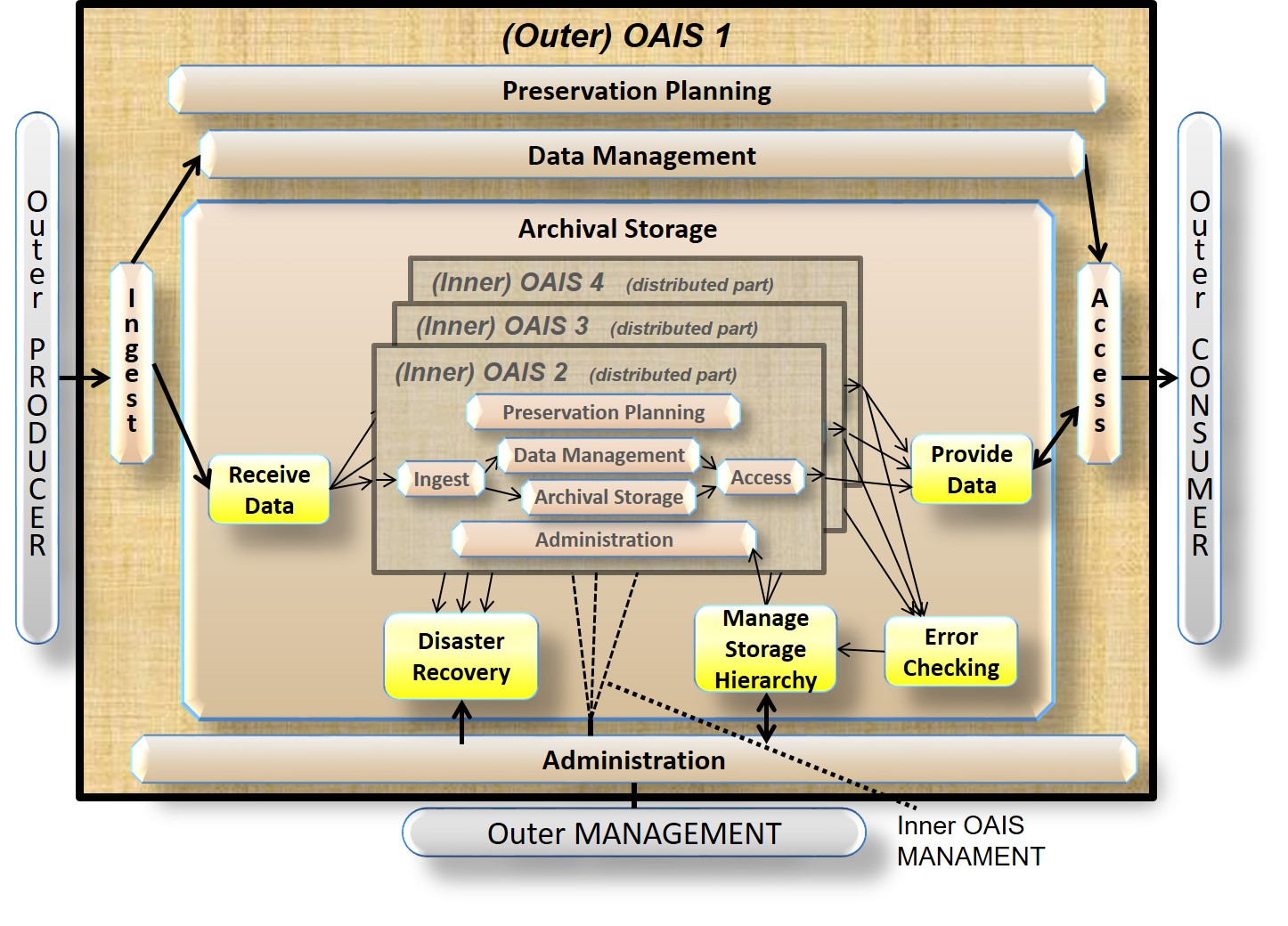


Figure 6-5: Example of a Distributed OAIS Archive for Archival Storage

Inthe examplefrom Figure 6-5, the inner OAISes contain different parts of the functions that constitute parts of the Archival Storage for the outer OAIS. The Designated Community, the Pr and the Consumer for the inner OAISes are in this example all represented by the outer OAIS. The Preservation Objectives for the inner OAISes are only to keep track on the bit sequences being intact and accessible, as well as handle Provenance Information. In this example, each of the inner OAISes need to have a minimum of Preservation Planning in form of Monitor Technology of the media used for storage, media migration planning etc. Data Management and Administration handle Provenance Information and Reference Information. In this example, all Preservation Planning concerning media migration and functions related to Replace Media are distributed to the inner OAISs as well as parts of the other Archival Storage functions needed for a simple bit preservation solution. However in this example, the outer OAIS still needs the function Developing Preservation Strategies to ensure that threats to copies are mitigated, functions to Manage Storage Hierarchy and initiate repairs through Disaster Recovery, in case Error Checking e.g. discover differences between copies placed in the different Inner OAISes. In this example, it is clear that the outer OAIS is not an OAIS without the inner OAISes, whereas the inner OAISes individually and independently can be regarded as an OAIS with a very specific Designated Community and very narrow Preservation Objectives that are not concerned with functional interpretation of the bits.

Figure 6-6 illustrates an example where two of the inner OAISes from the previous example are reused. In this example, the two OAISes depend on distribution of inner functional areas with parts managed in other organisations.

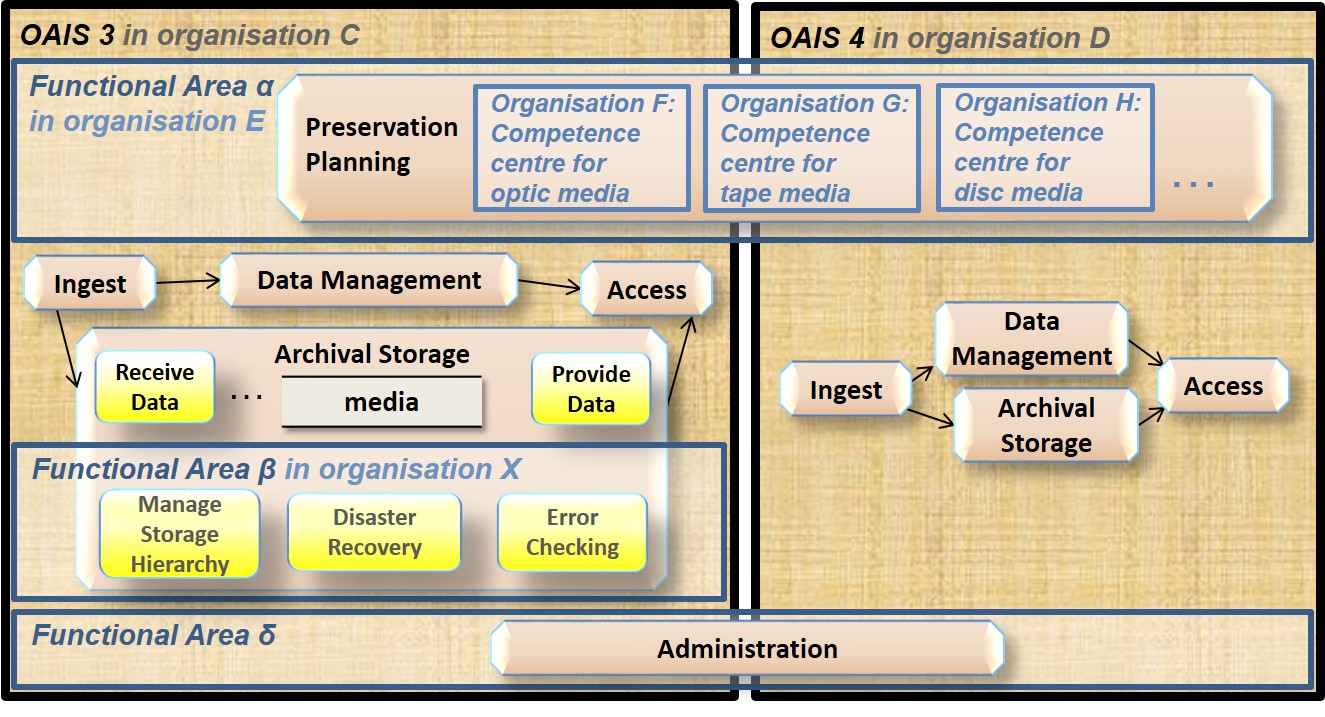


Figure 6-6: Example of Distributed Functional Areas

In this example, the inner functional area α is managed by organisation E, and involving a number of different competence centres managed by other organisations (F, G, H, …) which in the illustration is specialized on media and planning functions. The inner functional area β is managed by organisation X and the inner functional area δ is another way to illustrate that the archives OAIS 2 and 3 have shared functionality as described in section 6-4. Also in this example, the two OAISes are not complete without the contribution from the (inner) Functional Areas.

The use of denotations ‘outer OAIS’ and ‘inner OAIS’ focus means that depending on the focus, an outer OAIS can be an inner OAIS for other OAISes, and vice versa. The OAIS 3 and OAIS 4 are denoted and acting as inner OAISes in Figure 6-5, while in Figure 6-6, they are the outer OAISes depending on inner functional areas. Another example is illustrated in Figure 6-7, where the outer OAIS 1 from Figure 6-5 has different Preservation Objectives: In this example, the OAIS 1 only represents an OAIS performing advanced bit preservation, i.e. without Preservation Objectives related to the use of the bit stream.

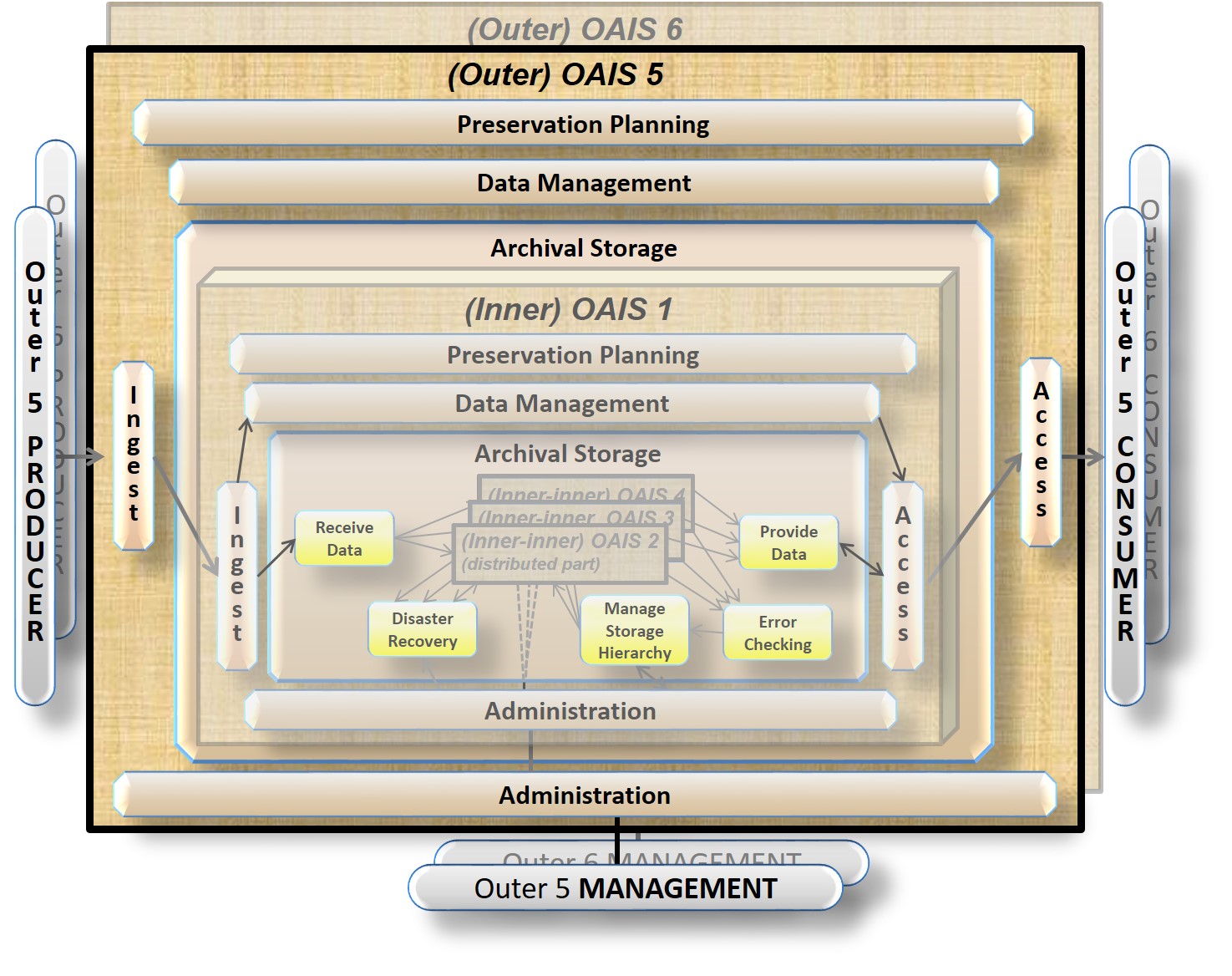


Figure 6-7: Another Example of a Distributed OAIS Archive for Archival Storage

In this example, the focus on OAIS 5 would mean that the OAIS 5 is the outer OAIS depending on the inner OAIS 1. A similar case can be made for OAIS 6 as an outer OAIS that also depends on the OAIS 1 as an inner OAIS.

A third example could be an inner OAIS that had the Preservation Objectives to preserve all systems software needed for emulation. Outer OAIS AIPs relying on emulation would in this case have parts of the Representation Information preserved in the inner OAIS.

A fourth example of Distributed OAISes using inner OAISes could be an inner OAIS that acts as a preliminary archive for SIPs that are in risk of getting lost due to a long processing period before the AIP can be formed in an acceptable way. The inner OAIS would then serve a Designated Community that only wants bit preservation of the SIPs with a minimum of metadata.