

Space data and information transfer systems — Open archival information systems — Reference model

ICS 49.140

National foreword

This British Standard reproduces verbatim ISO 14721:2003 and implements it as the UK national standard.

The UK participation in its preparation was entrusted by Technical Committee ACE/68, Space systems and operations, to Subcommittee ACE/68/-/7, Space data and information transfer systems, which has the responsibility to:

- aid enquirers to understand the text;
- present to the responsible international/European committee any enquiries on the interpretation, or proposals for change, and keep the UK interests informed;
- monitor related international and European developments and promulgate them in the UK.

A list of organizations represented on this subcommittee can be obtained on request to its secretary.

Cross-references

The British Standards which implement international publications referred to in this document may be found in the *BSI Catalogue* under the section entitled “International Standards Correspondence Index”, or by using the “Search” facility of the *BSI Electronic Catalogue* or of British Standards Online.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Summary of pages

This document comprises a front cover, an inside front cover, the ISO title page, pages ii and iii, a blank page, pages 1 to 156, an inside back cover and a back cover.

The BSI copyright date displayed in this document indicates when the document was last issued.

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 15 May 2003

Amendments issued since publication

Amd. No.	Date	Comments

© BSI 15 May 2003

ISBN 0 580 41862 6

INTERNATIONAL STANDARD

ISO 14721

First edition
2003-03-01

Space data and information transfer systems — Open archival information system — Reference model

*Systèmes de transfert des informations et données spatiales —
Système ouvert d'archivage de l'information — Modèle de référence*



Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

International Standard ISO 14721 was prepared by the Consultative Committee for Space Data Systems (CCSDS) (as CCSDS 650.0-B-1, January 2002) and was adopted (without modifications except those stated in clause 2 of this International Standard) by Technical Committee ISO/TC 20, *Aircraft and space vehicles*, Subcommittee SC 13, *Space data and information transfer systems*.

Space data and information transfer systems — Open archival information system — Reference model

1 Scope

This International Standard specifies a reference model for an open archival information system (OAIS). The purpose of this International Standard is to establish a system for archiving information, both digitalized and physical, with an organizational scheme composed of people who accept the responsibility to preserve information and make it available to a designated community.

The scope and field of application are furthermore detailed in subclauses 1.1 and 1.2 of the enclosed CCSDS publication.

2 Requirements

Requirements are the technical recommendations made in the following publication (reproduced on the following pages), which is adopted as an International Standard:

CCSDS 650.0-B-1, January 2002, *Reference Model for an Open Archival Information System (OAIS)*.

For the purposes of international standardization, the modifications outlined below shall apply to the specific clauses and paragraphs of publication CCSDS 650.0-B-1.

Pages i to v

This part is information which is relevant to the CCSDS publication only.

Page D-1

Add the following information to the references indicated:

- [5] Document IEEE 1003.0, February 1995, is equivalent to ISO/IEC/TR 14252:1996.
- [8] Document CCSDS 620.0-B-2, May 1992, is equivalent to ISO 12175:1994.
- [10] Document CCSDS 644.0-B-2, November 2000, is equivalent to ISO 15889:2003.
- [11] Document CCSDS 647.1-B-1, June 2001, is equivalent to ISO 21961:2003.
- [12] Document CCSDS 647.2-B-1, June 2001, is equivalent to ISO 21962:2003.

3 Revision of publication CCSDS 650.0-B-1

It has been agreed with the Consultative Committee for Space Data Systems that Subcommittee ISO/TC 20/SC 13 will be consulted in the event of any revision or amendment of publication CCSDS 650.0-B-1. To this end, NASA will act as a liaison body between CCSDS and ISO.

Consultative Committee for Space Data Systems

**RECOMMENDATION FOR SPACE
DATA SYSTEM STANDARDS**

Reference Model for an Open Archival Information System (OAIS)

CCSDS 650.0-B-1

BLUE BOOK

January 2002



(Blank page)

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

AUTHORITY

Issue:	Blue Book, Issue 1
Date:	January 2002
Location:	Not Applicable

This document has been approved for publication by the Management Council of the Consultative Committee for Space Data Systems (CCSDS) and represents the consensus technical agreement of the participating CCSDS Member Agencies. The procedure for review and authorization of CCSDS Recommendations is detailed in *Procedures Manual for the Consultative Committee for Space Data Systems*, and the record of Agency participation in the authorization of this document can be obtained from the CCSDS Secretariat at the address below.

This Recommendation is published and maintained by:

CCSDS Secretariat
Program Integration Division (Code M-3)
National Aeronautics and Space Administration
Washington, DC 20546, USA

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

STATEMENT OF INTENT

The Consultative Committee for Space Data Systems (CCSDS) is an organization officially established by the management of member space Agencies. The Committee meets periodically to address data systems problems that are common to all participants, and to formulate sound technical solutions to these problems. Inasmuch as participation in the CCSDS is completely voluntary, the results of Committee actions are termed **Recommendations** and are not considered binding on any Agency.

This **Recommendation** is issued by, and represents the consensus of, the CCSDS Plenary body. Agency endorsement of this **Recommendation** is entirely voluntary. Endorsement, however, indicates the following understandings:

- o Whenever an Agency establishes a CCSDS-related **standard**, this **standard** will be in accord with the relevant **Recommendation**. Establishing such a **standard** does not preclude other provisions which an Agency may develop.
- o Whenever an Agency establishes a CCSDS-related standard, the Agency will provide other CCSDS member Agencies with the following information:
 - The **standard** itself.
 - The anticipated date of initial operational capability.
 - The anticipated duration of operational service.
- o Specific service arrangements shall be made via memoranda of agreement. Neither this **Recommendation** nor any ensuing **standard** is a substitute for a memorandum of agreement.

No later than five years from its date of issuance, this **Recommendation** will be reviewed by the CCSDS to determine whether it should: (1) remain in effect without change; (2) be changed to reflect the impact of new technologies, new requirements, or new directions; or, (3) be retired or canceled.

In those instances when a new version of a **Recommendation** is issued, existing CCSDS-related Agency standards and implementations are not negated or deemed to be non-CCSDS compatible. It is the responsibility of each Agency to determine when such standards or implementations are to be modified. Each Agency is, however, strongly encouraged to direct planning for its new standards and implementations towards the later version of the Recommendation.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

FOREWORD

This document is a technical Recommendation for use in developing a broader consensus on what is required for an archive to provide permanent, or indefinite long-term, preservation of digital information.

This Recommendation establishes a common framework of terms and concepts which comprise an Open Archival Information System (OAIS). It allows existing and future archives to be more meaningfully compared and contrasted. It provides a basis for further standardization within an archival context and it should promote greater vendor awareness of, and support of, archival requirements.

Through the process of normal evolution, it is expected that expansion, deletion, or modification of this document may occur. This Recommendation is therefore subject to CCSDS document management and change control procedures which are defined in *Procedures Manual for the Consultative Committee for Space Data Systems*. Current versions of CCSDS documents are maintained at the CCSDS Web site:

<http://www.ccsds.org/>

Questions relating to the contents or status of this document should be addressed to the CCSDS Secretariat at the address indicated on page i.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

At time of publication, the active Member and Observer Agencies of the CCSDS were

Member Agencies

- Agenzia Spaziale Italiana (ASI)/Italy.
- British National Space Centre (BNSC)/United Kingdom.
- Canadian Space Agency (CSA)/Canada.
- Centre National d'Etudes Spatiales (CNES)/France.
- Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR)/Germany.
- European Space Agency (ESA)/Europe.
- Instituto Nacional de Pesquisas Espaciais (INPE)/Brazil.
- National Aeronautics and Space Administration (NASA)/USA.
- National Space Development Agency of Japan (NASDA)/Japan.
- Russian Space Agency (RSA)/Russian Federation.

Observer Agencies

- Austrian Space Agency (ASA)/Austria.
- Central Research Institute of Machine Building (TsNIIMash)/Russian Federation.
- Centro Tecnico Aeroespacial (CTA)/Brazil.
- Chinese Academy of Space Technology (CAST)/China.
- Commonwealth Scientific and Industrial Research Organization (CSIRO)/Australia.
- Communications Research Centre (CRC)/Canada.
- Communications Research Laboratory (CRL)/Japan.
- Danish Space Research Institute (DSRI)/Denmark.
- European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)/Europe.
- European Telecommunications Satellite Organization (EUTELSAT)/Europe.
- Federal Service of Scientific, Technical & Cultural Affairs (FSST&CA)/Belgium.
- Hellenic National Space Committee (HNSC)/Greece.
- Indian Space Research Organization (ISRO)/India.
- Institute of Space and Astronautical Science (ISAS)/Japan.
- Institute of Space Research (IKI)/Russian Federation.
- KFKI Research Institute for Particle & Nuclear Physics (KFKI)/Hungary.
- MIKOMTEK: CSIR (CSIR)/Republic of South Africa.
- Korea Aerospace Research Institute (KARI)/Korea.
- Ministry of Communications (MOC)/Israel.
- National Oceanic & Atmospheric Administration (NOAA)/USA.
- National Space Program Office (NSPO)/Taipei.
- Swedish Space Corporation (SSC)/Sweden.
- United States Geological Survey (USGS)/USA.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

DOCUMENT CONTROL

Document	Title	Date	Status and Substantive Changes
CCSDS 650.0-B-1	Reference Model for an Open Archival Information System (OAIS)	January 2002	Original Issue

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

CONTENTS

<u>Section</u>	<u>Page</u>
1 INTRODUCTION.....	1-1
1.1 PURPOSE AND SCOPE.....	1-1
1.2 APPLICABILITY.....	1-2
1.3 RATIONALE.....	1-3
1.4 CONFORMANCE.....	1-3
1.5 ROAD MAP FOR DEVELOPMENT OF RELATED STANDARDS.....	1-4
1.6 DOCUMENT STRUCTURE.....	1-4
1.7 DEFINITIONS.....	1-6
2 OAIS CONCEPTS.....	2-1
2.1 OAIS ENVIRONMENT.....	2-2
2.2 OAIS INFORMATION.....	2-3
2.3 OAIS HIGH-LEVEL EXTERNAL INTERACTIONS.....	2-7
3 OAIS RESPONSIBILITIES.....	3-1
3.1 MANDATORY RESPONSIBILITIES.....	3-1
3.2 EXAMPLE MECHANISMS FOR DISCHARGING RESPONSIBILITIES.....	3-1
4 DETAILED MODELS.....	4-1
4.1 FUNCTIONAL MODEL.....	4-1
4.2 INFORMATION MODEL.....	4-18
4.3 INFORMATION PACKAGE TRANSFORMATIONS.....	4-48
5 PRESERVATION PERSPECTIVES.....	5-1
5.1 INFORMATION PRESERVATION.....	5-1
5.2 ACCESS SERVICE PRESERVATION.....	5-10
6 ARCHIVE INTEROPERABILITY.....	6-1
6.1 TECHNICAL LEVELS OF INTERACTION BETWEEN OAIS ARCHIVES.....	6-2
6.2 MANAGEMENT ISSUES WITH FEDERATED ARCHIVES.....	6-8

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

CONTENTS (continued)

<u>Section</u>	<u>Page</u>
ANNEX A EXAMPLES OF EXISTING ARCHIVES	A-1
ANNEX B RELATIONSHIPS WITH OTHER STANDARDS OR EFFORTS	B-1
ANNEX C BRIEF GUIDE TO THE UNIFIED MODELING LANGUAGE (UML)	C-1
ANNEX D INFORMATIVE REFERENCES	D-1
ANNEX E A MODEL FOR SOFTWARE USE IN REPRESENTATION INFORMATION	E-1
ANNEX F COMPOSITE FUNCTIONAL VIEW	F-1

<u>Figure</u>	<u>Page</u>
2-1 Environment Model of an OAIS	2-2
2-2 Obtaining Information from Data	2-4
2-3 Information Package Concepts and Relationships	2-5
2-4 OAIS Archive External Data	2-8
4-1 OAIS Functional Entities	4-1
4-2 Functions of Ingest	4-5
4-3 Functions of Archival Storage	4-7
4-4 Functions of Data Management	4-9
4-5 Functions of Administration	4-10
4-6 Functions of Preservation Planning	4-13
4-7 Functions of Access	4-15
4-8 OAIS Data Flow Diagram	4-17
4-9 Administration Context Diagram	4-18
4-10 Information Object	4-20
4-11 Representation Information Object	4-22
4-12 Information Object Taxonomy	4-24
4-13 Information Package Contents	4-31
4-14 Information Package Taxonomy	4-32
4-15 Archival Information Package (AIP)	4-34
4-16 Preservation Description Information (PDI)	4-34
4-17 Package Description	4-35
4-18 Archival Information Package (Detailed View)	4-37
4-19 Archival Specialization of the AIP	4-38
4-20 Archival Specialization of the Package	4-39
4-21 Archival Information Unit (AIU)	4-40
4-22 Unit Description	4-41
4-23 Archive Information Collections Logical View	4-43

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

CONTENTS (continued)

4-24 Collection Descriptions 4-44

4-25 Data Management Information 4-47

4-26 High-Level Data Flows in an OAIS 4-48

5-1 Conceptual View of Relationships Among Names and AIP Components 5-3

6-1 Cooperating Archives with Mutual Exchange Agreement 6-3

6-2 Cooperating Archives with Standard Ingest and Access Methods 6-4

6-3 An OAIS Federation Employing a Common Catalog 6-5

6-4 Archives with Shared Storage 6-8

C-1 Key to UML Relationships C-1

E-1 Layered Information Model E-1

F-1 Composite of Functional Entities F-2

Table

4-1 Examples of PDI Types 4-29

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

1 INTRODUCTION

1.1 PURPOSE AND SCOPE

The purpose of this document is to define the International Organization for Standardization (ISO) Reference Model for an **Open Archival Information System (OAIS)**. An OAIS is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a **Designated Community**. It meets a set of such responsibilities as defined in this document, and this allows an OAIS archive to be distinguished from other uses of the term ‘archive’. The term ‘Open’ in OAIS is used to imply that this Recommendation, as well as future related Recommendations and standards, are developed in open forums, and it does not imply that access to the archive is unrestricted.

The information being maintained has been deemed to need **Long Term Preservation**, even if the OAIS itself is not permanent. **Long Term** is long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. Long Term may extend indefinitely. In this reference model there is a particular focus on digital information, both as the primary forms of information held and as supporting information for both digitally and physically archived materials. Therefore, the model accommodates information that is inherently non-digital (e.g., a physical sample), but the modeling and preservation of such information is not addressed in detail. This reference model:

- provides a framework for the understanding and increased awareness of archival concepts needed for Long Term digital information preservation and access;
- provides the concepts needed by non-archival organizations to be effective participants in the preservation process;
- provides a framework, including terminology and concepts, for describing and comparing architectures and operations of existing and future archives;
- provides a framework for describing and comparing different long term preservation strategies and techniques;
- provides a basis for comparing the data models of digital information preserved by archives and for discussing how data models and the underlying information may change over time;
- provides a foundation that may be expanded by other efforts to cover long-term preservation of information that is NOT in digital form (e.g., physical media and physical samples);
- expands consensus on the elements and processes for long-term digital information preservation and access, and promotes a larger market which vendors can support;
- guides the identification and production of OAIS-related standards.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

The reference model addresses a full range of archival information preservation functions including ingest, archival storage, data management, access, and dissemination. It also addresses the migration of digital information to new media and forms, the data models used to represent the information, the role of software in information preservation, and the exchange of digital information among archives. It identifies both internal and external interfaces to the archive functions, and it identifies a number of high-level services at these interfaces. It provides various illustrative examples and some 'best practice' recommendations. It defines a minimal set of responsibilities for an archive to be called an OAIS, and it also defines a maximal archive to provide a broad set of useful terms and concepts.

1.2 APPLICABILITY

The OAIS model in this document may be applicable to any archive. It is specifically applicable to organizations with the responsibility of making information available for the Long Term. This includes organizations with other responsibilities, such as processing and distribution in response to programmatic needs.

This model is also of interest to those organizations and individuals who create information that may need Long-Term Preservation and those that may need to acquire information from such archives.

The model, including the functional and information modeling concepts, is relevant to the comparison and design of facilities which hold information, on a temporary basis, for two reasons:

- When taking into consideration the rapid pace of technology changes or possible changes in a Designated Community, there is the likelihood that facilities, thought to be holding information on a temporary basis, will in fact find that some or much of their information holdings will need Long Term Preservation attention.
- Although some facilities holding information may themselves be temporary, some or all of their information may need to be preserved indefinitely. Such facilities need to be active participants in the Long Term Preservation effort.

Standards developers are expected to use this model as a basis for further standardization in this area. A large number of related standards are possible. A road map for such development is briefly addressed in 1.5.

This reference model does not specify a design or an implementation. Actual implementations may group or break out functionality differently.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

1.3 RATIONALE

A tremendous growth in computational power, and in networking bandwidth and connectivity, has resulted in an explosion in the number of organizations making digital information available. Transactions among all types of organizations are being conducted using digital forms that are taking the place of more traditional media such as paper.

Preserving information in digital forms is much more difficult than preserving information in forms such as paper and film. This is not only a problem for traditional archives, but also for many organizations that have never thought of themselves as performing an archival function. It is expected that this reference model, by establishing minimum requirements for an OAIS archive along with a set of archival concepts, will provide a common framework from which to view archival challenges, particularly as they relate to digital information. This should enable more organizations to understand the issues and take the proper steps to ensure Long Term information preservation. It should also provide a basis for more standardization and, therefore, a larger market that vendors can support in meeting archival requirements.

1.4 CONFORMANCE

A conforming OAIS archive implementation shall support the model of information described in 2.2. The OAIS Reference Model does not define or require any particular method of implementation of these concepts.

A conforming OAIS archive shall fulfill the responsibilities listed in 3.1. Subsection 3.2 provides examples of the mechanisms that may be used to discharge the responsibilities identified in 3.1. These mechanisms are not required for conformance.

It is assumed that implementers will use this reference model as a guide while developing a specific implementation to provide identified services and content. This document does not assume or endorse any specific computing platform, system environment, system design paradigm, system development methodology, database management system, database design paradigm, data definition language, command language, system interface, user interface, technology, or media required for implementation.

A conformant OAIS archive may provide additional services to users that are beyond those required of an OAIS.

The OAIS Reference Model is designed as a conceptual framework in which to discuss and compare archives. As such, it attempts to address all the major activities of an information-preserving archive in order to define a consistent and useful set of terms and concepts. A standard and other documents that claim to be conformant to the OAIS Reference Model shall use the terms and concepts defined in the OAIS Reference Model in the same manner.

1.5 ROAD MAP FOR DEVELOPMENT OF RELATED STANDARDS

This Reference Model, developed by CCSDS Panel 2 in response to ISO TC20/SC 13, serves to identify areas suitable for the development of OAIS-related standards. Some of these standards may be developed by Panel 2; others may be developed by other standardization bodies. However, any such work undertaken by other bodies should be coordinated in order to minimize incompatibilities and efforts. Areas for potential OAIS-related standards include:

- standard(s) for the interfaces between OAIS type archives;
- standard(s) for the submission (ingest) methodology used by an archive;
- standard(s) for the submission (ingest) of digital data sources to the archive;
- standard(s) for the delivery of digital sources from the archive;
- standard(s) for the submission of digital metadata, about digital or physical data sources, to the archive;
- standard(s) for the identification of digital sources within the archive;
- protocol standard(s) to search and retrieve metadata information about digital and physical data sources;
- standard(s) for media access allowing replacement of media management systems without having to rewrite the media;
- standard(s) for specific physical media;
- standard(s) for the migration of information across media and formats;
- standard(s) for recommended archival practices;
- standard(s) for accreditation of archives.

1.6 DOCUMENT STRUCTURE

1.6.1 HOW TO READ THIS DOCUMENT

All readers should read the Purpose and Scope (1.1), Applicability (1.2), and Conformance (1.4) subsections to obtain a view on the objectives and applicability of the document.

Those who want just an overview of the major concepts should also read OAIS Concepts (section 2) and OAIS Responsibilities (section 3).

Those who will implement OAIS archives or administer them on a daily basis should read the entire document.

1.6.2 ORGANIZATION BY SECTION

Section 1 provides purpose, scope, applicability, and definitions sections typical of many standards. It also provides rationale for the effort, conformance requirements, and a road map for development of related standards.

Section 2 provides a high-level overview of the major concepts involved in an OAIS archive. It provides a view of the environment of an OAIS archive and the roles played by those who

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

interact with it. It discusses what is meant by ‘information’ and what is necessary to preserve it for the Long Term. It contains key information concepts relevant to OAIS-conforming implementations.

Subsection 3.1 defines mandatory responsibilities an OAIS archive must discharge in preserving its information, and 3.2 provides clarifying material of the types of activities that may be needed in many archives to discharge these responsibilities.

Section 4 provides model views needed for a detailed understanding of an OAIS archive. It breaks down the OAIS into a number of functional areas and it identifies some high-level services at the interfaces. It also provides detailed data model views of information using Unified Modeling Language (UML) diagrams.

Section 5 provides some perspectives on the issues of information preservation using digital migration across media and across new formats or representations. It also provides some perspectives on the issues of preserving access services to digital information using software porting, wrapping, and emulation of hardware.

Section 6 is an introduction to the various alternatives for archive-to-archive associations to provide increased or more cost-effective services.

The annexes are not part of the Recommendation and are provided for the convenience of the reader:

- annex A provides scenarios of existing archive operations;
- annex B relates parts of this reference model to other standards work;
- annex C provides a brief tutorial on the Unified Modeling Language (UML);
- annex D provides a list of informative references;
- annex E provides a layered model of information;
- annex F provides a composite diagram of the detailed functional areas described in 4.1.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

1.7 DEFINITIONS**1.7.1 ACRONYMS AND ABBREVIATIONS**

AIC	Archival Information Collection
AIP	Archival Information Package
AIU	Archival Information Unit
ASCII	American Standard Code for Information Interchange
CAD	Computer-Automated Design
CCSDS	Consultative Committee for Space Data Systems
CDO	Content Data Object
CD-ROM	Compact Disk - Read Only Memory
CEOS	Committee on Earth Observation Satellites
CRC	Cyclical Redundancy Check
DED	Data Entity Dictionary
DBMS	Data Base Management System
DDL	Data Description Language
DIP	Dissemination Information Package
DVD	Digital Video Disk
EBCDIC	Extended Binary Coded Decimal Interchange Code
ECS	EOSDIS Core System
EOSDIS	Earth Observing System Data and Information System
FITS	Flexible Image Transfer System
GIF	Graphics Interchange Format
HFMS	Hierarchical File Management System
ICS	Interoperable Catalogue System
IEEE	Institute of Electrical and Electronic Engineers
IMS	Information Management System
ISBN	International Standard Book Number
ISO	International Organization for Standardization
LSDA	Life Sciences Data Archive
NARA	National Archives and Records Administration
NASA	National Aeronautics and Space Administration
NSSDC	National Space Science Data Center
OAIS	Open Archival Information System
ODL	Object Description Language
PDI	Preservation Description Information
PDMP	Project Data Management Plan
PDS	Planetary Data System
PSDD	Planetary Science Data Dictionary
SIP	Submission Information Package
UML	Unified Modeling Language
UNICODE	Universal Code
WWW	World-Wide Web

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

1.7.2 TERMINOLOGY

There are many terms used in this reference model which need to have well-defined meanings, and these are defined in this subsection. When first used in the text, they are shown in bold and are capitalized. Subsequent use employs capitalization only.

As this reference model is applicable to all disciplines and organizations that do, or expect to, preserve and provide information in digital form, these terms can not match all of those familiar to any particular discipline (e.g., traditional archives, digital libraries, science data centers). Rather, the approach taken is to use terms that are not already overloaded with meaning so as to reduce conveying unintended meanings. Therefore we expect all disciplines and organizations will find that they need to map some of their more familiar terms to those of the OAIS Reference Model. This should not be difficult and is viewed as a contribution, rather than a deterrent, to the success of the reference model. For example, archival science focuses on preservation of the ‘record’. This term is not used in the OAIS Reference Model, but one mapping might approximately equate it with ‘Content Information within an Archival Information Package’ (see definitions below, as well as 2.2 and 4.2 for context). A few such mappings are provided in annex B.

Access: The OAIS entity that contains the services and functions which make the archival information holdings and related services visible to Consumers.

Access Aid: A software program or document that allow Consumers to locate, analyze, and order Archival Information Packages of interest.

Access Collection: A collection of AIPs that is defined by a Collection Description but for which there is no Packaging Information for the collection in Archival Storage.

Access Method: A method for retrieving an Archival Information Package based on its name or identifier, which is available to authorized users.

Access Software: A type of software that presents part of or all of the information content of an Information Object in forms understandable to humans or systems.

Adhoc Order: A request that is generated by a Consumer for information the OAIS has indicated is currently available.

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

Archival Information Collection (AIC): An Archival Information Package whose Content Information is an aggregation of other Archival Information Packages.

Archival Information Package (AIP): An Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an OAIS.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

Archival Information Unit (AIU): An Archival Information Package whose Content Information is not further broken down into other Content Information components, each of which has its own complete Preservation Description Information. It can be viewed as an ‘atomic’ AIP. An example of an AIU would be a table of numbers representing temperatures in a certain region with all the associated documentation describing how and where the temperatures were measured, what instruments were used to make the measurements, who made the measurements, why they were made, what processing has been performed on the measurements and who has had custody of these measurements since they were first created, how the measurements relate to other information, how the measurements can be uniquely referenced by others, etc.

Archival Storage: The OAIS entity that contains the services and functions used for the storage and retrieval of Archival Information Packages.

Archive: An organization that intends to preserve information for access and use by a Designated Community.

Associated Description: The information describing the content of an Information Package from the point of view of a particular Access Aid.

Client: An application which exchanges information with another application (see also Consumer).

Collection Description: A type of Package Description that is specialized to provide information about an Archival Information Collection for use by Access Aids.

Common Services: The supporting services such as inter-process communication, name services, temporary storage allocation, exception handling, security, and directory services necessary to support the OAIS.

Consumer: The role played by those persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail. This can include other OAISs, as well as internal OAIS persons or systems.

Content Data Object: The Data Object, that together with associated Representation Information, is the original target of preservation.

Content Information: The set of information that is the original target of preservation. It is an Information Object comprised of its Content Data Object and its Representation Information. An example of Content Information could be a single table of numbers representing, and understandable as, temperatures, but excluding the documentation that would explain its history and origin, how it relates to other observations, etc.

Context Information: The information that documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

Co-operating Archives: Those archives that have Designated Communities with related interests. They may order and ingest data from each other. At a minimum, Co-operating Archives must agree to support at least one common Submission Information Package (SIP) and Dissemination Information Package (DIP) for inter-archive requests.

Data: A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing. Examples of data include a sequence of bits, a table of numbers, the characters on a page, the recording of sounds made by a person speaking, or a moon rock specimen.

Data Dictionary: A formal repository of terms used to describe data.

Data Dissemination Session: A delivered set of media or a single telecommunications session that provides data to a Consumer. The Data Dissemination Session format/contents is based on a data model negotiated between the OAIS and the Consumer in the Request Agreement. This data model identifies the logical constructs used by the OAIS and how they are represented on each media delivery or in the telecommunication session.

Data Management: The OAIS entity that contains the services and functions for populating, maintaining, and accessing a wide variety of information. Some examples of this information are catalogs and inventories on what may be retrieved from Archival Storage, processing algorithms that may be run on retrieved data, Consumer access statistics, Consumer billing, Event Based Orders, security controls, and OAIS schedules, policies, and procedures.

Data Management Data: The data created and stored in Data Management persistent storage that refer to operation of an archive. Some examples of this data are accounting data for Consumer billing and authorization, policy data, Event Based Order (subscription) data for repeating requests, preservation process history data, and statistical data for generating reports to archive management.

Data Object: Either a Physical Object or a Digital Object.

Data Submission Session: A delivered set of media or a single telecommunications session that provides Data to an OAIS. The Data Submission Session format/contents is based on a data model negotiated between the OAIS and the Producer in the Submission Agreement. This data model identifies the logical constructs used by the Producer and how they are represented on each media delivery or in the telecommunication session.

Derived AIP: An AIP generated by extracting or aggregating information from one or more source AIPs.

Descriptive Information: The set of information, consisting primarily of Package Descriptions, which is provided to Data Management to support the finding, ordering, and retrieving of OAIS information holdings by Consumers.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

Designated Community: An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities.

Digital Migration: The transfer of digital information, while intending to preserve it, within the OAIS. It is distinguished from transfers in general by three attributes:

- a focus on the preservation of the full information content;
- a perspective that the new archival implementation of the information is a replacement for the old; and
- an understanding that full control and responsibility over all aspects of the transfer resides with the OAIS.

Digital Object: An object composed of a set of bit sequences.

Dissemination Information Package (DIP): The Information Package, derived from one or more AIPs, received by the Consumer in response to a request to the OAIS.

Edition: An attribute of an AIP whose information content has been increased or improved from a source AIP and is therefore a candidate to replace the source AIP.

Event Based Order: A request that is generated by a Consumer for information that is to be delivered periodically on the basis of some event or events.

Federated Archives: A group of archives that has agreed to provide access to their holdings via one or more common finding aids.

Finding Aid: A type of Access Aid that allows a user to search for and identify Archival Information Packages of interest.

Fixity Information: The information which documents the authentication mechanisms and provides authentication keys to ensure that the Content Information object has not been altered in an undocumented manner. An example is a Cyclical Redundancy Check (CRC) code for a file.

Global Community: An extended Consumer community, in the context of Federated Archives, that accesses the holdings of several archives via one or more common Finding Aids.

Independently Understandable: A characteristic of information that has sufficient documentation to allow the information to be understood and used by the Designated Community without having to resort to special resources not widely available, including named individuals.

Information: Any type of knowledge that can be exchanged. In an exchange, it is represented by data. An example is a string of bits (the data) accompanied by a description

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

of how to interpret a string of bits as numbers representing temperature observations measured in degrees Celsius (the representation information).

Information Object: A Data Object together with its Representation Information.

Information Package: The Content Information and associated Preservation Description Information which is needed to aid in the preservation of the Content Information. The Information Package has associated Packaging Information used to delimit and identify the Content Information and Preservation Description Information.

Ingest: The OAIS entity that contains the services and functions that accept Submission Information Packages from Producers, prepares Archival Information Packages for storage, and ensures that Archival Information Packages and their supporting Descriptive Information become established within the OAIS.

Knowledge Base: A set of information, incorporated by a person or system, that allows that person or system to understand received information.

Local Community: The original Designated Community, in the context of Federated Archives, served by an archive.

Long Term: A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing user community, on the information being held in a repository. This period extends into the indefinite future.

Long Term Preservation: The act of maintaining information, in a correct and Independently Understandable form, over the Long Term.

Management: The role played by those who set overall OAIS policy as one component in a broader policy domain.

Member Description: An Associated Description that describes a member of a collection.

Metadata: Data about other data.

Open Archival Information System (OAIS): An archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community. It meets a set of responsibilities, as defined in 3.1, that allows an OAIS archive to be distinguished from other uses of the term ‘archive’. The term ‘Open’ in OAIS is used to imply that this Recommendation and future related Recommendations and standards are developed in open forums, and it does not imply that access to the archive is unrestricted.

Order Agreement: An agreement between the archive and the Consumer in which the physical details of the delivery, such as media type and format of Data, are specified.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

Ordering Aid: An application that assists the Consumer in discovering the cost of, and in ordering, AIPs of interest.

Overview Description: A specialization of the Collection Description that describes the collection as a whole.

Package Description: The information intended for use by Access Aids.

Packaging Information: The information that is used to bind and identify the components of an Information Package. For example, it may be the ISO 9660 volume and directory information used on a CD-ROM to provide the content of several files containing Content Information and Preservation Description Information.

Physical Object: An object (such as a moon rock, bio-specimen, microscope slide) with physically observable properties that represent information that is considered suitable for being adequately documented for preservation, distribution, and independent usage.

Preservation Description Information (PDI): The information which is necessary for adequate preservation of the Content Information and which can be categorized as Provenance, Reference, Fixity, and Context information.

Producer: The role played by those persons, or client systems, who provide the information to be preserved. This can include other OAISs or internal OAIS persons or systems.

Provenance Information: The information that documents the history of the Content Information. This information tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. Examples of Provenance Information are the principal investigator who recorded the data, and the information concerning its storage, handling, and migration.

Reference Information: The information that identifies, and if necessary describes, one or more mechanisms used to provide assigned identifiers for the Content Information. It also provides identifiers that allow outside systems to refer, unambiguously, to a particular Content Information. An example of Reference Information is an ISBN.

Reference Model: A framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. A reference model is based on a small number of unifying concepts and may be used as a basis for education and explaining standards to a non-specialist.

Refreshment: A Digital Migration where the effect is to replace a media instance with a copy that is sufficiently exact that all Archival Storage hardware and software continues to run as before.

Repackaging: A Digital Migration in which there is an alteration in the Packaging Information of the AIP.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

Replication: A Digital Migration where there is no change to the Packaging Information, the Content Information, and the PDI. The bits used to represent these Information Objects are preserved in the transfer to the same or new media instance.

Representation Information: The information that maps a Data Object into more meaningful concepts. An example is the ASCII definition that describes how a sequence of bits (i.e., a Data Object) is mapped into a symbol.

Representation Network: The set of Representation Information that fully describes the meaning of a Data Object. Representation Information in digital forms needs additional Representation Information so its digital forms can be understood over the Long Term.

Representation Rendering Software: A type of software that displays Representation Information of an Information Object in forms understandable to humans.

Result Set: The set of descriptive records for those AIPs in an OAIS which match the criteria stated in a Consumer query, or other results from a search on Data Management.

Retrieval Aid: An application that allows authorized users to retrieve the Content Information and PDI described by the Package Description.

Search Session: A session initiated by the Consumer with the archive during which the Consumer will use the archive Finding Aids to identify and investigate potential holdings of interest.

Structure Information: The information that imparts meaning about how other information is organized. For example, it maps bit streams to common computer types such as characters, numbers, and pixels and aggregations of those types such as character strings and arrays.

Submission Agreement: The agreement reached between an OAIS and the Producer that specifies a data model for the Data Submission Session. This data model identifies format/contents and the logical constructs used by the Producer and how they are represented on each media delivery or in a telecommunication session.

Submission Information Package (SIP): An Information Package that is delivered by the Producer to the OAIS for use in the construction of one or more AIPs.

Transformation: A Digital Migration in which there is an alteration to the Content Information or PDI of an Archival Information Package. For example, changing ASCII codes to UNICODE in a text document being preserved is a Transformation.

Unit Description: A type of Package Description that is specialized to provide information about an Archival Information Unit for use by Access Aids.

Version: An attribute of an AIP whose information content has undergone a transformation on a source AIP and is a candidate to replace the source AIP.

(Blank page)

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

2 OAIS CONCEPTS

The purpose of this section is to motivate and describe several key high-level OAIS concepts. A more complete view, and a formal modeling of these concepts, is given in section 4.

The term ‘archive’ has come to be used to refer to a wide variety of storage and preservation functions and systems. Traditional archives are understood as facilities or organizations which preserve records, originally generated by or for a government organization, institution, or corporation, for access by public or private communities. The archive accomplishes this task by taking ownership of the records, ensuring that they are understandable to the accessing community, and managing them so as to preserve their information content and authenticity. Historically, these records have been in such forms as books, papers, maps, photographs, and film, which can be read directly by humans, or read with the aid of simple optical magnification and scanning aids. The major focus for preserving this information has been to ensure that they are on media with long-term stability and that access to this media is carefully controlled.

The explosive growth of information in digital forms has posed a severe challenge not only for traditional archives and their information providers, but also for many other organizations in the government, commercial and non-profit sectors. These organizations are finding, or will find, that they need to take on the information preservation functions typically associated with traditional archives because digital information is easily lost or corrupted. The pace of technology evolution is causing some hardware and software systems to become obsolete in a matter of a few years, and these changes can put severe pressure on the ability of the related data structures or formats to continue effective representation of the full information desired. Because much of the supporting information necessary to preserve this information is more easily available or only available at the time when the original information is produced, these organizations need to be active participants in the long-term preservation effort, and they need to follow the principles espoused in this OAIS reference model to ensure that the information can be preserved for the Long Term. Participation in these efforts will minimize the lifecycle costs and enable effective long-term preservation of the information.

The explosion of computer processing power and digital media has resulted in many systems where the Producer role and the archive role are the responsibility of the same entity. These systems, which are sometimes known as Active Archives, should subscribe to the goals of Long Term Preservation discussed in this document. The design process must realize that some of the Long Term Preservation activities may conflict with the goals of rapid production and dissemination of products to Consumers. The designers and architects of such systems should document the solutions that have been reached.

A major purpose of this reference model is to facilitate a much wider understanding of what is required to preserve and access information for the Long Term. To avoid confusion with simple ‘bit storage’ functions, the reference model defines an Open Archival Information System (OAIS) which performs a long-term information preservation and access function.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

An OAIS archive is one that intends to preserve information for access and use by a Designated Community, and it meets the requirements given in section 3. It includes archives that have to keep up with steady input streams of information as well as those that experience primarily aperiodic inputs. It includes archives that provide a wide variety of sophisticated access services as well as those that support only the simplest types of requests. For the remainder of this document, the term archive and OAIS are equivalent and understood to refer to an OAIS archive, unless the context makes it clear otherwise (e.g., traditional archives).

The OAIS model recognizes the already highly distributed nature of digital information holdings and the need for local implementations of effective policies and procedures supporting information preservation. This allows, in principle, a wide variety of organizational arrangements, including various roles for traditional archives, in achieving this preservation. It is expected that organizations attempting to preserve information will find that using OAIS terms and concepts will assist them in achieving their information preservation goals.

2.1 OAIS ENVIRONMENT

The simple model shown in figure 2-1 depicts the environment surrounding an OAIS.

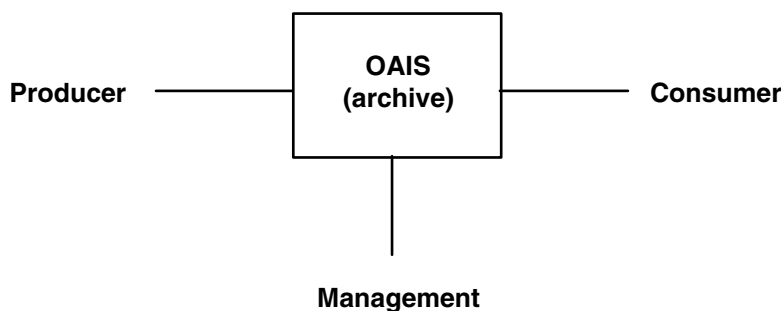


Figure 2-1: Environment Model of an OAIS

Outside the OAIS are **Producers**, **Consumers**, and **Management**.

- Producer is the role played by those persons, or client systems, which provide the information to be preserved.
- Management is the role played by those who set overall OAIS policy as one component in a broader policy domain. In other words, Management control of the OAIS is only one of Management's responsibilities. Management is not involved in day-to-day archive operations. The responsibility of managing the OAIS on a day-to-day basis is included within the OAIS in an administrative functional entity that will be described in 4.1.
- Consumer is the role played by those persons, or client systems, that interact with OAIS services to find and acquire preserved information of interest. A special class

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

of Consumers is the Designated Community. The Designated Community is the set of Consumers who should be able to understand the preserved information.

Other OAIS archives are not shown explicitly. Such archives may establish particular agreements among themselves consistent with Management and OAIS needs. Other archives may interact with a particular archive for a variety of reasons and with varying degrees of formalism for any pre-arranged agreements. One OAIS may take the role of Producer to another OAIS; an example is when the responsibility for preserving a type of information is to be moved to this other archive. One OAIS may take the role of Consumer to another OAIS; an example is when the first OAIS decides to rely on the other OAIS for a type of information it seldom needs and chooses not to preserve locally. Such reliance should have some formal basis that includes the requirement for communication between the archives of any policy changes that might affect this reliance. The range of possible interactions between OAIS archives is discussed in section 6, Archive Interoperability.

2.2 OAIS INFORMATION

2.2.1 INFORMATION DEFINITION

A clear definition of information is central to the ability of an OAIS to preserve it. While formal modeling of information is provided in section 4, some key concepts are provided in this subsection.

A person, or system, can be said to have a **Knowledge Base**, which allows them to understand received information. For example, a person who has a Knowledge Base that includes an understanding of English will be able to read, and understand, an English text.

Information is defined as any type of knowledge that can be exchanged, and this information is always expressed (i.e., represented) by some type of data. For example, the information in a hardcopy book is typically expressed by the observable characters (the data) which, when they are combined with a knowledge of the language used (the Knowledge Base), are converted to more meaningful information. If the recipient does not already include English in its Knowledge Base, then the English text (the data) needs to be accompanied by English dictionary and grammar information (i.e., **Representation Information**) in a form that is understandable using the recipient's Knowledge Base.

Similarly, the information stored within a CD-ROM file is expressed by the bits (the data) it contains which, when they are combined with the Representation Information for those bits, are converted to more meaningful information as long as the Representation Information is understandable using the recipient's Knowledge Base. For example, assume the bits represent an ASCII table of numbers giving the coordinates of a location on the Earth measured in degrees latitude and East longitude. The Representation Information will typically include the definition of ASCII together with descriptions of the format of the numbers and their locations in the file, their definitions as latitude and longitude, and the definition of their units as degrees. It may also include additional meaning that is assigned to

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

the table. In general, it can be said that ‘Data interpreted using its Representation Information yields Information’, and this is shown schematically in figure 2-2.

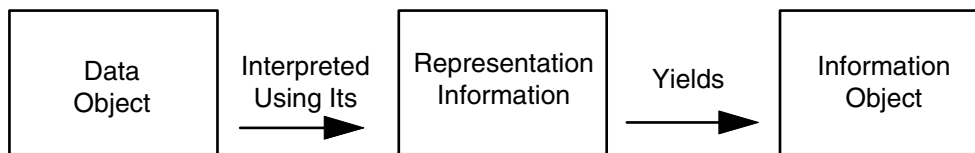


Figure 2-2: Obtaining Information from Data

In order for this **Information Object** to be successfully preserved, it is critical for an OAIS to clearly identify and understand the **Data Object** and its associated Representation Information. For digital information, this means the OAIS must clearly identify the bits and the Representation Information that applies to those bits. This required transparency to the bit level is a distinguishing feature of digital information preservation, and it runs counter to object-oriented concepts which try to hide these implementation issues. This presents a significant challenge to the preservation of digital information.

As a further complication, the recursive nature of Representation Information, which typically is composed of its own data and other Representation Information, typically leads to a network of Representation Information objects. Since a key purpose of an OAIS is to preserve information for a Designated Community, the OAIS must understand the Knowledge Base of its Designated Community to understand the minimum Representation Information that must be maintained. The OAIS should then make a decision between maintaining the minimum Representation Information needed for its Designated Community, or maintaining a larger amount of Representation Information that may allow understanding by a larger Consumer community with a less specialized Knowledge Base. Over time, evolution of the Designated Community’s Knowledge Base may require updates to the Representation Information to ensure continued understanding.

As a practical matter, software is used to access the Information Object, and it will incorporate some understanding of the network of Representation Information objects involved. However, this software should not be used as rationale for avoiding identifying and gathering readily understandable Representation Information that defines the Information Object, because it is harder to preserve working software than to preserve information in digital or hardcopy forms.

The OAIS reference model emphasizes the preservation of information content. As digital technology evolves, multimedia technology and the dependency on complex interplay between the data and presentation technologies will lead some organizations to require that the look and feel of the original presentation of the information be preserved. This type of preservation requirement may necessitate that the software programs and interfaces used to access the data be preserved. This problem may be further complicated by the proprietary nature of some of the software. Various techniques for preserving the look and feel of information access are currently the subject of research and prototyping. These techniques, which include hardware level emulation, emulation of various common service APIs, and the

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

development of a virtual machine, investigate the preservation of the original bit stream and software across technology. Though the OAIS reference model does not focus on these emerging techniques, it should provide architectural basis for the prototyping and comparison of these techniques. A more detailed discussion of the issues involved in the preservation of look and feel of information access can be found in 5.2 of this document.

2.2.2 INFORMATION PACKAGE DEFINITION

The definition of an Information Object is applicable to all the information types discussed in this and the following subsections. In other words, they all have associated Representation Information, although this is usually not shown explicitly.

Every submission of information to an OAIS by a Producer, and every dissemination of information to a Consumer, occurs as one or more discrete transmissions. Therefore, it is convenient to define the concept of an **Information Package**.

An Information Package is a conceptual container of two types of information called **Content Information** and **Preservation Description Information (PDI)**. The Content Information and PDI are viewed as being encapsulated and identifiable by the **Packaging Information**. The resulting package is viewed as being discoverable by virtue of the **Descriptive Information**.

These Information Package relationships are shown schematically in figure 2-3.

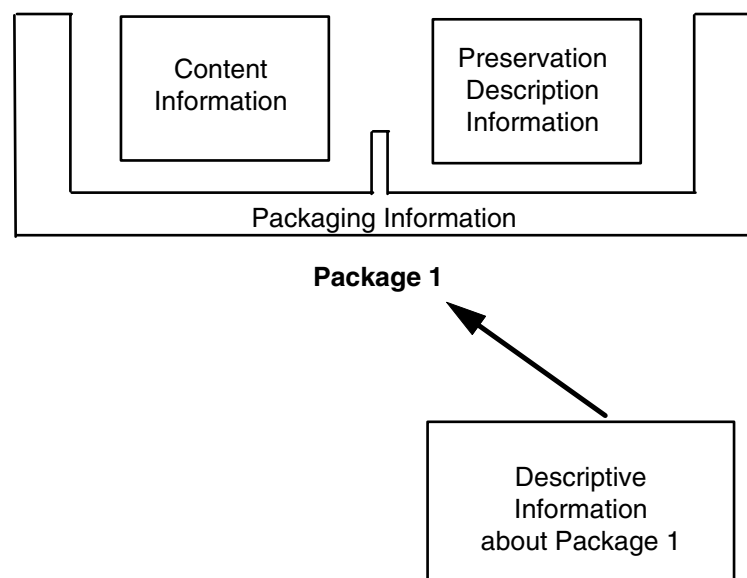


Figure 2-3: Information Package Concepts and Relationships

The Content Information is that information which is the original target of preservation. It consists of the Content Data Object (Physical Object or Digital Object, i.e., bits) and its associated Representation Information needed to make the Content Data Object understandable to the Designated Community. For example, the CDO may be an image that

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

is provided as the bit content of one CD-ROM file together with other files, on the same CD-ROM, that contain Representation Information.

Only after the Content Information has been clearly defined can an assessment of the Preservation Description Information be made. The Preservation Description Information applies to the Content Information and is needed to preserve the Content Information, to ensure it is clearly identified, and to understand the environment in which the Content Information was created. The Preservation Description Information is divided into four types of preserving information called Provenance, Context, Reference, and Fixity. Briefly, they are the following:

- Provenance describes the source of the Content Information, who has had custody of it since its origination, and its history (including processing history).
- Context describes how the Content Information relates to other information outside the Information Package. For example, it would describe why the Content Information was produced, and it may include a description of how it relates to another Content Information object that is available.
- Reference provides one or more identifiers, or systems of identifiers, by which the Content Information may be uniquely identified. Examples include an ISBN number for a book, or a set of attributes that distinguish one instance of Content Information from another.
- Fixity provides a wrapper, or protective shield, that protects the Content Information from undocumented alteration. For example, it may involve a check sum over the Content Information of a digital Information Package.

The Packaging Information is that information which, either actually or logically, binds, identifies and relates the Content Information and PDI. For example, if the Content Information and PDI are identified as being the content of specific files on a CD-ROM, then the Packaging Information would include the ISO 9660 volume/file structure on the CD-ROM, as well as the names and directory information of the files on CD-ROM disk.

The Descriptive Information is that information which is used to discover which package has the Content Information of interest. Depending on the setting, this may be no more than a descriptive title of the Information Package that appears in some message, or it may be a full set of attributes that are searchable in a catalog service.

2.2.3 INFORMATION PACKAGE VARIANTS

It is necessary to distinguish between an Information Package that is preserved by an OAIS and the Information Packages that are submitted to, and disseminated from, an OAIS. These variant packages are needed to reflect the reality that some submissions to an OAIS will have insufficient Representation Information or PDI to meet final OAIS preservation requirements. In addition, these may be organized very differently from the way the OAIS organizes the information it is preserving. Finally, the OAIS may provide information to

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

Consumers that does not include all the Representation Information or all the PDI with the associated Content Information being disseminated. These variants are referred to as the Submission Information Package (SIP), the Archival Information Package (AIP), and the Dissemination Information Package (DIP).

The **Submission Information Package (SIP)** is that package that is sent to an OAIS by a Producer. Its form and detailed content are typically negotiated between the Producer and the OAIS. Most SIPs will have some Content Information and some PDI, but it may require several SIPs to provide a complete set of Content Information and associated PDI to form an AIP. A single SIP may contain information that is to be included in several AIPs. The Packaging Information will always be present in some form.

Within the OAIS one or more SIPs are transformed into one or more **Archival Information Packages (AIP)** for preservation. The AIP has a complete set of PDI for the associated Content Information. The AIP may also contain a collection of other AIPs, and this is discussed and modeled in section 4. The Packaging Information of the AIP will conform to OAIS internal standards, and it may vary as it is managed by the OAIS.

In response to a request, the OAIS provides all or a part of an AIP to a Consumer in the form of a **Dissemination Information Package (DIP)**. The DIP may also include collections of AIPs, and it may or may not have complete PDI. The Packaging Information will necessarily be present in some form so that the Consumer can clearly distinguish the information that was requested. Depending on the dissemination media and Consumer requirements, the Packaging Information may take various forms.

2.3 OAIS HIGH-LEVEL EXTERNAL INTERACTIONS

The following subsections present a high-level view of the interaction between the entities identified in the OAIS environment. Figure 2-4 is a data flow diagram that represents the operational OAIS archive external data flows. This diagram concentrates on the flow of information among Producers, Consumers and the OAIS and does not include flows that involve Management. These flows are dealt with further in section 4.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

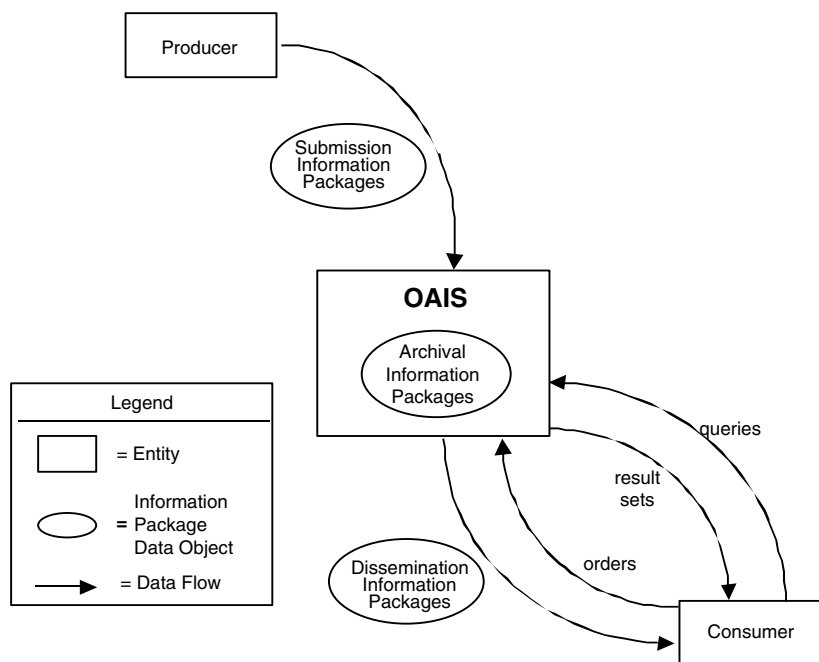


Figure 2-4: OAIS Archive External Data

2.3.1 MANAGEMENT INTERACTION

Management provides the OAIS with its charter and scope. The charter may be developed by the archive, but it is important that Management formally endorse archive activities. The scope determines the breadth of both the Producer and Consumer groups served by the archive.

Some examples of typical interactions between the OAIS and Management include:

- Management is often the primary source of funding for an OAIS and may provide guidelines for resource utilization (personnel, equipment, facilities).
- Management will generally conduct some regular review process to evaluate OAIS performance and progress toward long-term goals.
- Management determines, or at least endorses, pricing policies, as applicable, for OAIS services.
- Management participates in conflict resolution involving Producers, Consumers and OAIS internal administration.

Effective Management should also provide support for the OAIS by establishing procedures that assure OAIS utilization within its sphere of influence. For example, management policies should require that all funded activities within its sphere of influence submit data products to the archive and also adhere to archive standards and procedures.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

2.3.2 PRODUCER INTERACTION

The first contact between the OAIS and the Producer is a request that the OAIS preserve the data products created by the Producer. This contact may be initiated by the OAIS, the Producer or Management. The Producer establishes a **Submission Agreement** with the OAIS, which identifies the SIPs to be submitted and may span any length of time for this submission. Some Submission Agreements will reflect a mandatory requirement to provide information to the OAIS, while others will reflect a voluntary offering of information. Even in the case where no formal Submission Agreement exists, such as a World Wide Web (WWW) site, a virtual Submission Agreement may exist specifying the file formats and the general subject matter the site will accept.

Within the Submission Agreement, one or more **Data Submission Sessions** are specified. There may be significant time gaps between the Data Submission Sessions. A Data Submission Session will contain one or more SIPs and may be a delivered set of media or a single telecommunications session. The Data Submission Session content is based on a data model negotiated between the OAIS and the Producer in the Submission Agreement. This data model identifies the logical components of the SIP (e.g., the Content Information, PDI, Packaging Information, and Descriptive Information) that are to be provided and how (and whether) they are represented in each Data Submission Session. All data deliveries within a Submission Agreement are recognized as belonging to that Submission Agreement and will generally have a consistent data model, which is specified in the Submission Agreement. For example, a Data Submission Session may consist of a set of Content Information corresponding to a set of observations, which are carried by a set of files on a CD-ROM. The Preservation Description Information is split between two other files. All of these files need Representation Information which must be provided in some way. The CD-ROM and its directory/file structure are the Packaging Information, which provides encapsulation and identification of the Content Information and PDI in the Data Submission Session. The Submission Agreement indicates how the Representation Information for each file is to be provided, how the CD-ROM is to be recognized, how the Packaging Information will be used to identify and encapsulate the SIP Content Information and PDI, and how frequently Data Submission Sessions (e.g., one per month for two years) will occur. It also gives other needed information such as access restrictions to the data.

Each SIP in a Data Submission Session is expected to meet minimum OAIS requirements for completeness. However, in some cases multiple SIPs may need to be received before an acceptable AIP can be formed and fully ingested within the OAIS. In other cases, a single SIP may contain data to be included in many AIPs. A Submission Agreement also includes, or references, the procedures and protocols by which an OAIS will either verify the arrival and completeness of a Data Submission Session with the Producer or question the Producer on the contents of the Data Submission Session.

2.3.3 CONSUMER INTERACTION

There are many types of interactions between the Consumer and the OAIS. These interactions include questions to a help desk, requests for literature, catalog searches, orders

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

and order status requests. The ordering process is of special interest to the OAIS Reference Model, since it deals with the flow of archive holdings between the OAIS and the Consumer.

The Consumer establishes an **Order Agreement** with the OAIS for information. This information may currently exist in the archive or be expected to be ingested in the future. The Order Agreement may span any length of time, and under it one or more **Data Dissemination Sessions** may take place. A Data Dissemination Session may involve the transfer of a set of media or a single telecommunications session. The Order Agreement identifies one or more AIPs of interest, how those AIPs are to be transformed and mapped into Dissemination Information Packages (DIPs) and how those DIPs will be packaged in a Data Dissemination Session. The Order Agreement will also specify other needed information such as delivery information (e.g., name or mailing address), and any pricing agreements as applicable. There are two common order types initiated by Consumers: the **Event Based Order** and the **Adhoc Order**.

In the case of an Adhoc Order, the Consumer establishes an Order Agreement with the OAIS for information available from the archive. If the Consumer does not know *a priori* what specific holdings of the OAIS are of interest, the Consumer will establish a **Search Session** with the OAIS. During this Search Session the Consumer will use the OAIS **Finding Aids** that operate on **Descriptive Information**, or in some cases on the AIPs themselves, to identify and investigate potential holdings of interest. This may be accomplished by the submission of queries and the return of result sets to the Consumer. This searching process tends to be iterative, with a Consumer first identifying broad criteria and then refining these criteria based on previous search results. Once the Consumer identifies the OAIS AIPs of interest, the Consumer may provide an Order Agreement that documents the identifiers of the AIPs the Consumer wishes to acquire, and how the DIPs will be acquired from the OAIS. If the AIPs are available, an Adhoc Order will be placed. However if the AIPs desired are not yet available, an Event Based Order may be placed.

In the case of an Event Based Order, the Consumer establishes an Order Agreement with the OAIS for information expected to be received on the basis of some triggering event. This event may be periodic, such as a monthly distribution of any AIPs ingested by the OAIS from a specific Producer, or it may be a unique event such as the ingestion of a specific AIP. The Order Agreement will also specify other needed information such as the trigger event for new Data Dissemination Sessions and the criteria for selecting the OAIS holdings to be included in each new Data Dissemination Session.

The Order Agreement does not have to be a formal document. In general an OAIS will have a general pricing policy and maintain an information base of the electronic and physical mailing addresses of its users. In this case, the process of developing an Order Agreement may be no more than the completion of a World Wide Web form to specify the AIPs of interest.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

3 OAIS RESPONSIBILITIES

Subsection 3.1 identifies the minimal responsibilities that shall be discharged by an OAIS. Subsection 3.2 gives some examples of mechanisms to discharge these responsibilities, although not all of these will be applicable to all OAISs.

3.1 MANDATORY RESPONSIBILITIES

This subsection establishes mandatory responsibilities that an organization must discharge in order to operate an OAIS archive.

The OAIS must:

- Negotiate for and accept appropriate information from information Producers.
- Obtain sufficient control of the information provided to the level needed to ensure Long-Term Preservation.
- Determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided.
- Ensure that the information to be preserved is **Independently Understandable** to the Designated Community. In other words, the community should be able to understand the information without needing the assistance of the experts who produced the information.
- Follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, and which enable the information to be disseminated as authenticated copies of the original, or as traceable to the original.
- Make the preserved information available to the Designated Community.

3.2 EXAMPLE MECHANISMS FOR DISCHARGING RESPONSIBILITIES

This subsection provides example mechanisms for discharging the responsibilities of 3.1. Not all of these mechanisms will be applicable to all OAISs.

3.2.1 NEGOTIATES FOR AND ACCEPTS INFORMATION

An organization operating an OAIS will have established some criteria that aids in determining the types of information that it is willing to, or it is required to, accept. These criteria may include, among others, subject matter, information source, degree of uniqueness or originality, and the nature of the techniques used to represent the information (e.g., physical media, digital media, format). The information may, in general, be submitted using a wide variety of common and not-so-common forms, such as books, documents, maps, data sets, and moon rocks using a variety of communication paths including networks, mail, and special delivery.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

The OAIS negotiates with the Producer to ensure it acquires appropriate Content Information and associated PDI for its mission and the Designated Community. Considerable iteration may be required to agree on the right information to be submitted, and to get it into forms acceptable to the OAIS. For example, this may include digitizing audio or video material, or scanning text.

The OAIS should extract, or otherwise obtain, sufficient Descriptive Information to assist the Designated Community in finding the Content Information of interest. It also should ensure that the information meets all OAIS internal standards.

3.2.2 OBTAINS SUFFICIENT CONTROL FOR PRESERVATION

It is important for the OAIS to recognize the separation that can exist between physical ownership or possession of Content Information and ownership of intellectual property rights in this information. If it has created the information and is the legal owner of the Content Information, the OAIS already has the independence to do what is required to preserve the information and make it available. When acquiring the Content Information from any other producer or entity, the OAIS must ensure that there is a legally valid transfer agreement that either transfers intellectual property rights to the OAIS, or clearly specifies the rights granted to the OAIS and any limitations imposed by the rightsholder(s). The OAIS must ensure that its subsequent actions to preserve the information and make it available conform with these rights and limitations. When the OAIS does not acquire the intellectual property rights, the agreement should specify what involvement the rightsholder(s) will have in preservation, management or release of the information. In most cases, it will be preferable for the OAIS to negotiate an agreement that specifies the rightsholder(s) requirements and authorizes the OAIS to act in accordance with those requirements without active involvement of the rightsholder(s) in individual cases.

The OAIS must assume sufficient control over the Content Information and Preservation Description Information so that it is able to preserve it for the Long Term. There is no issue with the AIP's Packaging Information because, by definition, this is under internal OAIS control. The problems of assuming sufficient control of the Content Information and Preservation Description Information, when they are largely digital, are addressed in three related categories, as follows:

- copyright implications; intellectual property and other legal restrictions on use;
- authority to modify Representation Information;
- agreements with external organizations.

Copyright implications, intellectual property and other legal restrictions on use: An archive will honor all applicable legal restrictions. These issues occur when the OAIS acts as a custodian. An OAIS should understand the copyright concepts and applicable laws prior to accepting copyright materials into the OAIS. It can establish guidelines for ingestion of information and rules for dissemination and duplication of the information when necessary. It is beyond the scope of this document to provide details of national and international copyright laws.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

Authority to modify Representation Information: Although the Fixity information within the Preservation Description Information of an AIP ensures that the Content Information-related bits have not been altered, there will come a time when Content Information bits are not in a form that is convenient for the Consumers from the Designated Community. The Content Information bits may be fully documented in available hardcopy forms, so technically the information has not been lost, but practically the information has become inaccessible. The OAIS needs the authority to migrate the Content Information to new representation forms. If it is acting as a custodian, it may need to seek additional permission to make such changes. If the information is copyrighted, the OAIS should already have negotiated permission to make the changes needed to meet preservation objectives. It may employ subject matter experts, from outside the OAIS, to help ensure that information is not lost. Ideally, when this situation arises, both the original AIPs (fully described) and new AIPs will be retained. Digital Migration issues are addressed more fully in section 0.

Agreements with external organizations: An OAIS may establish a variety of agreements with other organizations to assist in its preservation objectives. For example, it may establish an agreement with another OAIS so that it does not have to preserve all the common Representation Information objects related to its Content Information objects. Agreements with other organizations should be monitored to be sure they are being followed and remain useful.

3.2.3 DETERMINES DESIGNATED CONSUMER COMMUNITY

The submission, or planned submission, of Content Information and associated PDI requires a determination as to who the expected Consumers, or Designated Community, of this information will be. This is necessary in order to determine if the information, as represented, will be understandable to that community. For example, an archive may decide that certain Content Information should be understandable to the general public and, therefore, this becomes the Designated Community.

For some scientific information, the Designated Community of Consumers might be described as those with a first year graduate level education in a related scientific discipline. This is a more difficult case as it is less clear what degree of specialized scientific terminology might actually be acceptable. The Producers of such specialized information are often familiar with a narrowly recognized set of terminology, so it is especially critical to clearly define the Designated Community for their information and to make the effort to ensure that this community can understand the information.

The possible evolution of the definition of the Designated Community also needs consideration. Information originally intended for a narrowly defined community may need to be made more widely understandable at some future date. For example, information originally intended to be understandable to a particular scientific community may need to be made understandable to the general public. This is likely to mean adding explanations in support of the Representation Information and the Preservation Description Information, and it can become increasingly difficult to obtain this information over time. Selecting a broader definition of the Designated Community (e.g., general public) when the information is first

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

proposed for Long-Term Preservation can reduce this concern and also improve the likelihood that the information will be understandable to all in the original community.

3.2.4 ENSURES INFORMATION IS INDEPENDENTLY UNDERSTANDABLE

The degree to which Content Information and its associated PDI conveys information to a Designated Community is, in general, quite subjective. Nevertheless, it is essential that an archive make this determination in order to maximize information preservation. Digital Content Information and PDI need adequate Representation Information to be Independently Understandable to the Designated Community. Typically there are multiple Representation Information objects involved, and this is discussed in 4.2.

For example, consider Content Information from a digital set of observations of rainfall, temperature, pressure, wind velocities, and other parameters measured all over the world for a year. This type of information is very extensive, and is not usually in a form intended for direct human browsing or reading, but it is in a form appropriate to searching and manipulation by application software. Such content may only be understandable to the original Producers, unless there is adequate documentation of the meaning of the various fields and their inter-relationships, and how the values relate back to the original instrumentation that made the observations. In such specialized fields extra effort is needed to ensure that the Content Information and the Preservation Description Information are understandable to a Designated Community. If the archive does not have this level of expertise in-house, it may have outside community representatives review the information for long-term understandability. Otherwise some of the information may be understandable to only a few specialists and be lost when they are no longer available.

Even when a set of information has been determined to be understandable to a particular Designated Community, over time the Knowledge Base of this community may evolve to the point that important aspects of the information may no longer be readily understandable. At this point it may be necessary for the OAIS to enhance the associated Representation Information so that it is again readily understandable to the Designated Community.

As another example, a manuscript's Content Information may be written in English and therefore its content may be generally understandable to a wide audience. However, unless the purpose for which it was created is clearly documented, much of its meaning may be lost. This 'purpose' information is part of its Context and must be provided in the Preservation Description Information.

Software is needed for efficient access to Digital Content Information. However, maintaining Content Information-specific software over the Long Term has not yet been proven cost effective due to the narrow application of such software. The danger of information loss is great when such software is relied upon for information preservation and understanding because it may cease to function under only small changes to the hardware and software environment. This may not be recognized unless there is a vigorous, ongoing, testing and validation program.

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

3.2.5 FOLLOWS ESTABLISHED PRESERVATION POLICIES AND PROCEDURES

It is essential for an OAIS to have documented policies and procedures for preserving its AIPs, and it should follow those procedures. The appropriate policies and procedures will depend, at minimum, on the nature of the AIPs and any 'backup' relationships the archive may have with other archives. For example, migrations that alter any Content Information or PDI will need to be carefully monitored and the appropriate PDI fully updated. This attention to detail, while also ensuring against processing errors, requires that strong policies and procedures be in place and that they be executed.

The Producer and Consumer communities should be provided with submission and dissemination standards, policies, and procedures to support the preservation objectives of the OAIS.

The Designated Community should be monitored to be sure the Content Information is still understandable to them. The Designated Community may lose its familiarity with some terminology, and the definition of the community may be broadened to include other members with different backgrounds. For example, a periodic review with participants representing the Designated Community could assist in this process.

A long-term technology usage plan, updated as technology evolves, is essential to avoid being caught with very costly system maintenance, emergency system replacements, and costly data representation transformations.

3.2.6 MAKES THE INFORMATION AVAILABLE

By definition, an OAIS makes its AIPs visible and available to its Designated Communities. Multiple views of its holdings, supported by various search aids that may cut across collections of AIPs, may be provided. Some AIPs may only exist as the output of algorithms operating on other AIPs. They appear as DIPs that, upon dissemination, should include documentation on how they were derived from other AIPs. The expectations of OAIS Consumers regarding access services will vary widely among archives and over time as technology evolves. Pressures for more effective access must be balanced with the requirements for preservation under the available resource constraints.

Some AIPs may have restricted access and therefore may only be disseminated to Consumers who meet access restrictions. The OAIS should have published policies on access and restrictions so that the rights of all parties are protected.

In general, DIPs may be distributed by all varieties of communication paths, including networks and physical media.

(Blank page)

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

4 DETAILED MODELS

The purpose of this section is to provide a more detailed model view of the functional entities of the OAIS and the information handled by the OAIS. This aids OAIS designers of future systems and provides a more precise set of terms and concepts for discussion of current systems.

4.1 FUNCTIONAL MODEL

The OAIS of figure 2-1 is separated in figure 4-1 into six functional entities and related interfaces. Only major information flows are shown. The lines connecting entities identify communication paths over which information flows in both directions. The lines to Administration and Preservation Planning are dashed only to reduce diagram clutter.

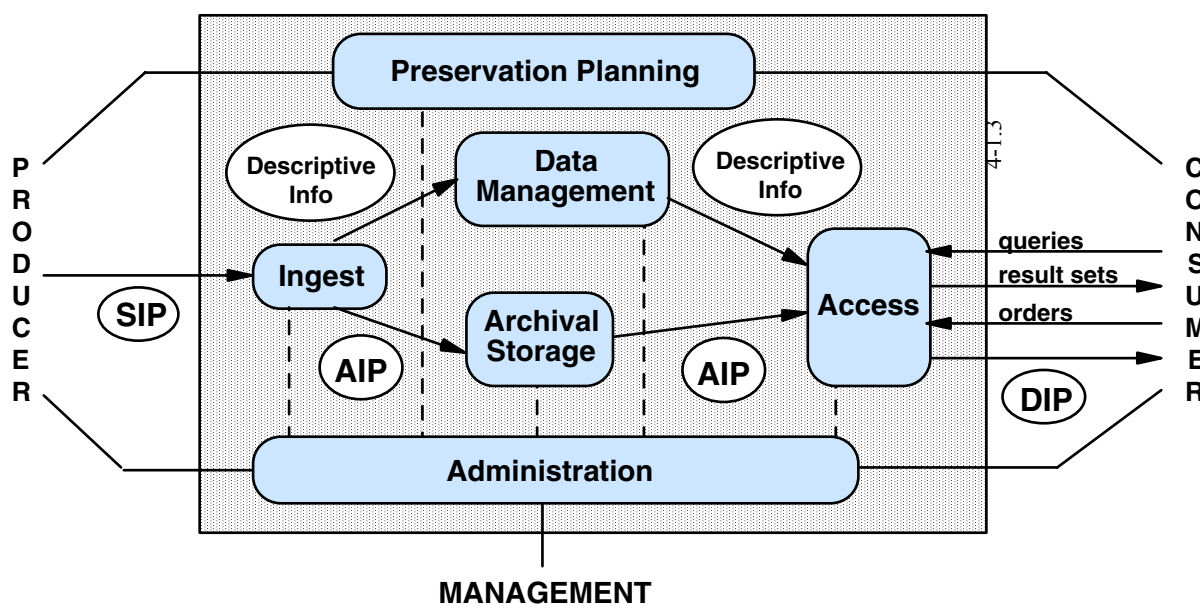


Figure 4-1: OAIS Functional Entities

The role provided by each of the entities in figure 4-1 is described briefly as follows:

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. Ingest functions include receiving SIPs, performing quality assurance on SIPs, generating an Archival Information Package (AIP) which complies with the archive's data formatting and documentation standards, extracting Descriptive Information from the AIPs for inclusion in the archive database, and coordinating updates to Archival Storage and Data Management.

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

CCSDS RECOMMENDATION FOR AN OAIS REFERENCE MODEL

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.

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. Data Management functions include administering the archive database functions (maintaining schema and view definitions, and referential integrity), performing database updates (loading new descriptive information or archive administrative data), performing queries on the data management data to generate result sets, and producing reports from these result sets.

Administration: This entity provides the services and functions for the overall operation of the archive system. Administration functions include soliciting and negotiating submission agreements with Producers, auditing submissions to ensure that they meet archive standards, and maintaining configuration management of system hardware and software. It also provides system engineering functions to monitor and improve archive operations, and to inventory, report on, and migrate/update the contents of the archive. It is also responsible for establishing and maintaining archive standards and policies, providing customer support, and activating stored requests.

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 functions include evaluating the contents of the archive and periodically recommending archival information updates to migrate current archive holdings, developing recommendations for archive standards and policies, and monitoring changes in the technology environment and in the Designated Community's service requirements and Knowledge Base. Preservation Planning also designs IP templates and provides design assistance and review to specialize these templates into SIPs and AIPs for specific submissions. . Preservation Planning also develops detailed Migration plans, software prototypes and test plans to enable implementation of Administration migration goals.

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. Access functions include communicating with Consumers to receive requests, applying controls to limit access to specially protected information, coordinating the execution of requests to successful completion, generating responses (Dissemination Information Packages, result sets, reports) and delivering the responses to Consumers.

In addition to the entities described above, there are various **Common Services** assumed to be available. These services are considered to constitute another functional entity in this model. This entity is so pervasive that, for clarity, it is not shown in figure 4-1.