Open Archival Information System (OAIS) Reference Model

Christopher A. Lee
School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, U.S.A.

Abstract
The Reference Model for an Open Archival Information System (OAIS) describes components and services required to develop and maintain archives, in order to support long-term access to and understanding of the information in those archives. This entry discusses the context in which the OAIS was initiated and provides a chronology of the OAIS development process, including its transformation from a space data standard into a document of much wider scope. The author explains the nature of reference models as a particular type of standard, and then describes the major components and concepts of the OAIS Reference Model. The primary mechanisms that the Reference Model uses to convey the aspects of an OAIS are its functional model and information model. The entry also summarizes numerous (completed, ongoing, and emerging) initiatives that have adopted and expanded on the OAIS. Finally, the author discusses major implications of the OAIS.

INTRODUCTION

The Reference Model for an Open Archival Information System (OAIS) describes components and services required to develop and maintain archives, in order to support long-term access to and understanding of the information in those archives.[1] The development of the OAIS took place within a standards development organization called the Consultative Committee for Space Data Systems (CCSDS), whose formal purview is the work of space agencies, but the effort reached far beyond the traditional CCSDS interests and stakeholders.[2] It has become a fundamental component of digital archive research and development in a variety of disciplines and sectors.

This entry discusses the context in which the OAIS was initiated and provides a chronology of the OAIS development process, including its transformation from a space data standard into a document of much wider scope. It then describes the major components and concepts of the OAIS Reference Model. The entry also summarizes numerous initiatives that have adopted and expanded on the OAIS. Finally, the author discusses major implications of the OAIS.

HISTORICAL CONTEXT

Digital preservation has long been an area of concern for those responsible for repositories of digital objects. Data mismanagement, technological dependency, media degradation, and technological obsolescence have all threatened the long-term accessibility of resources stored in digital formats. In the 1950s, 1960s, and 1970s, organizations began increasingly to rely on collections of computer-dependent data. Several streams of activity gradually emerged to address parts of the digital preservation problem (e.g., care and physical properties of physical storage media; hardware and software interoperability; management and provision of access to digital library collections), but there was often little communication or coordination across the streams. The streams of activity often developed their own distinct forums (journals, conferences, consortia) and sets of funding mechanisms (government budget areas, research agendas, foundation support).

Two trends that began in the 1960s and 1970s, but became much more prominent during the 1980s and early 1990s, were: 1) actors with long traditions of preserving physical artifacts (e.g., archivists, librarians, museum curators) increasingly recognizing that information which fell within the scope of their responsibility was now digital; and 2) actors with long traditions of managing computer-dependent data sets (e.g., scientific data center personnel, corporate information technology managers) increasingly recognizing that information which fell within the scope of their responsibility had long-term preservation value. The effort to develop the OAIS came at a time when the separate streams of activity were making important progress but they were only beginning to identify points of intersection between the streams. During the years immediately preceding and throughout the OAIS development effort, participants in digital preservation work increasingly recognized that they were addressing similar issues.

The 1990s were characterized by a broadening societal awareness of both the importance of standards in supporting the infrastructure that underlies various activities and

the challenges of long-term digital preservation. Two closely connected factors were the development of widely distributed computer networks and an industry trend toward commercial off-the-shelf (COTS) equipment. Rather than depending on the compatibility of an entire suite of hardware and software from a single vendor, producers and consumers of computer equipment came to rely on conventions for interchange of data between a heterogeneous set of components. Both the International Organization for Standardization (ISO) and Internet Engineering Task Force (IETF) developed layered architectures to which hardware and software producers could conform in order to ensure that their products could interchange data with other products on the Internet. In the early 1990s, the adoption of the World Wide Web reflected and contributed to a dramatic growth in the base of consumers who had a stake in the interchange of data over computer networks. The Year 2000 (Y2K) conversion effort was one widely recognized example of dependence on a widespread assortment of hardware and software components that usually interoperated in ways invisible to most people, but could cause significant problems when they failed to interoperate. In 1994, the Commission on Preservation and Access (CPA) and Research Libraries Group (RLG) created a Task Force on Digital Archiving. The Task Force issued a report in 1996 called *Preserving Digital Information*, which was frequently cited by subsequent literature on digital preservation.[3] Two works that brought digital preservation to popular attention were a 1995 article by Jeff Rothenberg in *Scientific American* called "Ensuring the longevity of digital documents"[4] and a movie called *Into the Future*,[5] which was released in 1997 and shown on Public Broadcasting Service (PBS) across the United States in 1998.

Prior to the development of the OAIS, there had been numerous calls in professional literature for both development and adoption of open standards in support of long-term digital preservation. The most active standardization and consensus was associated with physical storage media and storage conditions. In other areas of digital preservation, most of the existing standards had served primarily to advance work within specific streams of activities, rather than spanning multiple professions.

OAIS DEVELOPMENT PROCESS

The historical context discussed in the previous section is essential to understanding the character and impact of the OAIS. The Reference Model was both a product of and an influential factor in the evolution of digital preservation activities in the mid- to late-1990s—reflecting preexisting notions and helping to define new notions. A major factor in the success of the OAIS was the timing of its development. Actors within several streams of activity related to digital preservation perceived the need for a high-level model but had not themselves developed one. At the same time, several actors now felt they had knowledge from their own recent digital archiving efforts, which could inform the development of the OAIS. Despite this growing body of practical experience and understanding of the functions associated with digital archiving, one essential element that was missing was a common vocabulary. Problems had often stemmed from terms—such as archives/archiving or metadata—that were used so widely and for so many different purposes that it was difficult to determine if they were being used in the same way by different actors. The combination of pressing need, available expertise, and inconsistent language meant the time was ripe for developing a reference model that could codify and support greater consistency in discussions of digital archives.

The development of the OAIS took place within the CCSDS, whose formal purview was specifically support for study of the terrestrial and space environments. However, the OAIS development effort took on a much wider scope than one may have reasonably predicted, given its CCSDS origins. The OAIS development process ultimately involved and gained visibility among a much broader set of stakeholders than simply members of the CCSDS. The word "Open" in the acronym OAIS—meant to indicate that the standard was "developed in open forums"—was a defining feature of its evolution. The leader of the OAIS development process was Don Sawyer, Computer Scientist and head of the U.S. National Aeronautics and Space Administration (NASA) Office of Standards and Technology (NOST) at the Goddard Space Flight Center (GSFC) in Greenbelt, Maryland. Sawyer indicated early and continued to reiterate that the process should be open and inclusive, in order to get sufficient input and buy-in. The core development team actively recruited commentary on the drafts of the document and gave dozens of public presentations about the Reference Model to a diverse range of professional groups.

The formal process began on April 5, 1994, when Gael Squibb of the Jet Propulsion Laboratory (JPL) proposed a New Work Item (NWI) related to "archiving space data" to ISO Technical Committee 20 (Aircraft and Space Vehicles), Subcommittee 14 (Space Systems and Operations). This proposal ultimately found a home in another subcommittee of ISO TC 20: Subcommittee 13 (Space Data and Information Transfer Systems). The formal Secretariat for TC 20/SC 13 is the American National Standards Institute (ANSI), and it is administered by NASA. The CCSDS is a liaison organization to TC 20/SC 13 (see Fig. 1). Sawyer made the case for this effort to NASA management. As the NASA representative, Sawyer also submitted a document to Panel 2 (Information Interchange Processes) of the CCSDS on April 25, 1995, proposing a new "work package." In May, Panel 2 created "WP [Work Package] 700 Archiving" with Sawyer as the leader and the initial subtask being "WP 710 Archiving

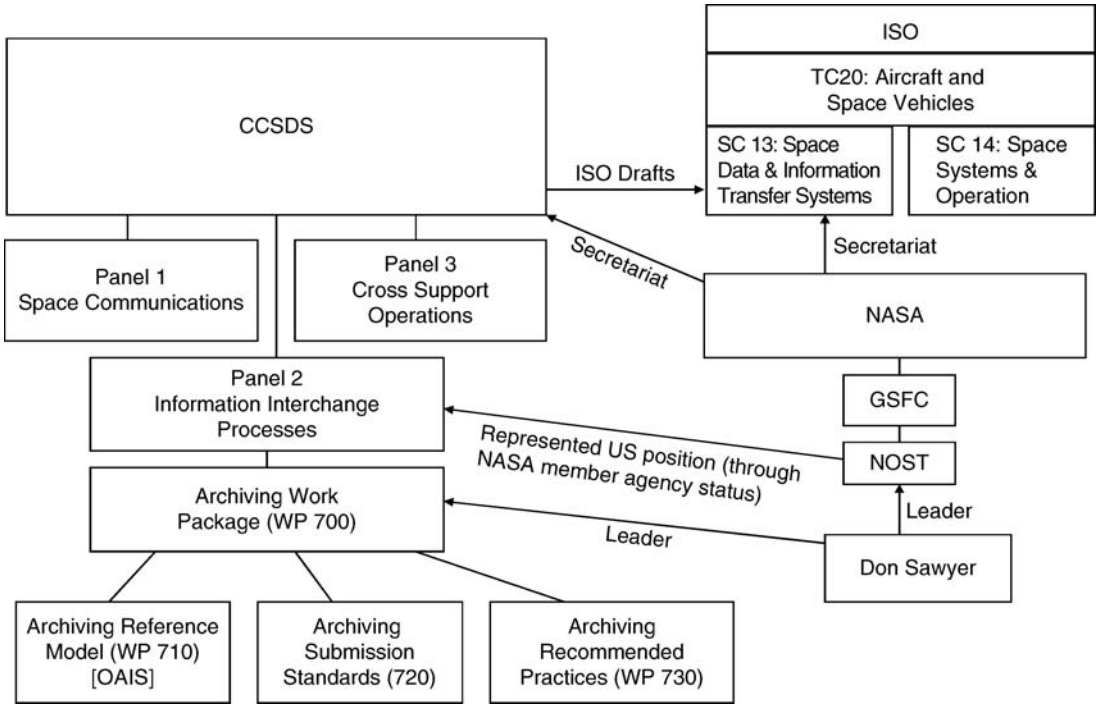

Fig. 1 Organizational context of OAIS development.
Source: Adapted from C. Lee, Defining *Digital Preservation Work: A Case Study of the Development of the Reference model for an Open Archival Information System*. Doctoral Dissertation, University of Michigan: Michigan, 2005.[2] CCSDS, Consultative Committee for Space Data Systems; GSFC, Goddard Space Flight Center; ISO, International Organization for Standardization; NASA, U.S. National Aeronautics and Space Administration; NOST, NASA Office of Standards and Technology.

Reference Model." Sawyer and Lou Reich, who worked under contract for the GSFC as Senior Consulting Engineer for the Computer Sciences Corporation, coauthored a preliminary discussion draft of a reference model document. Sawyer and Reich remained the primary authors and editors of later versions, but the process also increasingly drew from contributions (both text and figures) from other participants. John Garrett, who was a contractor for the NSSDC as Senior Analyst at Hughes STX (ST Systems Corporation), helped to set up the initial meeting, and he served as a significant source of administrative support and documentation throughout the OAIS development process.

Development of the OAIS was markedly different from the previous standards development efforts of the CCSDS by being both broader in scope and inclusive of a more diverse set of actors. In order to support this unusually inclusive effort, the leaders set up a unique meeting and decision-making structure. In addition to the well-established semiannual CCSDS meetings, the OAIS development effort also involved a set of 18 U.S. Workshops, devoted primarily to document development, and a set of more open meetings (one in France, two in the United Kingdom, and two in the United States), designed to gather input and review from a wider set of actors. Most of the U.S. Workshops took place at the Archives II facility of the U.S. National Archives and Records Administration (NARA); Bruce Ambacher, Archives Specialist at NARA, was one of the most active participants in the OAIS development process. The creation of the Reference Model involved a relatively small and stable set of core actors, but it also involved a much larger set of actors who had more limited Workshop participation. The latter played an extremely important role in the development, review, and visibility of the Reference Model. There were 306 individuals who participated in one or more of the Workshops.

Development of the OAIS involved negotiation over issues such as the scope of the Reference Model, its intended purpose, and the definition of basic terms. The development process also involved considerable borrowing and adaptation of ideas and documents already in existence. Over time, common notions about the content of the Reference Model became more established, and the number and extent of revisions to drafts of the Reference Model decreased.

After being requested to do so by CCSDS Panel 2, ISO SC 13 voted on May 22, 2000 to allow circulation of Red Book 1[6] as a Draft International Standard (DIS) to all ISO national member bodies. By releasing the Reference Model as a Red Book, the CCDS was indicating that the document was "technically mature and ready for extensive and formal review by appropriate technical organizations within each Member Agency." Comments received

in response to Red Book 1 implied substantive changes to the document. SC 13 sent a request to the ISO in January 2001 to "reprocess" the DIS, meaning a revised document would be resubmitted to ISO member bodies for another vote. Red Book 1.1 (April 20, 2001), Red Book 1.2 (June 2001), and Red Book 2 (July 2001) were the result of efforts to address the comments received in response to Red Book 1. On October 23, 2001, SC 13 approved the submission of Red Book 2 for ISO vote. There were then a few minor editorial changes to the document before it was circulated (as the Blue Book)[7] to ISO member bodies for balloting. A Blue Book is a CCSDS Recommendation, which "reflects resolution of official comments from Member Agencies during formal reviews, and, as such, represents the consensus of appropriate implementing organizations within each Member Agency. Member Agency approval of a Blue Book implies an intent to reflect its provisions in future data systems standards developed through internal mechanisms." The ISO balloting process, which involved ISO national member bodies, ran from January 24 to April 5, 2002. There were eight affirmative votes, no negative votes, and two abstentions. The chair of SC 13 formally reported the results of voting on the Reference Model to ISO on August 20, 2002. On February 24, 2003, the Reference Model was finally published by ISO as an International Standard (ISO 14721:2003).[8] The Blue Book version, which is freely available from the CCSDS Web site, is identical in content to the International Standard, except that the current Blue Book version reflects a September 2007 "editorial correction," involving slight alterations to several figures to correct "display anomalies" (problems with encoding of line breaks had resulted in extraneous square blocks appearing next to some text elements) and a "typographical error" in one of the figures.

Many important activities in the development of the OAIS took place outside the context of formal ISO Archiving meetings and review processes. Teleconferencing, e-mail, and the Web greatly facilitated work on the OAIS and allowed review and use of the Reference Model drafts by those who were not necessarily able to attend the Workshops. Core members of the OAIS development effort gave dozens of presentations related to the Reference Model at conferences and other professional events. Several of the actors in the ISO Archiving Workshops also took part in mass storage systems and technologies (MSST) conferences, which served as forums for both the dissemination of information about the development of the Reference Model and recruitment of actors into the process. Individuals and organizations involved in initiatives that attempted to apply and test the Reference Model—e.g., CEDARS (CURL [Consortium of Research Libraries] Exemplars in Digital Archives) in the United Kingdom, NEDLIB (Networked European Deposit Library), British Library, and PANDORA (Preserving and Accessing Networked Documentary Resources of Australia)—played an essential role in the process. The archives certification work that built off of the Reference Model also provided important input into the development of the OAIS and demonstrated the Reference Model's potential value to a variety of stakeholders.

Development of the Reference Model drew from many other sources. These included concepts, terminology, models, and strings of text and images from guidelines, reports, and standards. In some cases—such as the "Z39.50 profile for access to digital collections" (PDC),[9] *Preserving Digital Information*, *Planetary Data System Data Preparation Workbook*,[10] and *IEEE Guide to the POSIX Open System Environment (OSE)*[11]—it is possible to identify specific terms or concepts that were incorporated into the Reference Model. Other sources were discussed during Workshops and provided varying degrees of conceptual background for the work on the Reference Model.

OAIS AS A REFERENCE MODEL

The OAIS is a reference model, which is a very particular type of standard. Carl Cargill provides a framework for describing standards, which distinguishes providers of information technology products and services from the users of those products and services. According to Cargill, these two groups have very different motivations and needs in the standardization process.

> On the provider side is the global model that describes all of the potentials that the IT industry will need to satisfy all users over a long time in nearly all situations, and that serves as a reference for all providers. This reference model, if it is correctly constructed, includes some present and future technologies, a road map function, and some of the methodologies of the thought processes that occurred when it was constructed. The time span covered is up to ten years, and the model is applicable to all technical disciplines that deal in this area. On the IT user side is a description of a solution implementation that is immediate and particular to that user's application problems (p. 1-11).[1]

As explained by Cargill above, reference models operate at a higher level of abstraction than other types of standards and are purposely designed to be "implementation-independent." The OAIS itself provides the following definition of 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 nonspecialist.[12]

Cargill explains that bridging the gap between reference models and application implementations requires a chain of standards at increasing levels of specificity (p. 1-12).[1] Within the arena of digital archives, this means that one would expect more specific standards to emerge, which indicate how to apply the concepts of the OAIS in more specific contexts. As explained later in the section on "OAIS Adoption, Extension, and Future Directions," such follow-on standardization has been taking place.

CONTENT OF THE OAIS

The OAIS is a 148-page document, composed of six sections and six annexes. Section 1 frames the content to follow, by providing discussions of purpose, scope, applicability, definitions, rationale, and conformance requirements. It also situates the document in a larger context by including a "road map for development of related standards." Section 2 lays out several core concepts that are then modeled in more detail in Section 4. These include archive; information (as distinct from data); interfaces between an archive, Producers, Consumers, and Management; and Information Package and its subtypes: Submission Information Package (SIP), Archival Information Package (AIP), and Dissemination Information Package (DIP). Section 3 discusses the responsibilities of an Open Archival Information System and "some examples of mechanisms to discharge these responsibilities." Section 4 presents a "more detailed model view" of the concepts previously laid out in the document. The section includes a functional model (including a high-level view and then unpacking of each entity and data flows between the entities) and an information model, which provides a hierarchical set of views and explanations for what logical elements should be stored and managed in association with a data object. Section 4 also provides an account of the "transformations, both logical and physical, of the Information Package and its associated objects as they follow a lifecycle from the Producer to the OAIS, and from the OAIS to the Consumer." Section 5 provides some discussion—intended to be implementation agnostic—of technical issues and strategies that an archive can potentially use to address changes in underlying hardware, software, formats, and access services. Section 6 discusses potential arrangements between multiple archives. The annexes that follow Section 6 are not considered part of the Reference Model's normative content but are instead "provided for the convenience of the reader." The annexes include a set of five "scenarios" that use OAIS terminology and concepts to describe specific existing archives; explanations of how the Reference Model relates to other standards and projects; a brief Unified Modeling Language (UML) tutorial; list of references; a layered model of how software could be used to support Representation Information; and a "composite diagram" that presents in one place the detailed interfaces between each of the entities in the functional model.

Fundamental Terms and Concepts

The Reference Model defines an OAIS as "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." Many of the requirements for an OAIS are based on the needs of its Designated Community, which is the set of one or more "user communities" that the OAIS is serving. An OAIS is responsible for digital information over the "long term," which 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" (p. 1-11).[1] One of the fundamental challenges of an OAIS is to ensure that the digital information in its care is "independently understandable" to the Designated Community, which means that the Designated Community can understand it "without having to resort to special resources not widely available, including named individuals" (p. 1-10).[1] The Designated Community has a certain "knowledge base," which can change over time.

One of the most important insights embedded in the Reference Model is that the "Content Information" to be preserved by an archive is composed not only of a "set of bit sequences" (the "data object") but also associated sufficient "Representation Information" to allow the bits to be rendered, used, and understood.

> 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 (p. 2-4).[1]

The three main roles played by the external entities with which an OAIS interacts are Producer, Consumer, and Management. Producers are "persons, or client systems, who provide the information to be preserved." Consumers are "persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail." Management is "the role played by those who set overall OAIS policy as one component in a broader policy domain (p. 1-11)."[1]

A diversity of workflow and collaborative arrangements can be mapped to the Reference Model by representing

"other OAISs" and "internal OAIS persons or systems" as Producers and Consumers (pp. 1-8, 1-12).[1] The Reference Model describes "four categories of archive association" based on "successively higher degrees of interaction":

- Independent—"no management or technical interaction."
- Cooperating—"potential common producers, common submission standards, and common dissemination standards, but no common finding aids."
- Federated—serve "both a Local Community (i.e., the original Designated Community served by the archive) and a Global community (i.e., an extended Designated Community) which has interests in the holdings of several OAIS archives and has influenced those archives to provide access to their holdings via one or more common finding aids."
- Shared resources—have agreed to share resources between the archives, which "requires various standards internal to the archive (such as ingest-storage and access-storage interface standards), but does not alter the user community's view of the archive" (p. 6-2).[1]

The primary mechanisms that the Reference Model uses to convey the aspects of an OAIS are its functional model and information model. Roughly speaking, the former indicates what an OAIS must do, and the latter indicates what the OAIS must have in its collections.

Functional Model

The functional model is composed of seven main functional entities and the interfaces between them: Access, Administration, Archival storage, Common Services, Data Management, Ingest, and Preservation Planning. The single figure from the Reference Model that has received the most attention in the digital preservation literature is a representation of six of the functional entities (see Fig. 2). Not directly represented in Fig. 2 are Common Services, which are the "supporting services" that must be in place for computer systems to operate and perform properly, including interprocess communication, name services, temporary storage allocation, exception handling, backup, directory services, as well as other aspects of operating system, network, and security services. Although Common Services are necessary for an OAIS, they are not a major focus of the Reference Model, because they are "assumed to be available" (p. 4-2).[1] As illustrated in Fig. 2, the OAIS functional model also identifies between Administration and management, but it does not elaborate any function within management itself.

Many aspects of the functional model rest on the distinction between SIPs, AIPs, and DIPs. SIPs are what the OAIS receives from Producers. AIPs are what the OAIS manages and preserves. DIPs are "derived from one or more AIPs, [and are] received by the Consumer in response to a request to the OAIS" (p. 1-10).[1]

Ingest is the entity that receives SIPs, performs quality assurance on the SIPs, generates AIPs, extracts Descriptive Information from AIPs, and coordinates updates to Archival Storage and Data Management. Archival Storage is responsible for "receiving AIPs from Ingest and adding them to permanent storage, managing the storage hierarchy, refreshing the media on which archive holdings are stored, performing routine and special error checking, providing disaster recovery capabilities, and providing AIPs to access to fulfill orders" (pp. 4-1, 4-2).[1] Data Management supports the populating, maintenance and accessing of both Descriptive Information and administrative data associated with the OAIS holdings; this includes database administration, database updates, performing queries on the data, and producing reports that result from

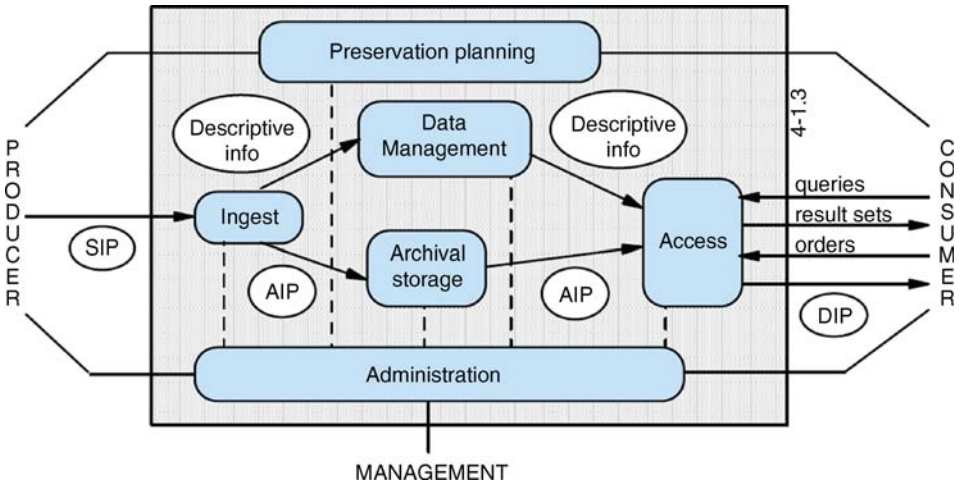

Fig. 2 OAIS Functional Entities.
Source: Reference model for an Open Archival Information System (OAIS); Consulative Committee for Space Data System: Washington, D.C, 2002; 4:1. [SIP = Submission Information Package; AIP = Archival Information Package; DIP = Dissemination Information Package].

the queries. Administration is responsible for "the overall operation of the archive system," which includes soliciting and negotiating submission agreements, auditing submissions, configuration management, system engineering, activating stored requests, and the establishment and maintenance of standards and policies (p. 4-2).[1] Preservation Planning monitors the environment for important changes in technology or the needs of the Designated Community; evaluates the implications of those changes to the archive's holdings; designs Information Package templates; "provides design assistance and review to specialize these templates into SIPs and AIPs for specific submissions"; "develops detailed Migration plans, software prototypes and test plans," and provides periodic recommendations for "archival information updates," standards and policies (p. 4-2).[1] Access both provides and appropriately restricts Consumers' ability to discover, request, and receive information from the archive, including DIPs, "result sets" and reports (p. 4-2).[1]

Information Model

The information model defines and describes "the types of information that are exchanged and managed within the OAIS" (p. 4-18).[1] It is based on the recognition that long-term preservation of digital information requires an archive to "store significantly more than the contents of the object it is expected to preserve" (p. 4-19).[1] The Reference Model uses the term "Information Package" to refer to the logical unit that includes both a digital object and the other types of information that should be associated with the digital object in order to preserve and provide meaningful access to it over time. AIPs are the information packages that are managed internally by the OAIS. An AIP can be either an Archival Information Collection (AIC), "whose Content Information is an aggregation of other Archival Information Packages" (p. 1-7),[1] or Archival Information Unit (AIU) "whose Content Information is not further broken down into other Content Information components."[24] Fig. 3 presents the main types of information that constitute and are associated with an AIP. The Package Description is information about an Information Package, which is used by Access Aids. An Access Aid is "a software program or document that allow[s] Consumers to locate, analyze, and order Archival Information Packages of interest" (p. 1-7).[1] Packaging Information, in contrast, is not intended for direct use by Consumer but is instead "used to bind and identify the components of an Information Package" (e.g., volumes and directory information for the components) (p. 1-12).[1]

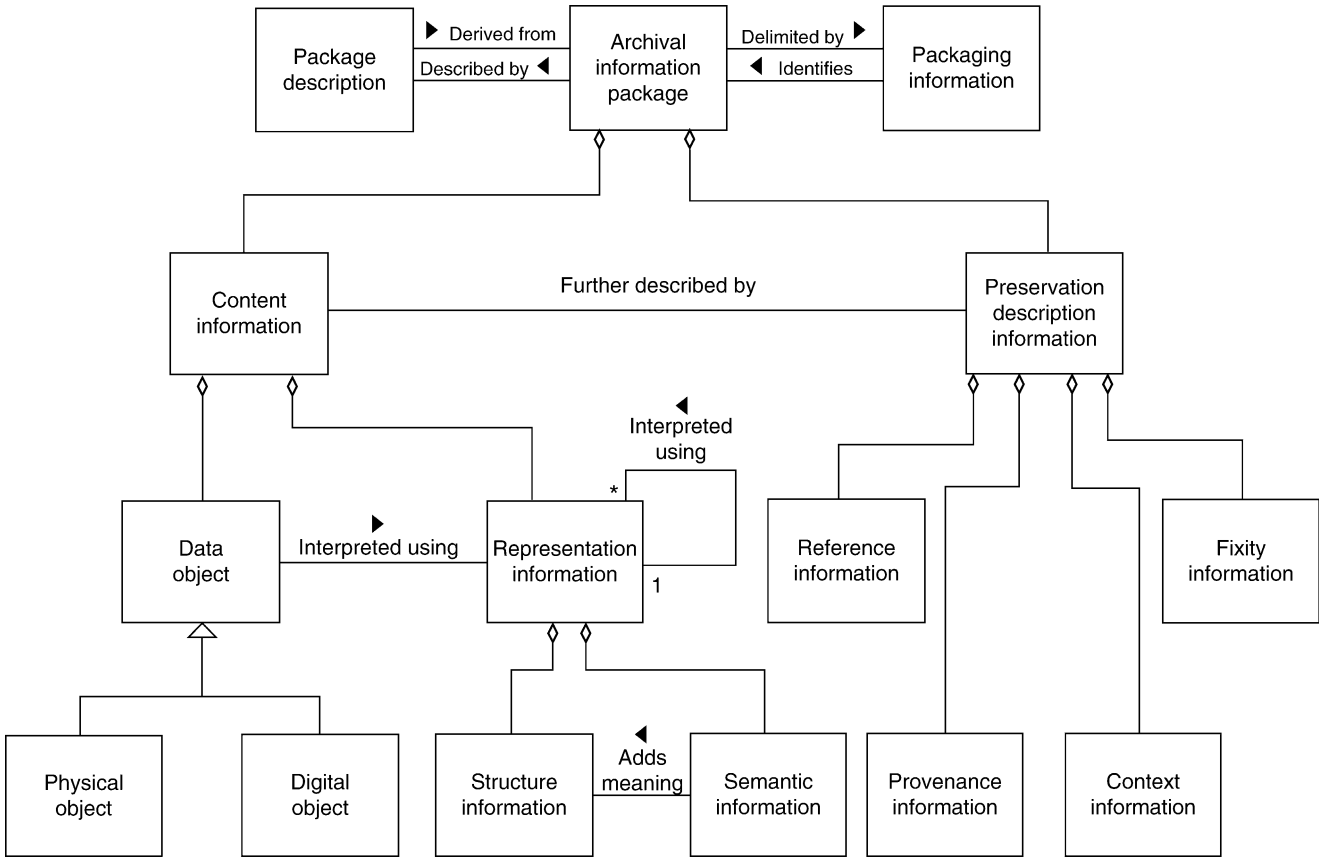

Fig. 3 Archival Information Package (Detailed view).
Source: Reference Model for an Open Archival Information System (OAIS); Consulative Committee for Space Data Systems: Washington, D.C, 2002-4-37.

The Content Information is the "the original target of preservation" (p. 1-8).[1] As described earlier, Content Information is composed of both the Data Object (for digital information, one or more sequences of bits) and Representation Information, which "allows for the full interpretation of the data into meaningful information" (p. 4-19).[1] Representation Information can be either Structure Information, which "imparts meaning about how other information is organized," such as mapping "bit streams to common computer types such as characters, numbers, and pixels and aggregations of those types such as character strings and arrays" (p. 1-13),[1] or Semantic Information, which indicates the meaning of language used in either the Structure Information or other parts of the Content Information.

Representation Information is often conveyed using Digital Objects, which then require their own Representation Information, resulting in what is called a Representation Network. For example, a string of bits (Digital Object) could represent an array of pixels, but a Consumer would also need information about the file format (Representation Information), in order to render and use that Digital Object as an image. If the image were encoded using Scalable Vector Graphics (SVG), which is based on Extensible Markup Language (XML), then rendering and using the image would require not only the SVG specification but also the specification for XML, as well as other specifications for character encoding and Uniform Resource Identifiers (URIs) upon which XML itself depends.

As illustrated in Fig. 3, Content Information is "further described by" Preservation Description Information (PDI), which "is necessary for adequate preservation of the Content Information." PDI is composed of:

- Reference Information—"identifies, and if necessary describes, one or more mechanisms used to provide assigned identifiers for the Content Information" and "provides identifiers that allow outside systems to refer, unambiguously, to a particular Content Information."[14]
- Provenance Information—"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" (p. 1-12).[1]
- Context Information—"documents the relationships of the Content Information to its environment" including "why the Content Information was created and how it relates to other Content Information objects" (p. 1-8).[1]
- Fixity Information—"documents the authentication mechanisms and provides authentication keys to ensure that the Content Information object has not been altered in an undocumented manner" (p. 1-10).[1]

PDI is an extension and elaboration of what *Preserving Digital Information* called the "features that determine information integrity and deserve attention for archival purposes": "content, fixity, reference, provenance, and context."[13]

OAIS ADOPTION, EXTENSIONS, AND FUTURE DIRECTIONS

Several years before it had reached formally approved status within the CCSDS or ISO, the OAIS was already receiving considerable attention from those engaged in digital preservation research and development. An extensive survey of the English-language literature up to April 2005 identified 335 sources that cite or discuss the OAIS, beginning in 1997 and increasing every year thereafter.[14] Over the past several years, the OAIS has come to be a widely assumed basis for research and development on digital archives, with conference papers, articles, and reports very often presenting their findings within the context of the OAIS.

The OAIS has become "the reference model of choice of those involved in digital preservation worldwide,"[15] serving as a "galvanizing force"[16] and a "major factor in the advancement of digital archiving efforts."[17] It has contributed "a common language and concepts for different professional groups involved in digital preservation and developing archiving systems."[18] "The reference model represented common ground upon which to consolidate understanding of the needs and requirements of digital preservation: an opportunity to gather the strands of isolated digital preservation activities, merging them into a shared (albeit highly conceptual) characterization of the problem's boundaries."[19]

A large number of research and development projects have either based their work directly on or claimed that their final products conform to the OAIS, many of which are listed in Table 1.

Professional development has also been strongly influenced by the OAIS. One of the earliest activities of the Digital Preservation Coalition (DPC) in the United Kingdom after its formation in 2001 was to develop, along with the British National Space Centre (BNSC), a seminar to discuss and "raise the profile of" the OAIS. The Cornell University Library offered a highly acclaimed workshop series from 2003 to 2006, and disseminated an associated tutorial, both called Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Problems, which used the OAIS as a foundation.[20] In 2008, Inter-University Consortium for Political and Social Research (ICPSR) became the host of the Digital Preservation Management workshop and tutorial, which continue to be based heavily on the OAIS.

From its initial conception, the OAIS was intended to serve as the basis for further development of more specific digital archives standards, and the OAIS has indeed played that role. It has served as the basis for several very

Table 1 Examples of research and development activities with stated OAIS influence or requirements

Major institutions, organizations, and repositories	British Library
	Digital Curation Centre in the United Kingdom
	European Union
	French Archive Institute
	French National Archiving Center for University and Scientific Publications (CINES)
	Inter-University Consortium for Political and Social Research (ICPSR)
	JSTOR (Journal Storage)
	National Library of Australia
	National Library of France (BnF)
	National Library of New Zealand
	National Library of the Netherlands (KB)
	Online Computer Library Center (OCLC)
	Portico
	Project Euclid—Cornell University Library and Duke University Press
	U.S. Government Printing Office
	U.S. Library of Congress, National Library of China
	U.S. National Archives and Records Administration
	U.S. National Library of Medicine
	U.S. National Oceanic and Atmospheric Administration
	United Nations Educational, Scientific and Cultural Organization (UNESCO)
	Numerous university libraries and space agencies
Funding bodies emphasizing OAIS in solicitations and research agendas	U.S. National Science Foundation
	Joint Information Systems Committee (JISC)—United Kingdom
Digital collection management systems and preservation platforms	aDORe Archive—Los Alamos National Laboratory
	CONTENTdm—OCLC
	DSpace—Massachusetts Institute of Technology
	Flexible Extensible Digital Object and Repository Architecture (Fedora)
	Digital Information Archiving System (DIAS)—IBM
	LOCKSS (Lots of Copies Keep Stuff Safe)—Stanford University Libraries
	Integrated Rule-Oriented Data System (iRODS)—Data Intensive Cyber Environments (DICE) Group
	Preservation and Long-term Access through Networked Services (PLANETS) Testbed
	Securing a Hybrid Environment for Research Preservation and Access (SHERPA)
	Digital Preservation (DP) Service—Arts and Humanities Data Service (AHDS)

prominent digital preservation metadata initiatives, including CEDARS, NEDLIB, and two joint Research Libraries Group (RLG)/Online Computer Library Center (OCLC) efforts—the Working Group on Preservation Metadata and then the Preservation Metadata Implementation Strategies (PREMIS) Working Group. The CCSDS also has been coordinating the development of follow-on standards based on the OAIS, which provide more detailed guidance related to parts of the OAIS functional and information models. The *Producer–Archive Interface Methodology Abstract Standard* (PAIMAS) "defines the methodology for the structure of actions that are required from the initial time of contact between the producer and the archive until the objects of information are received and validated by the archive."[21] PAIMAS was issued as a CCSDS Blue Book in May 2004 and was published as an ISO standard (ISO 20652) in 2006.[22] The *Producer–Archive Interface Specification* (PAIS) focuses on an even finer level of granulating, by presenting "a standard method to formally define the digital information objects to be transferred by an information Producer to an Archive and for effectively transferring these objects in the form of Submission Information Packages (SIPs)." As of the writing of this entry, the PAIS was in a tenth White Book version (April 2009), undergoing consideration for Red Book status. A White Book is "a preliminary draft of a planned CCSDS Recommendation or Report" that is "under development" and "not necessarily endorsed by any CCSDS Member or Observer Agency or given any CCSDS external distribution."[23]

RLG and NARA formed a Digital Repository Certification Task Force, whose efforts were explicitly tied to the OAIS. An initiative by the Center for Research Libraries (CRL) is extending the RLG/NARA certification work, with funding from the Andrew W. Mellon Foundation. A Birds of a Feather group is also attempting to develop an ISO standard for digital repository audit and certification, through the same channels as the OAIS, viz. ISO TC 20, SC 13, and the CCSDS.

The OAIS is currently undergoing a 5-years review within the CCSDS and then within the ISO. During the

comment period, which ended in October 2006, the CCSDS received 11 separate documents containing comments. At a meeting on October 4, 2007, the Digital Archive Ingest (DAI) Working Group began reviewing the OAIS review comments, and identified several action items related to specific parts of the text. In 2003, the CCSDS underwent a major reorganization. The digital archives standardization activities that were previously part of Panel 2, including the OAIS, were moved into a new area called Mission Operations and Information Management Services (MOIMS). The DAI working group, is part of MOIMS. The DAI Working Group determined an initial set of responses to the comments, shared the potential responses with those who submitted the original comments, and then revised the OAIS document based on the comments and follow-on interchanges. On May 5, 2009, John Garrett (Chair) and David Giaretta (Deputy Chair) of the DAI Working Group disseminated a proposed draft of the revised OAIS Reference Model through several electronic mailing lists, "seeking primarily to identify errors" in preparation for submitting the revised Reference Model for ISO review and balloting.

CONCLUSION

The OAIS has been a focal point and foundation for discussion of digital preservation and digital curation across many professional boundaries. It is thus valuable for information professionals not only to understand the content of the document but also to learn lessons about how the Reference Model was developed and promulgated. A reference model is a very high-level, conceptual standard. Rather than providing detailed specifications of mandatory data elements or file formats, the OAIS has introduced a set of interrelated terms and concepts. Its development occurred during a period when many professional communities had a need for such a high-level standard but had not yet developed one themselves. There are various ways in which the Reference Model is able to provide detailed concepts, while still remaining relatively "implementation-independent": presentation of the functional model and information model in terms of high-level entities; extensive use of figures; and elaboration of implementation details in nonnormative appendices to the document, including several organization-specific "scenarios."

Much of the value of the Reference Model comes from its coherence, clarity, and synthesis of ideas. The Reference Model does define many new terms and introduces many original concepts, but its development has also been characterized by significant adaptation and reuse of preexisting sources.

Finally, the word "Open" in the acronym OAIS—meant to indicate that the standard was "developed in open forums"—was a defining feature of its evolution. The leaders of the effort presented draft products to many professional forums and actively recruited contributions and commentary. In addition to the traditional set of CCSDS stakeholders (space agencies and their contractors), many other organizations and individuals also contributed to the OAIS development process.

The OAIS is playing a major role in many current standardization, research, and development initiatives. Its impact on professional conversations surrounding digital preservation and digital curation continues to grow, 7 years after the Reference Model was approved as an International Standard.

ACKNOWLEDGMENTS

The author would like to thank Don Sawyer for sharing information about the OAIS five-year review process and the CCSDS for granting rights to use two of the figures from the OAIS.

REFERENCES

1. *Reference Model for an Open Archival Information System (OAIS)*; CCSDS 650.0-B-1; Consultative Committee for Space Data Systems: Washington, DC, 2002.
2. Lee, C.A. Defining digital preservation work: A case study of the development of the reference model for an open archival information system; Doctoral dissertation, University of Michigan, Michigan, 2005.
3. Waters, D.; Garrett, J. *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*; Commission on Preservation and Access and Research Libraries Group: Washington, DC, 1996.
4. Rothenberg, J. Ensuring the Longevity of Digital Documents. Sci. Am. **1995**, *272* (1), 42–47.
5. Sanders, T.; MacNeil, R. *Into the Future: On the Preservation of Knowledge in the Electronic Age*; Commission on Preservation and Access, and American Council of Learned Societies, 1997.
6. Procedures Manual for the Consultative Committee for Space Data Systems; CCSDS A00.0-Y-8; Consultative Committee for Space Data Systems: Washington, DC, November 2003; 1–4.
7. Procedures Manual for the Consultative Committee for Space Data Systems, CCSDS A00.0-Y-8; Consultative Committee for Space Data Systems: Washington, DC, November 2003; 1–4.
8. *Space Data and Information Transfer Systems—Open Archival Information System—Reference Model*; ISO 14721:2003; International Organization for Standardization: Geneva, Switzerland, 2003.
9. *Z39.50 Profile for Access to Digital Collections*; Library of Congress: Washington, DC, 1996.
10. *Planetary Data System Data Preparation Workbook*; Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, 1995.

11. *IEEE Guide to the POSIX Open System Environment (OSE)*; Institute of Electrical and Electronics Engineers: New York, NY, 1995.
12. Cargill, C.F. *Open Systems Standardization: A Business Approach*; Prentice Hall: Upper Saddle River, NJ, 1997; 90.
13. Waters, D.; Garrett, J. *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*; Commission on Preservation and Access and Research Libraries Group: Washington, DC, 1996, 12.
14. Lee, C.A. English-language literature citing or discussing the OAIS. In *Defining Digital Preservation Work: A Case Study of the Development of the Reference Model for an Open Archival Information System*; Doctoral dissertation, University of Michigan: Michigan, 2005; 227–253.
15. Greenstein, D.; Smith, A. Digital preservation in the United States: Survey of current research, practice, and common understandings. In *New-Model Scholarship: How Will It Survive?*; Smith, A., Ed.; Council on Library and Information Resources: Washington, DC, 2003; 40–48, 43.
16. Waters, D. Good archives make good scholars: Reflections on recent steps toward the archiving of digital information. In *The State of Digital Preservation: An International Perspective*; Council on Library and Information Resources: Washington, DC, 2002; 78–95, 80.
17. Hodge, G.M. Digital Preservation: Overview of Current Developments. Inform. Serv. Use **2002**, *22* (2/3), 73–82.
18. Beagrie, N. *National Digital Preservation Initiatives: An Overview of Developments in Australia, France, the Netherlands, and the United Kingdom and of Related International Activity*; Council on Library and Information Resources and the Library of Congress: Washington, DC, 2003, 45.
19. Lavoie, B.F. *The Open Archival Information System Reference Model: Introductory Guide*; OCLC Online Computer Library Center and Digital Preservation Coalition: Dublin, OH, 2004; 2.
20. Kenney, A.R.; McGovern, N.Y.; Entlich, R.; Kehoe, W.R.; Olsen, E.; Buckley, E.; DeMello, C. *Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Problems*; Cornell University Library: Ithaca, NY, 2005, http://www.library.cornell.edu/iris/tutorial/dpm/ (accessed November 14, 2007).
21. *Producer–Archive Interface Methodology Abstract Standard*; CCSDS 651.0-B-1; Consultative Committee for Space Data Systems: Washington, DC, 2004.
22. *Space Data and Information Transfer Systems—Producer–Archive Interface—Methodology Abstract Standard*; ISO 20652:2006; International Organization for Standardization: Geneva, Switzerland, 2006.
23. Procedures Manual for the Consultative Committee for Space Data Systems; CCSDS A00.0-Y-8; Consultative Committee for Space Data Systems: Washington, DC, November 2003; 1–3.