Open Archival Information System (OAIS): lights and shadows of a reference model

José Ramón Cruz-Mundet.
Carmen Díez-Carrera.

Article received on:
January 13, 2014.
Accepted on:
May 14, 2015.

Abstract

The Open Archive Information System (OAIS) reference model is an ISO standard originally developed by the Consultative Committee for Space Data Systems (CCSDS). The model serves to define the processes for effective, long-term preservation of information, while ensuring access to them. This model also provides a common language to describe these objects and has been widely accepted in the digital preservation community. Although it is not an application architecture, many experts have called into question its use. This paper examines the features of the model and the scientific controversy arising from

* Ambos autores pertenecen ala Universidad Carlos III de Madrid, España. (joseramoncruz@uc3m.es) ; (carmen.diez@uc3m.es)

its adaptability to diverse scenarios and wide-spread implementation.

Keywords: OAIS; Digital Preservation; ISO; Information Architecture; Electronic records; Standardization

Resumen

Sistema de Información de Archivo Abierto (OAIS): luces y sombras de un modelo de referencia
José Ramón Cruz-Mundet.y Carmen Díez-Carrera.

El Modelo de Referencia OAIS: Sistema de Información de Archivo Abierto es una norma ISO, desarrollada originalmente por el Consultative Committee for Space Data Systems (CCSDS), que define los procesos necesarios para preservar y acceder a los objetos de información de forma efectiva y a largo plazo, y establece un lenguaje común que los describe. Ha sido aceptado por buena parte de la comunidad implicada en la preservación digital, aunque no se trata de una arquitectura de aplicación, lo que ha llevado a numerosos expertos a poner en duda su utilización. El trabajo analiza las características del modelo y la controversia científica producida por su puesta en práctica y su capacidad de adaptarse a los diferentes escenarios.

Palabras clave: OAIS; Preservación Digital; ISO; Arquitectura de la información; Documentos electrónicos; Normalización

Introduction

OAIS is a trending topic in the professional environment; and although the search engine par excellence will ask us if we meant "oasis" when entering the acronym, results exceed by far 800,000 hits. Attached to the "conformity," hits change hardly at all, something difficult to understand as we address the matter: But is conformity to OAIS possible? To answer this

DOI: http://dx.doi.org/10.1016/j.libbai.2016.10.010

question preliminarily, let us clarify that this is a reference scheme on how a file information system for permanent digital preservation should be. If we move our focus from the professional literature to the information of the numerous preservation projects provided by software vendors, the impression does not change. This work attempts to account for this apparent misunderstanding. This study aims to describe, elucidate and analyze the model in the context of other emerging models, thereby contributing to an adequate understanding and knowledge of it. This endeavor is the result of a larger research effort focusing on permanent digital preservation and its economic sustainability, whose results shall be published in installments.

OAIS is a complex model mainly because its many nuances give rise to terms, mostly recent coinages, requiring effort, concentration and comprehension on the part of the reader and the help of a list of acronyms with their development. An OCLC paper (perhaps the first to address such matters), already has pointed this out.[1] The systematics are understandable, especially for an archivist and those who understand the functions of a document and archive management system. What we might call "small print," the detailed content of the standard; however, configures a text of enormous complexity and is elaborated fundamentally by and targeted at engineers and systems experts. In fact, the model has been developed by the ISO TC/20 specialized in the aerospace. As such, for most users, even those abreast of digital preservation, reading it might remind us of the contracting party scene in the Marx Brothers movies "A Night at the Opera."

The Open Archival Information System can suggest two ideas: open archival information system and open system of archival information. The first is the most frequently used and produced by the translators, and in any case the adjective (open) refers to the archive information system, not just to the archive. The model itself makes this clear, by stating that the term "open" is used to imply that it has been developed in open forums and does not mean that access to the file is unlimited (p.1-1, since we use the original text, references cites the original pagination, with the first digit indicating the section and the second, the page).

OAIS is a reference model that defines the processes necessary to preserve and access information objects effectively and in the long term. Moreover, it establishes a common language to describe such objects. As its name implies,

1 Meg Bellinger, "Understanding Digital Preservation: A Report from OCLC".

it is a only a model, and does not specify how things are done, but rather provides the framework for performing tasks successfully, while describing the basic functionalities and types of information required of the preservation environment.[2] OAIS identifies the mandatory duties as well as the paper and digital interactions negotiated among producers, consumers and document managers. It provides a standardized method for describing repository functionality, detailed models of information and archival functions.[3]

This model emerges as a result of the work of the Consultative Committee for Space Data Systems (CCSDS) – a standards body focused on earth and space data. Its development has embraced a broader scope and has gained acceptance across a wider range of stakeholders than the CCSDS.[4] The participants in the creation of OAIS tried apply it to a wide variety of repositories types. In this sense, OAIS became a widely adopted lingua franca for archival information systems, because it enables effective communication between national and international projects.[5]

The OAIS reference model represents a rare case of success in the history of the use of ICT methods. The model has enjoyed widespread great acceptance across a diverse range of professional communities, facilitating the conceptual and practical exchange of information among them.[6]

OAIS is omnipresent, almost a mantra, in the literature and theoretical and practical developments in the field of digital preservation. However, as evidenced by the use and interpretations of it in the Hispanic community, not everyone seems to fully grasp the reference model. As such we hope to provide a certain level of detail and help those interested grasp the true dimension and scope of the same, a task to which we will dedicate the epigraphs 2 to 7. We begin by describing the origins and the context in which it unfolds, and then proceed to analyze its object and scope, the foundational conceptual model, the model itself and further developments. The final section deals with the more controversial aspects of the model. We know that reader may want to know more about such matters. We have performed an exhaustive review of the

2 Raivo Ruusalepp et al., "Standards Alignement", 119 ss
3 Gregory S. Hunter, Custody of Digital Records: Options and Implementation Considerations.
4 Christopher A. Lee, "Open Archival Information System (OAIS) Reference Model".
5 Jens Klump, "Criteria for the Trustworthiness of Data Centres".
6 Achim Oßwald, "Das Referenzmodell. OAIS – Open Archival Information System. Einführung".

DOI: http://dx.doi.org/10.1016/j.libbai.2016.10.010

specialized literature up to September 2014, and very little it appears has been written on the subject.

Origins and context of the model

This practice is recommended by the Consultative Committee for Space Data Systems (CCSDS), founded in 1982 by leading aerospace agencies around the world as a multinational forum whose purpose is to study issues of space data systems and offer technical solutions in the form of recommendations for the development of data communication systems and spaceflight standards. In 1990 the CCSDS reached an agreement with ISO, allowing the recommendations of CCSDS, after a review and voting process, to become formal standards following.

At the request of ISO, work in this regard began in 1994 in an open, cooperative work environment described by Lavoie in his report for the Digital Preservation Coalition (2004), and even more exhaustively by Lee in his doctoral thesis devoted to the subject.[7] The *Encyclopedia of Library and Information Sciences*[8] also addresses this matter in more condensed entry. The Committee found that there was no generally accepted framework serving as the basis for policy development. For example, there are no shared concepts and terminology on digital preservation, basic functions of a digital archive system, and attributes of information objects to which preservation efforts can be aimed.

In the absence of a common framework, CCSDS began by developing a reference model defining the basic functional components of a system devoted to the permanent preservation of digital information, the internal and external key interfaces of the system, and characterization of information objects managed by the system, all of which are based on a set of well-defined terms and concepts that transcend the vocabularies of a specific domain and that should list the minimum requirements to be met by the archive system. The reference model needs to be a comprehensive and consistent framework for describing and analyzing digital preservation projects, while also providing a solid foundation and guiding principles for future policy developments.

7 Christopher A. Lee, Defining Digital Preservation Work: A Case Study of the Development of the Reference Model for an Open Archival Information System.

8 Christopher A. Lee, "Open Archival…".

The first draft for review was issued in May 1997, and in May 1999 the second draft was released. The latter was approved as a draft by ISO in 2000. After eight years of effort, the final version was released in January 2002[9] and the current version in June 2012; Both of these were approved as ISO 14721. The text of the standard can be downloaded from the CCSDS web page,[10] where it appears under the heading of Recommended Practices. These recommendations are descriptive in nature and are intended to provide general guidance on how to deal with a problem associated with support for space missions. The previous version came under the heading of Recommended Standards, which are prescriptive in nature and indicate how the infrastructure supporting space missions should operate and interoperate with one another.

Faced with the need to solve the problems associated with safeguarding digital or analog data and accustomed to solving complex problems and developing pioneering solutions, space agencies within the CCSDS decided to tackle the issue of digital preservation and the concept of electronic archiving. The solution comes in the form of a reference standard or framework for the long-term preservation of digital information, rather than any kind of application. Therefore, OAIS is a recommendation developed by and intended for a scientific population of engineers, physicists and computer scientists, for whom the challenge was not technological in nature, but rather conceptual, i.e., how to structure and integrate the processes of an electronic archiving system.

The OAIS text is organized into six sections: Introduction, Concepts, Responsibilities, Model, Perspectives of Preservation and Interoperability. The text also includes six appendices.

Purpose and scope

As defined by the document itself:

An OAIS is an archive consisting of an organization, which may form part

9 Brian Lavoie, The Open Archival Information System Reference Model: Introductory Guide, 3.

10 http://public.ccsds.org/publications/archive/650x0m2.pdf [Dated consulted: 16 de enero de 2015]

DOI: http://dx.doi.org/10.1016/j.libbai.2016.10.010

of a larger organization, persons and systems in charge of preserving information and making it accessible to a designated community.[11] (P.1-1)

This reference model:

- provides a framework for growing understanding and awareness of the archival concepts needed for the preservation and access of long-term digital information;
- provides the concepts necessary for non-archive organizations to be effective participants in preservation processes;
- provides a framework, including terminology and concepts, to describe and compare existing and future architectures and file operations;
- provides a framework for describing and comparing different long-term preservation strategies and techniques;
- provides the basis for comparing data models of the digital information preserved in archives and for discussing how data models and underlying information can change over time;
- provides a framework that can subsequently be expanded to cover the long-term preservation of non-digital information (e.g., physical media or samples);
- Expands consensus regarding components and processes for the preservation and long-term access of digital information, and promotes a larger market that vendors can support;
- guides the identification and production of standards related to OAIS.

The reference model addresses a wide range of digital information preservation functions including data entry, installation, data management, access and dissemination. It also deals with the migration of digital information to new media and forms, the data models used to represent information, the role of software in preserving information, and the exchange of digital information among archives. Identifies internal and external interfaces to file functions, and identifies a number of high-level services on those interfaces. The model provides several illustrative examples and an array of good practice recommendations, while defining a minimum set of requirements for an archive to be deemed an OAIS, and

11 Designated Community is defined in the model glossary as: "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. A designated community is defined by the archive and can change over time." The specific subset of information is what the model calls the knowledge base, defined in the model as: "A set of information, incorporated by a person or system that allows that person or system to understand received information."

> a maximum archive in order to provide a broad set of useful terms and concepts. (pp. 1-1 and 1-2)

Defined in its own terms, it is essentially a conceptual model of necessary knowledge for the task of preservation. It is a descriptive framework that allows of comparison present and future architectures, models, strategies, techniques and preservation operations; and includes all types of information in all types of supports and formats. The model promotes consensus regarding the permanent preservation and production of other standards that develop the model itself and complement it.

Since it is a model created by companies in response to their specific needs of preservation and long-term access to aerospace data. It can be applied to any repository, specifically in those organizations with responsibilities to keep the information accessible in the long term. It is also of interest to those organizations and individuals who produce information that must be preserved in the long term, as well as to those who need to obtain it from those archives.

Conceptual basis of OAIS

The model assumes that it is more difficult to preserve information digitally that committed to paper or film supports, because the associated technology is dogged by obsolescence which poses information loss hazards. The model treats this as an organizational, legal, industrial, scientific and cultural issue, not merely a technological one, warning that ignoring the problems posed by the preservation of information in digital form would lead inevitably to its loss.

The model empowers the archive, understood as a variety of functions and systems of storage and preservation, to achieve its objectives while overcoming the problem of the obsolescence of digital information and minimizing costs.

The purpose is to preserve the content information, and to do so in a way comprehensible to the designated community. As such the information must be represented in line with the knowledge base of that community. This entails harmonizing the access tools with the knowledge of the users, without losing sight of how this knowledge evolves.

DOI: http://dx.doi.org/10.1016/j.libbai.2016.10.010

Preservation is embodied in what is known as the information package, which brings together the content of the information and the description of the preservation description information, in order to ensure that the content is identified and the environment in which it was created can be understood. Both content and description are encapsulated, linked and identified by the packaging information and made accessible through descriptive information, consisting of a simple heading or a complete set of cataloging attributes. In sum, content is the object of preservation. The content, in turn, is accompanied by a series of metadata, thereby establishing the information package needed for its retrieval and use.

The information package comes in three variants representing as many states or versions of a document:

1. Submission Information Package (SIP), which is the original object transferred by the producer to the archive under the terms agreed by the two parties.
2. Archival Information Package (AIP) in which a SIP is transformed for preservation. It is an object stored in the archive.
3. Dissemination Information Package (DIP) is the object provided in response to a request from the user.

There are three distinct states of the document, or information: the information that enters, the information archived and the information disseminated.

The archive operates within and interacts with a three-part environment (see Figure 1). These parts are producer, management and consumer. The relationships with the producer are based on the submission agreement, which identifies the packages, the timing of shipments, and it materializes with the data submission sessions, in accord with the data models negotiated between the parties. Management refers to the fact that OAIS is part of a broader policy domain, which it constitutes, and for which it institutes governance, objectives and financing, where the designated community (consumers who are presumed to be capable of understanding preserved information) are the central factor. This includes issues such as reference, catalogues, searches, applications, etc. These three parts entail a series of mandatory responsibilities for OAIS (page 3-1), including:

1. Negotiates with the producer and accepts information.
2. Gain sufficient control over the information it archives.
3. Determine who makes up the designated community.
4. Ensure that the information is understandable in itself.
5. Follow established policies and procedures to ensure that it pre serves information against all reasonable contingencies.
6. Make information visible and available to its designated communities.

The model

The model has three components: the functional model, the information model and the information packaging transformations

·The functional model (pp. 4-1 to 4-20)

The functional model is repeated each time the subject is addressed, as shown in *Figure 1.* It is easy to understand, but this hides a certain complexity, since each of the entities comprising it has a specific functional development and its corresponding graphic representation:

As described above, producer, consumer and management constitute the environment in which an OAIS operates and interacts. The functional entities are:

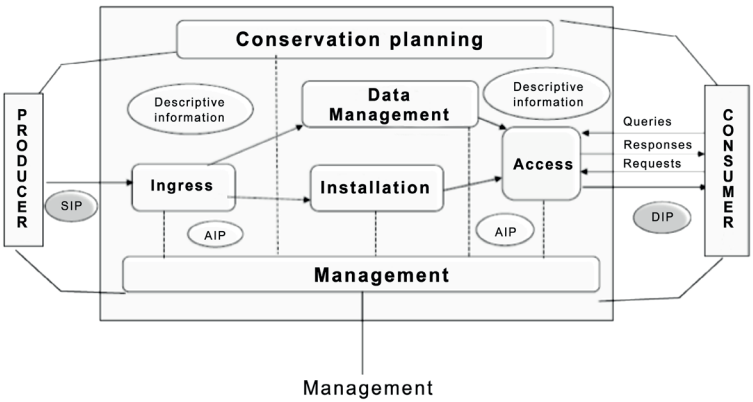

Figura 1. OAIS. Entidades funcionales. Fuente: autores a partir del original

DOI: http://dx.doi.org/10.1016/j.libbai.2016.10.010

- Ingest Functional Entity provides the services and functions to accept information transfer packages from producers (SIP) and prepare the contents for management and storage in the archive. For each SIP the input function implies a string that begins with reception, whose quality is assured, from which an information file package (AIP) is generated, described and transferred to the functional entity facility.
- The Archival Storage Functional Entity deals with the storage, maintenance and retrieval of the AIP, which entails receiving a storage request and an AIP, conveying it to permanent storage in the archive, managing the storage hierarchy, replacing the media as warranted, checking errors, duplicating contents and storing them in a separate installation for retrieval in the event of catastrophe, and providing data to the functional entity access to fulfill requests.
- Data Management Functional Entity provides the services and functions for completing, maintaining and accessing descriptive information, which identifies and documents the collection in an archive, as well as the administrative data used to manage the archive. It includes database managing and upkeep, receiving and answering queries, and generating reports.
- Administration Functional Entity deals with OAIS global operations. It includes negotiating transfer agreements, managing system configuration, updating archival information, controlling physical access, setting rules and policies, auditing transfers, activating requests and maintaining customer service.
- Preservation Planning Functional Entity monitors the OAIS environment, provides recommendations and preservation plans to ensure that stored information remains accessible and understandable for the designated community in the long term, even after the original computer environment were to become obsolete. The functions of this entity include targeted community and technology monitoring, development of preservation strategies and standards, packaging designs and migration plans.
- Access Functional Entity helps consumers determine the existence, description, location and accessibility of information stored in the OAIS, allowing them to request and receive information products. It includes coordinating access activities, generating information dissemination packages (DIPs), and responding to requests.

In addition to these functional entities, several Common Services (pp. 4-3 through 4-5) are presumably available that constitute another functional entity in the model. Briefly these services are as follows:

- Operating system services: These provide the core services needed to operate and manage the application platform and provide an interface between the application software and the platform.
- Network services: These provide the capacities and mechanisms to support distributed applications requiring data access, and the interoperability of applications in heterogeneous network environments.
- Security services: These provide capacities and mechanisms to protect sensitive information and associated treatment of data in the information system.

The information model ***(pp. 4-20 to 4-49)***

This model describes the types of information in an OAIS and defines the specific information objects that are used to preserve and access the information entrusted to the archive. It is designed to aid future OAIS system architects or designers, and are concepts not directly applicable in practice. Schematically, this service is structured in three parts as follows:

- Logical model for archival information: This model defines the types of information objects (the data with associated representation) required in an OAIS in order to enable long-term preservation of information and effective access to it by the designated community. Such objects include information content, packaging information, and description.
- Logical model of information in an OAIS: This model uses the descriptions of information objects to model the conceptual structures of information necessary to carry out the functions of an OAIS. It aims to highlight the relationship between the types of information needed in the archival process and comprises the three variants of information packages, which are the conceptual structure supporting long-term preservation: transfer, archiving and dissemination. The file information units, which are like information atoms that the archive must store; their descriptions and the collections of archival information and associated descriptions are entailed in this model.

DOI: http://dx.doi.org/10.1016/j.libbai.2016.10.010

- Management information: All information needed for the operations of an archive could be stored in databases as persistent data classes. Archive Administration Information represents the full range of information necessary for the day-to-day operation of the archive. Without entailing a comprehensive relationship, this information includes information policy, request tracking, security, subscriptions, user profiles, preservation history, statistics, and accounting information.

Transformations of the information package *(pp. 4-50 to 4-55)*

If the model has so far dealt with functional and architecture, the purpose of this subsection is the subsequent operation followed by the information package and its associated objects from the producer to the archive and from the archive to the consumer. This cycle breaks down as follows:

- Transformations of data in the producing entity. Producer data is private and may be in the format it wishes; however, when the decisions is made to install such data in an OAIS, the producer responsible for the data reaches enters into a transfer agreement with the archivists. This agreement defines issues such as the content, format and expected arrival times of the information transfer packages (SIP).
- Data transformations in the ingest functional area. Once the SIP are in the OAIS, their form and content can be changed, in fact, they become AIP and package descriptors that can be accepted and stored by the installation and data management functional entities, in a process of variable complexity. Moreover, the ingest functional entity will sort the incoming information objects and determine the collections to which they belong, while also updating the descriptions and providing additional access information. It shall coordinate updates and recovery of glitches occurring between data management and installation.
- Transformations of data in the functional areas of installation and data management. The installation functional entity takes the ingest AIP and places it in the permanent archive. The data management entity takes the descriptions and augments them.
- Data flows and transformations in the access functional area. When a consumer wishes to use OAIS data, a description instrument can be used to locate the information of interest. Once the

desired information is located, the consumer makes a request, which, if accepted, gives triggers the access response of providing copies in a temporary storage space.

Perspectives of preservation and interoperability

OAIS's long-term preservation objective faces several daunting challenges associated with and ever-changing computer industry, the short-lived nature of electronic storage media, problems with software itself, and changes in the knowledge base of the designated community.

Among the possibilities available, the proposed strategy is based on migration, because it improves profitability, facilitates adaptation to new technologies and prevents deterioration of the media. However, migration takes time, entails costs and exposes OAIS to a greater risk of information loss.

Without recommending any single one over the others, the model analyzes the pros and cons the diverse types of migration (refreshment, replication, repackaging and transformation). In terms of preserving stable access and use services, the model recognizes the importance of respecting users' familiarity with interfaces, something that is often ignored.

It contemplates both the possibility of an OAIS existing in more than one geographic location and cooperation among diverse archives. Therefore, it considers the points of view of the users wishing to have common services; that of the producers with respect to unified operations, the viewpoint of administrators who value uniformity and quality; as well as matters associated with the interests of archive, such as cost reduction, user satisfaction, and competitiveness.

The model analyzes four possibilities from the perspective of interoperability: absence of interoperability (independent archive), cooperation, federation and shared resources; as well as from the functional point of view, rather than from a technological point of view. As in other sections, it does not prejudge any choice.Further Developments

So far (February 2014) OAIS has led to the development of tools focusing on the ingest functional entity that help materialize the model and:

DOI: http://dx.doi.org/10.1016/j.ibbai.2016.10.010

- Producer-Archive Interface Methodology Abstract Standard (PAIMAS).[12] This is a CCSDS recommended standard that identifies and provides a framework for interactions taking place between an information producer and a archive. It covers the early stages of the OAIS-defined entry process, with the aim of providing a standardized method for formally defining digital information objects to be transferred from an information producer to an archive and for the effective packaging of those objects in the form of SIP, which supports the effective transfer and validation of the data.
- (XFDU) Structure and construction rules[13] is a CCSDS recommended standard for data and metadata packaging, including software, in an individual package (e.g., a file or a message) to facilitate the transfer and the information file. Another of its purposes is to offer detailed specification of the central packaging structures and mechanisms that comply with the current CCSDS recommendations.
- Producer Archive Interface Specification (PAIS) developed by the CCSDS is a recommended standard issued February 2014[14] and implemented by PAIMAS, for the purpose of providing a standardized method for modeling data to be transferred from an information producer to an archive and subsequent validation by the latter.

Moreover, previous CCSDS developments have been incorporated and adapted to the OAIS philosophy. These standards include the Data Description Language-EAST Specification, the Data Entity Dictionary Specification Language (DEDSL), the Audit and Certification of Trustworthy Digital Repositories (2011) [15] (ISO 16363: 2012), Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories (2014) [16] and Reference Architecture for Space Information Management (2013) [17] Other products have also adapted to the same philosophy, such as the PREMIS Data Dictionary for Preservation Metadata.

12 http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206511R1/Attachments/651x1r1.pdf [Date consulted: 16 de enero de 2015]
13 http://public.ccsds.org/publications/archive/661x0b1.pdf [Date consulted: January 16, 2015]
14 http://public.ccsds.org/publications/archive/651x1b1.pdf [Fecha de consulta: January 16, 2015]
15 http://public.ccsds.org/publications/archive/652x0m1.pdf [Date consulted: January 16, 2015]
16 http://public.ccsds.org/publications/archive/652x1m2.pdf [Date consulted: January 16, 2015]
17 http://public.ccsds.org/publications/archive/312x0g1.pdf [Date consulted: January 16, 2015]

INVESTIGACIÓN BIBLIOTECOLÓGICA, 30 (69), May/August, 2016, México, ISSN: 0187-358X, 227-253.

There are reasonable doubts about the degree of correlation between those who claim to have understood and implemented the model (from institutions dedicated to the preservation of the digital heritage to software manufacturers) and those who actually have done so, since it is not an application and the standards and tools to implement and measure performance are still in the pipeline. Moreover, many authors invoke it like a mantra, but few have pointed outs its limitations. Still some have tried to state what OAIS is and what it is not, without prejudicing its validity.

OAIS is a conceptualization of the environment, functional components and information objects associated with a system designed for effective, long-term preservation of digital materials.[18] OAIS is an important step towards standardization in the field of digital preservation, "including the development of criteria and procedures for analyzing and evaluating archival preservation and dissemination practices."[19]

The practicality of OAIS as a high-level model with which to frame the structural organization of a repository has been proven. The conceptual framework serves as an independent model of community and technology, defining the essential components of a repository, including the people and automated systems needed to manage long-term digital content and make it accessible to the user community. As an abstract model it confers significant flexibility to system designers in their repositories, allowing them to use in a relevant way regardless of the field of knowledge and content of the repository.[20]

Among the advantages of the model is that it has encouraged the participation of the digital preservation community[21] in the development of standards and application tools, whose common purpose is to find ways to implement an abstract model without forfeiting qualities of universality and standardization.[22] Early on the report *Trusted Digital Repository: Roles and*

18 Brian Lavoie, The Open Archival…, 14.
19 Neil Schumann y Andreas Recker, "De-mystifying OAIS compliance: Benefits and challenges of mapping the OAIS reference model to the GESIS Data Archive", 6.
20 Gareth Knight y Mark Hedges, "Modelling OAIS Compliance for Disaggregated Preservation Services", 63.
21 Especially archivists, librarians, conservationists, experts and companies in the technological and administrative public sector
22 Nancy McGovern, "Aligning National Approaches to Digital Preservation: an Introduction"

DOI: http://dx.doi.org/10.1016/j.libbai.2016.10.010

Responsibilities[23] sought to support implementation of the model by identifying prerequisites to be fulfilled by an organization. Both works define an integrative context for digital preservation, for the first time emphasizing the organizational and technological aspects associated with the management of digital preservation.

In 2003, OAIS working groups produced the Producer-Archive Interface - Methodology Abstract Standard (PAIMAS), which three years later would be earn approval under ISO 20652: 2006. This focuses on detailing the relationship between the producer transferring the digital content and the archive assuming the duty of preserving it, providing the workflow for negotiating and coordinating transfers.

Another important point of collaborative development has been in the area of metadata, when in 2005 *Working Group on Preservation Metadata: Implementation Strategies* (PREMIS)of the OCLC / RLG published the first version of the well-known Preservation Metadata Dictionary known by the acronym of the group.[24]

In 2007, as a result of the efforts of an international group working through the technical committee of the ISO TC20 / SC13, the report *Trustworthy Repositories Audit & Certification* (TRAC: *Criteria and Checklist* was produced. Other certifications in this vein include *Digital Repository Audit Method Based on Risk Assessment* (DRAMBORA), a tool developed by the *Digital Curation Centre and Digital Preservation Europe* (DPE), a methodology for digital repositories administrators, to perform self-assessment of risks in the preservation activity, and the German project Nestor.

As already mentioned, OAIS is not an application and consequently does not prescribe any single architecture, technology or database design; hence the difficulty in verifying the suitability of an archive with the model: it also accounts for the origin of most of the criticisms directed at the same, i.e., the impossibility of adapting to an abstract model. Brian F. Lavoie, an early and insightful analyst, pointed out the ambiguity of OAIS compliance: "Because the reference model is a conceptual framework, rather than a concrete implementation, the meaning 'pursuant to OAIS' is necessarily vague. Compliance with the reference model may imply an explicit application of OAIS concepts,

23 Research Libraries Group (RLG) y Online Computer Library Center (OCLC), Trusted Digital Repositories: Attributes and Responsibilities.

24 PREMIS Editorial Committee, PREMIS Data Dictionary for Preservation Metadata.

terminology, and functional and informational models in the course of developing the system architecture and the data model of a digital repository. It can also mean; however, that OAIS concepts and models are 'recoverable' from implementation. In other words from a high level perspective, it is possible to assign the diverse components in the archive system with the corresponding elements of the reference model. More ambiguity is introduced when institutions and organizations claim compliance with OAIS without defining or clarifying what that means with respect to their particular implementations.[25] A point of view with which practically all scholars have agreed from the beginning[26] which has turned into a sort of critical consensus about the use of the model as a commonplace, slogan or label; and we are referring to its use, not to the value of the model itself. In other words, complying with OAIS necessarily entails implementing a set of requirements that need to be translated, interpreted and fulfilled.

The issue of conformity is usually linked to institutional self-portraits of digital files and computer package descriptions. Despite the influence of the model (or perhaps because of it) and the ubiquity of its terminology and concepts, we often find misconceptions as to what OAIS is and what it is for. These misconceptions reveal fundamental misinterpretation of what is a reference model is.[27] This leads to the cloud of confusion sometimes surrounding the reference model. In our opinion, the fact that the model has been published as an ISO standard serves to expand its prestige, largely because we are accustomed to the ISO issuing prescriptive technical and organizational standards intended for application, but not abstract models. This was a mistake that could be righted only through the progressive development of a normative body oriented to practical application.

> The question inevitably arises to what extent is OAIS applicable? Of course in a literal sense, it is not directly applicable, since it is not a list of requirements that can be checked off as an archive system is constructed. It is, rather, a series of functions and conditions that must be fulfilled, regardless of how. As such, it is impossible to certify, because in that sense it lacks the features of other ISO standards. However, with regard to the mandatory responsibilities described in the

25 Brian Lavoie, The Open Archival…

26 Hilary Beedham et al., Assessment of UKDA and TNA Compliance with OAIS AND METS Standards.

27 Neil Schumann, Andreas Recker (2012). De-mystifying OAIS compliance: Benefits and challenges of mapping the OAIS reference model to the GESIS Data Archive. IASSIST Quarterly , p. 6.

DOI: http://dx.doi.org/10.1016/j.jlbbai.2016.10.010

> model (section 3-1), "it would be difficult for any working archive not to comply with them."[28] For all this, "to speak of implementation of OAIS is misleading. While details of this may seem objectionable, it is important to understand that the OAIS reference model is not transferable to the real world as such, and that this has an impact on the notion of OAIS compliance as posed in the model. As such, OAIS must be conveyed and interpreted by the archive or preservation service provider."[29]

However, it is not appropriate to generalize, since on the other pan of the scale we find abundant examples of proper understanding and judicious use of the reference model. Among these we can cite *Cultural, Artistic and Scientific Knowledge for Preservation and Access and Retrieval*)[30] (CASPAR) funded by the EU under the 6th Framework Program, which was based on the OAIS Model for packaging, access and security management, digital rights management and digital information access; the *Electronic Records Archives* (ERA),[31] whose philosophy tracks with the OAIS model; the SHERPA project, a disaggregated model to provide preservation services to small institutional repositories,[32] or the SHAMAN project (*Sustaining Heritage Access through Multivalent Archives*) for a system of long-term digital preservation in a grid environment.[33] These efforts do not indent or state full conformity.

For some authors, the OAIS model and some of its subsequent developments, such as DRAMBORA, work within --not so much outside of-- traditional systems. ISO 16363: 2012 *(Space data and information transfer systems -- Audit and certification of trustworthy digital repositories)* is based on OAIS and its revision has moved in parallel. The DRAMBORA risk assessment method provides a catalogue of typical risks inherent to digital preservation environments. Both were developed specifically for traditional digital preservation scenarios:

> Its focuses on providing a system to address the problem of digital preservation as a whole. This makes it difficult to apply in non-traditional digital preservation configurations. These models provide guidance on compliance criteria to be met, but do not provide effective governance and control mechanisms, or clear guidance on how to improve an organization's processes associated with digital

28 Hilary Beedham et al., Assessment of UKDA…, 10.
29 Neil Schumann y Andreas Recker, "De-mystifying OAIS…", 7.
30 http://www.casparpreserves.eu/index.html [Dated consulted: January 16, 2015]
31 http://www.archives.gov/records-mgmt/era/ [Dated consulted: January 16, 2015]
32 Gareth Knight y Mark Hedges, "Modelling OAIS Compliance…", 71.
33 Jorg Brunsmann, Long-term Preservation of Product Lifecycle Metadata in OAIS Archives.

> preservation. However, digital preservation is increasingly a concern in non-traditional settings, where the organization's environment may not be suitable for the use of an OAIS-based digital preservation system, but instead requires the incorporation of digital preservation skills with existing processes and capabilities associated with the organizational and technological system in place.[34]

A key issue in digital preservation is the fact that long-term conservation must be realized and guaranteed by technologies that are, by nature, short-term and changing. Additionally, it entails components of security and risk management that often are overlooked. Also, the complexity of long-term digital preservation increases by the fact that each activity type and organization has its own particularities and special requirements. This makes it a process that his highly dependent on the environment. Legacy institutions adopting the OAIS reference model have experience in the preservation of tangible objects for a certain number of years, but this experience may not be suitable for organizations faced with emerging digital preservation needs, such as industrial design or e-science companies, where the problem is to develop systems in which digital preservation is a relevant property and life cycle a key element. One way to understand the implications of the context of a digital object is to analyze life cycle. OAIS is limited to the file, which may be insufficient in terms of the additional information needed to preserve the object. "A broader concept of the life cycle of the object is needed, so that all the knowledge necessary for the reuse of objects in the future will be retained."[35]

In this sense, [36] OAIS provides only a very high level, narrow view of the main functions of a *Trustworthy Digital Repository (TDR).* On the other hand, it does prescribe a type of architectural solution that does not necessarily fit the technological horizon of an organization. Even when it does not stand in opposition, the description of a quasi-monolithic, separate system for digital preservation complicates the concept of incremental addition of capabilities and components to an existing system, such as an *Enterprise Content Management System* (ECMS), in order to facilitate preservation. In short, this is a question of scalability.

OAIS is difficult to reconcile where other systems are in operation with

34 Christoph Becker et al., "A Capability Model for Digital Preservation Analyzing Concerns, Drivers, Constraints, Capabilities and Maturities", 1. (Traducción propia.)

35 Gonçalo Antunes, José Barateiro y José Borbinha, "A reference architecture for digital preservation", 229 ss.

36 Christoph Becker et al., "Modeling Digital Preservation Capabilities in Enterprise Architecture".

DOI: http://dx.doi.org/10.1016/j.lbbai.2016.10.010

which it can be piggybacked to perform digital preservation tasks and processes. This can occur in organizations with electronic document management systems (EMS) and with key models such as Moreq2010, which covers a broad spectrum of aspects across hundreds of requirements.[37] MoReq2010 is an important catalogue of functional requirements for an electronic document management system (EMS), which covers aspects ranging from classification, audit trail, backup, recovery and security, reference, search and retrieval of information. "It is much more grounded in formal modeling than OAIS, but its hundreds of requirement statements make it too unwieldy and complex. Additionally, it not only covers the essential digital preservation capability of an ERMS, but also the full range of its functionality, providing exhaustive details on the desired operation of specific components of an ERMS."[38]

Although an ECM system is compatible with and can be compliant with OAIS, and these models are in fact complementary, these architectures differ in several essential aspects,[39] which are summarized in an integrated model in the organization (ECM), as opposed to a non-integrated model (OAIS):

- The data entry function in an ECM collects all the content that an organization produces, whereas the OAIS entry function needs to be provided with the information to be preserved.
- ECMs are usually embedded in the organizational infrastructure, while OAIS are often outside organizations that assume responsibility for preserving the information that other organizations have produced.
- Data capture (ECM) collects metadata on ownership, access rights, and other information needed for the active part of the document lifecycle. Ingest (OAIS) specializes in preservation-related metadata, such as, for example, file formats, representational metadata and preservation.
- In an OAIS, the descriptive information is kept separate from the actual data, representing the information to be preserved.
- ECM does not provide preservation planning (an aspect widely discussed by Becker et al.),[40] but rather a continuous functional, or controlled, logical preservation.

37 Christoph Becker et al., "A Capability Model…", 2.
38 Christoph Becker et al., "Modeling Digital…", 85.
39 Joachim Korb y Stephan Strodl, "Digital Preservation for Enterprise Content: A Gap-Analysis between ECM and OAIS"; Stephano Cavagliero, "Digital Archiving Systems Confronted with the OAIS Reference Model".
40 Christoph Becker et al., "Systematic planning for digital preservation: Evaluating potential strategies and building preservation plans".

OAIS does not provide specific implementation guidance; therefore, intermediate specifications and models are needed to build an OAIS-based system.[41] To accomplish this, some authors point to the need for a reference architecture that provides such practical guidance

> To establish the minimum, mandatory requirements for policies, processes and metadata to measure and validate the reliability of the repository in terms of authenticity, integrity, reproduction, meaning and retrieval of preserved digital materials ... it is not a matter of specifying a particular way of implementing OAIS: it is more about the general implementation guidelines needed if the term "OAIS compliant" is to be taken as a valid indicator that an archive's digital preservation system attains and maintains an adequate or improved degree of operational reliability, consistency and long-term compatibility, which, moreover, is measurable, verifiable, manageable and reasonably future-proof.[42]

There is also no shortage of voices questioning the application of OAIS in certain systems. Thus, in the Preserving Virtual Worlds (PVW) project, devoted to preservation of computer games and interactive fiction, part of the effort was focused on creating compliant OAIS archive information packages (AIP) and placing them in preservation repositories operated by Stanford University libraries and the central library of the University of Illinois at Champaign-Urbana. The greatest difficulties are with the concept of information representation, because it is not the same to address a designated community possessed of a compact knowledge base, such as the scientific community to which OAIS was originally targeted, than the broad, fluid, heterogeneous general public. The information on file formats or contextual information requires institutions to cooperate in order to avoid duplication of effort and enjoy savings. Additionally, it is an extraordinarily complex task to preserve games such as Spacewar! (1962) and Second Life (2003) which depend on long-gone systems and programs.

This experience has even led to reticence to adopt OAIS in the general library area, since it is a standard developed by the space data community, later adopted by the library community with little mind to the fact that scientific data files and research libraries have missions and modes of operation that are quite unique.[43]

41 Luigi Briguglio, Carlo Meghini y C. David Giaretta, "Best Practices for an OAIS Implementation"; Al Egger, "Shortcomings of the Reference Model for an Open Archival Information System (OAIS)".

42 Dennis Nicholson y Milena Dobreva, "Beyond OAIS: towards a reliable and consistent digital preservation implementation framework", 1.

43 James McDonough, "'Knee-Deep in the Data': Practical Problems in Applying the OAIS Reference Model to the Preservation of Computer Games", 1625.

DOI: http://dx.doi.org/10.1016/j.libbai.2016.10.010

Finally, economic sustainability, so often minimized in the analysis of digital preservation, is a determining factor for the future of OAIS. This is especially important to organizations operating in competitive environments that absorb costs of such efforts.[44]

Conclusions

In conclusion, OAIS is a reference model, a recommendation to be followed for the design and implementation of an archival system for long-term preservation. It is also a self-explanatory conceptual and terminological model, allowing communication between systems, designers and managers adopting the model. It is not a toolbox, nor is it an application, and much less simply technology: It involves aligning organizational policy and legal, industrial, scientific and cultural requirements.

This has been a decisive step in the standardization of digital preservation, a field with a remarkable degree of dispersion, and it embodies a success story in terms of the wide acceptance it has had in all areas.

It has exerted considerable influence, but it is not the only reference available for proposed digital preservation projects. The criteria catalogues for authentication of repositories specify the requirements that a repository must meet to be reliable. These criteria include levels of technical and organizational responsibility, which are often difficult to assess. Likewise, the archive community has developed criteria and models to support archival projects in their quest for security in the authenticity and provenance of their collections.[45]

Perhaps the future success of OAIS lies in keeping it flexibility, breadth and abstraction. These features have been key in its success so far, and in a remarkable way, earning it an undisputed reputation as a standard model. Moreover, the need to implement it has given rise to a body of standards and developments that enjoy general consensus, and which reinforce the construction of a standardization framework digital preservation requires. Additionally, it is compatible and provides room for the creation of sub-models (e.g., ECM) adapted to diverse preservation environments.

44 Annemieke de Jong, Beth Delaney y Daniel Steinmeier, OAIS Compliant Preservation Workflows in an AV Archive. A requirements Project.

45 Grehory Hunter, Custody of Digital Records…

Reference

Almost all references are available online. We have refrained from providing links since these often expire. The reader can access sources using a search engine.

Antunes, Gonzalo, José Barateiro y José Borbinha. "A reference architecture for digital preservation", en *Proceedings iPRES2010. 7th International Conference on Preservation of Digital Objects,* 229-237. Viena: iPRES, 2010.

Becker, Christoph, Hannes Kulovits, Mark Guttenbrunner, Stephan Strodl, Andreas Rauber, Hans Hofman. "Systematic planning for digital preservation: Evaluating potential strategies and building preservation plans", *International Journal on Digital Libraries* 10 (2009) (4):133-157.

Gonçalo Antunes, José Barateiro y Ricardo Vieira. "A Capability Model for Digital Preservation Analyzing Concerns, Drivers, Constraints, Capabilities and Maturities", en *Proceedings of the iPRES 2011 8th International Conference on Preservation of Digital Objects*. Singapur: Singapur iPRES, 2011.

Gonçalo Antunes, José Barateiro, Ricardo Vieira y José Borbinha. "Modeling Digital Preservation Capabilities in Enterprise Architecture". *The Proceedings of the 12th Annual International Conference on Digital Government Research,* 84-93. Maryland: University of Marylenad, 2011.

Beedham, Hilary, Julie Missen, Matt Palmer y Raivo Ruusalepp. *Assessment of UKDA and TNA Compliance with OAIS AND METS Standards*. London: JISC, UK Data Archive, University of Essex, 2005.

Bellinger, Meg. "Understanding Digital Preservation: A Report from OCLC", en *The State of Digital Preservation: An International Perspective,* 38-48. Washington: Council on Library and Information Resources, 2002.

Briguglio, Luigi, Carlo Meghini y David Giaretta. "Best Practices for an OAIS Implementation", *ERCIM News, special theme: Digital Preservation* 80 (2010): 36-37.

Brunsmann, Jorg. "Long-term Preservation of Product Lifecycle Metadata in OAIS Archives". Tesis de grado, Fakultat fur Mathematik und Informatik der FernUniversitat in Hagen, La Haya: FernUniversitat in Hagen, 2012.

Cavagliero, Stephano. "Digital Archiving Systems Confronted with the OAIS Reference Model", en *Proceedings of The Memory of the World in the Digital Age: Digitization and Preservation. An international conference on permanent access to digital documentary heritage,* 451-456. Vancouver: UNESCO, 2012.

CCSDS. *XML Formatted Data Unit (XFDU) Structure and Construction Rules. Recom-*

DOI: http://dx.doi.org/10.1016/j.lbbai.2016.10.010

mended Standard, issue 1. Blue Book. Washington: CCSDS, 2008.

Audit and Certification of Trustworthy Digital Repositories. Recommended Practice, issue 1. Magenta Book. Washington: CCSDS, 2001.

Producer-Archive Interface Methodology Abstract Standard (PAIMAS). Recommended Practice, issue 1. Red Book. Washington: CCSDS, 2012.

Reference Model for an Open Archival Information System (OAIS). Recommended Practice, issue 2. Magenta Book. Washington: CCSDS, 2012.

Reference Architecture for Space Information Management. Informational Report, issue 1. Green Book. Washington: CCSDS, 2013.

Producer Archive Interface Specification (PAIS). Recommended Standard, issue 1. Blue Book. Washington: CCSDS, 2014.

Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories. Recomended Practice, issue 2. Magenta Book. Washington: CCSDS, 2014.

CRL y OCLC. *Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist.* Dublin (OH.): CRL, OCLC, 2007.

Egger, Al. "Shortcomings of the Reference Model for an Open Archival Information System (OAIS)", *TCDL Bulletin Current* 2 (2006) (2).

Hunter, Gregory S. *Custody of Digital Records: Options and Implementation Considerations.* Prepared for the South Carolina Department of Archives and History. CRM Hunter Information Management Services, Inc., 2006.

ISO 14721:2003 Space data and information transfer systems -- Open archival information system -- Reference model. International Standards Organization. Ginebra: ISO, 2003.

ISO 20652:2006 Space data and information transfer systems -- Producer-archive interface -- Methodology abstract standard. International Standards Organization. Ginebra: ISO, 2006.

ISO 13527:2010 Space data and information transfer systems -- XML formatted data unit (XFDU) structure and construction rules. International Standards Organization. Ginebra: ISO, 2010.

ISO 16363:2012 Space data and information transfer systems -- Audit and certification of trustworthy digital repositories. International Standards Organization. Ginebra: ISO, 2012.

Jong, Annemieke de, Beth Delaney y Daniel Steinmeier. *OAIS Compliant Preservation Workflows in an AV Archive. A requirements Project.* Amsterdam: Netherlands Institute for Sound and Vision, 2013.

Klump, Jens. "Criteria for the Trustworthiness of Data Centres", *D-Lib Magazine* 17 (1/2).

Knight, Gareth y Mark Hedges. "Modelling OAIS Compliance for Disaggregated Preservation Services", *The International Journal of Digital Curation* 1 (2007) (2): 62-72.

Korb, Joachim y Stephan Strodl. "Digital Preservation for Enterprise Content: A Gap-Analysis Between ECM and OAIS", en *iPRES 2010 Proceedings of the 7th International Conference on Preservation of Digital Objects*, 221-228. Vienna, Österreichische Computer Gesellschaft, 2010.

Lavoie, Brian. *The Open Archival Information System Reference Model: Introductory Guide.* Dublin (OH.): OCLC Online Computer Library Center, Inc. and Digital Preservation Coalition, 2004.

Lee, Christopher A. "Defining Digital Preservation Work: A Case Study of the Development of the Reference Model for an Open Archival Information System". Tesis doctoral, Michigan: University of Michigan, 2005.

"Open Archival Information System (OAIS) Reference Model", en *Encyclopedia of Library and Information Sciences, Third Edition*, Marcia J. Bates y Mary Niles Maack (ed.), 4020-4030. Boca Raton, (FL.): CRC Press, 2009.

McDonough, J. "'Knee-Deep in the Data': Practical Problems in Applying the OAIS Reference Model to the Preservation of Computer Games", en *45th Hawaii International Conference on System Sciences,* 1625-1634. Maui (HI): IEEE Computer Society, 2012.

McGovern, Nancy. "Aligning National Approaches to Digital Preservation: an Introduction", en *Aligning National Approaches to Digital Preservation,* N. McGovern (Volume editor) y K. Skinner (Series editor), 5-16. Atlanta (GE): USA; Educopia Institute Publications, 2012.

Nicholson, Dennis y Milena Dobreva. "Beyond OAIS: towards a reliable and consistent digital preservation implementation framework", en *16th International Conference on Digital Signal Processing (DSP 2009)*, Santorini, Greece, julio 5-7, 2009.

Oßwald, Achim. "Das Referenzmodell. OAIS – Open Archival Information System", en *nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3.,* Heike Neuroth, Achim Oßwald, Regine Scheffel, Stefan Strathmann y K. Huth. Gotinga: nestor, 2010.

PREMIS Editorial Committee, *PREMIS Data Dictionary for Preservation Metadata, version 2.2.* Washington: Library of Congress, 2012.

Research Libraries Group (RLG) y Online Computer Library Center (OCLC). *Trusted Digital Repositories: Attributes and Responsibilities.* Mountain View (CA): RLG, 2002.

Ruusalepp, Raivo, C. A. Lee, B. van der Werf y M. Woollard. "Standards Alignment", en *Aligning National Approaches to Digital Preservation,* N. McGovern (Volume editor), 115-166. Atlanta (GE): Educopia Institute Publications, 2012.

Schumann, Neil y Andreas Recker. "De-mystifying OAIS compliance: Benefits and challenges of mapping the OAIS reference model to the GESIS Data Archive", *IASSIST Quarterly* (2012): 6-11.

Para citar este artículo:

Cruz-Mundet, José Ramón y Carmen Díez-Carrera. 2016. "Sistema de Información de Archivo Abierto (OAIS): luces y sombras de un modelo de referencia." *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información 70*: 227-253. http://dx.doi.org/10.1016/j.ibbai.2016.10.010

DOI: http://dx.doi.org/10.1016/j.ibbai.2016.10.010