



OAIS Five-year review
Recommendations for update
Compiled on behalf of members of
The Digital Curation Centre
and
The Digital Preservation Coalition

by
Sarah Higgins
Standards Advisor: Digital Curation Centre
and
Najla Semple
Executive Secretary: Digital Preservation Coalition

31/10/2006

Contributors:

Julie Allinson	UKOLN
Sheila Anderson	Arts and Humanities Data Service (AHDS)
Kevin Ashley	University of London Computer Centre (ULCC)
Simon Bains	National Library of Scotland (NLS)
Steve Bordwell	National Archives of Scotland (NAS)
Lyn Lewis Dafis	National Library of Wales (NLW)
Michael Day	UKOLN
David Giaretta	Digital Curation Centre (DCC)
Sarah Higgins	Digital Curation Centre (DCC)
Carol Jackson	Digital Preservation Coalition (DPC)
Gareth Knight	Arts and Humanities Data Service (AHDS)
Andrew McHugh	Digital Curation Centre (DCC)
Rory McLeod	British Library (BL)
Grayham Mount	National Archives of Scotland (NAS)
John Robertson	Centre for Digital Library Research (CDLR) / Centre For Educational Technology Interoperability Standards (CETIS)
Chris Rusbridge	Digital Curation Centre (DCC)
Derek Sargeant	University of Leeds
Najla Semple	Digital Preservation Coalition (DPC)
Susan Thomas	Oxford University Library Services (OULS)
Paul Wheatly	British Library (BL) / Preserv
Andrew Wilson	Arts and Humanities Data Service (AHDS)
Matthew Woollard	UK Data Archive (UKDA)

External reviewer:

Zoë Smyth	Public Record Office of Northern Ireland
-----------	--

General Recommendations:

Below is a compilation of comments gathered from a joint meeting of the Networks Associates Members of the Digital Curation Centre and members of the Digital Preservation Coalition on 13 October 2006. It was agreed that the standard is seen as important and valuable but needs updating.

1. There is a formatting issue throughout the document. There should be a clear distinction between examples and use cases, which should be separated from the normative text in order to create a better document structure.
2. The standard should use the agreed ISO terminology definitions for prescriptive words such as shall and should.
3. The standard should provide supplementary documents for full understanding, such as OAIS-lite for managers. A self-testing manual to establish bench marks would be useful along with a detailed checklist of the steps required for an implementation. Best practice guides, at a national level, through national standards bodies such as the British Standards Institution would be helpful.
4. Better, more concrete and more up-to-date examples are needed. The CD-Rom example and moon rock example are particularly unhelpful.
5. The standard needs to decide whether it is applicable to digital data or physical objects. Just now it tries, unsuccessfully to cover both which is both unhelpful and confusing.
6. The model appears quite definitive in many places, and not always true to its high level roots: "This reference model does not specify a design or an implementation. Actual implementations may group or break out functionality differently" (page 1-2, final paragraph). The model is sometimes too prescriptive and this could constrain implementers.
7. There needs to be some re-iteration that it is not necessary to implement everything and section 3.1 should clearly establish what the minimal requirements are.
8. Any review should look at terminology mapping between OAIS, PREMIS and other relevant standards and models as there are currently clashes.
9. The concept of the designated community should be extended to take account of the user community which extends beyond it. The knowledge base of the designated community, and the designated community itself, will change over time.
10. Interaction with internal and external systems and services: the OAIS seems to imply an 'insular' stand-alone archive but in reality it's likely to be part of a bigger organisation or network.
11. The model should deal with inevitable degradation of the OAIS in the long-term as it is unrealistic that the archive can produce 'perfect' versions over the long-term. The model should also include an exit strategy.

12. Section 3 could helpfully provide some additional information concerning the first stages of the ingest process, and explicitly reference the *Producer-archive interface -- Methodology abstract standard* (ISO 20652:2006) and *Information and documentation – Records management* (ISO 15489:2001). Any “producer-archive interface model” of early ingest would have to relate to ISO 15489.
13. Section 5 is overly specific about details and insufficient on concepts. Most of the discussion of migration should be removed; it is simply storage management, and belongs in the reliable storage layer.
14. Section 6 is over-specified for its purpose. In particular, it should list less about archive interoperability in the forms suggested here, and more interoperability of key services such as Representation Information registries, persistent identifier resolution, etc. The Risk impact of such approaches forms an important but missing part of this specification.

Chapter 1

Updates needed for clarification

Section 1.1 (page 1-1) bullet 6, beginning *Provides a foundation...* - This paragraph implies that there is nothing else available which will address the problem. This is not true as different communities have been doing similar things for years. The document should make clear that it builds on work by other communities, such as libraries and archives.

Section 1.4 (p 1-3) paragraph beginning *"A conformant OAIS archive may provide..."* - This statement should be in bold and moved to be paragraph 3 in the section, as the content of the paragraph is required by implementers prior to the information in paragraphs 2 and 5.

Section 1.7.2 (p1-7 to 1-13) - New definitions are required for the following terms:

Active archive	<i>eg as used in 2 (page 2-1) paragraph beginning "The explosion of computer processing power..."</i>
Access services	<i>eg as used in 5.2 (page 5-10) section header</i>
Authenticated	<i>eg as used in 3.1 (page 3-1) bullet 5</i>
Emulation	<i>see comments on Chapter 5 below</i>
IP templates	<i>eg as used in 4.1 (page 4-2) in paragraph beginning "Preservation Planning..."</i>
Integrity	as used throughout the document
Original	<i>eg as used in 3.1 (page 3-1) bullet 5</i>
Preservation	see Long Term Preservation in table below
Preservation Planning Functional Entity	<i>see Functional entity in list below. A possible definition is "The OAIS entity contains the services and functions for monitoring the environment of the OAIS and providing recommendations to ensure that the information remains accessible to the Designated Community over the long term."</i>
Repository	this term is increasingly used by the digital curation community. All RLG work uses this term and it would be helpful if the terms were consistent across the user community
Reasonable contingencies	<i>as used in 3.1 (page 3-1) bullet 5</i>
Staging area	<i>eg as used in 4.1.1.2 (page 4-5) in paragraph beginning "The Quality Assurance function..."</i>
Security	eg as used throughout the document
Traceable	<i>eg as used in 3.1 (page 3-1) bullet 5</i>
Understand	<i>eg as used in chapter 2.2.1 (page 2-3) paragraph beginning "A person or system..."</i> A clear definition of this is required in conjunction with the term "Knowledge Base" (see below)
User Community	eg the wider community who may use the OAIS but who are not the designated community

Clearer definitions are required for the following terms:

Functional entities	<p>The definition for the terms defining functional entities should spell out that these are Functional Entities so that they cannot be mixed up with verbs for other possible actions <i>eg</i></p> <ul style="list-style-type: none"> • Access Functional Entity NOT <i>Access</i> • Administration Functional Entity NOT <i>Administration</i> • Ingest Functional Entity NOT <i>Ingest</i> • Archival Storage Functional Entity NOT <i>Archival Storage</i> • Data Management Functional Entity NOT <i>Data Management</i>
Archival Information Unit	This definition needs to make clear whether the granularity of an AIU is defined as the smallest thing which a particular OAIS implementation would archive or whether it is the smallest thing which it is possible for an OAIS implementation to archive.
Ingest	This term needs a separate definition for the verb <i>to ingest</i> rather than just a definition for the Ingest Functional Entity
Data	The moon rock example is unhelpful as it mixes together digital data and physical objects. The difference between these needs to be clear.
Designated community	The definition needs to make clear that the designated community is a part of the entire possible user community. It also needs to highlight that the designated community may change over time. The characteristics of the designated community may change and the OAIS may need to preserve data for a new designated community (e.g. biologists emphasise changes from wet lab to IT & genetics, or a new designated community can come into play, e.g. company records designated community changes from internal management to external researchers).
Digital Object	This definition could be replaced with either A Sequence of Bit Sequences or Digital Entity . This removes any problems related to the object-oriented computing definition of a data object as a structured association of data and methods.
Edition	This definition is grammatically incorrect and out of step with the definition of “edition” in other standards - <i>eg</i> in FRBR (Functional requirements for bibliographic records) terms an Edition equates to an Expression; a Version or a Manifestation
Fixity Information	The use of the word <i>authenticity</i> would appear to be more applicable in this context than <i>authentication</i> ; <i>eg</i> : "The information which documents the authenticity mechanisms that ensure the Content Information Object has not been altered in an un-documented manner. An example is..."
Independently understandable	This definition is problematic as information will not remain independently understandable forever in the way described here. Decoding it will eventually depend on special resources not widely available, such as dictionaries and grammars of 21st Century English. The definition should imply that an OAIS will perform the maintenance required to ensure that a digital object which is currently independently understandable will remain so.

Information	The definition given here and the one in <i>chapter 2.2.1 (page 2-3) paragraph beginning "Information is defined..."</i> (see below) are not the same. They need to be consistent with each other and dictionary definitions. This definition does not work for a physical object.
Knowledge Base	This is a critical definition. If it is to make sense, then a clear definition of the meaning of both Knowledge Base AND (??) "understand" in this context has to be given. It should be made clear that the Knowledge Base is defined by the Designated Community and this will change over time. The example is false.
Long term	Should "user community" be replaced by "designated community" (or have the term designated community added to the definition)?
Long Term Preservation	(Page 1-11) paragraph beginning "Long Term Preservation: The act of..." – Replace the sentence: "The act of maintaining information in a correct and Independently Understandable form..." with: "The act of maintaining information in an Independently Understandable form...". Correctness cannot always be assured, but at least fixity and provenance should tell you what has happened! There is no mention of the Designated Community in this definition which is critical.
Preservation Description Information	The definition given does not adequately distinguish Preservation Description Information from Representation Information.
Physical Object	The definition should clearly distinguish between a physical object and a digital object.
Representation Information	A clearer definition which is consistent with the definition in chapter 4.2.1.3 is needed. It is not clear what "map" and "meaningful concept" mean. The definition should describe descriptive features and deal with data, software and hardware issues. The example doesn't explain the concept adequately and a better example should be found.

Updates to add missing concepts or strengthen weak concepts

Section 1.2 (page 1-2) paragraph beginning "This reference model does not specify a design or an implementation..." - This statement is contradicted by the sense of the document, which appears at a number of places to be quite specific about recommended implementation. The document should clearly separate the theory and the conceptual model from actual implementation.

Identification of any outdated material

Section 1.5 (p 1-4) - The list in this section needs updating and should point to relevant sections of appendix B. It should give a list of standards which it recommends and should point to work which builds on OAIS such as the PREMIS Preservation Metadata: Implementation Strategies Working Group, *PREMIS Data Dictionary* version 1.0 (OCLC and RLG, 2005) and the RLG, *An Audit Checklist for the Certification of Trusted Digital Repositories*, 2005.

Chapter 2

Updates needed for clarification

Chapter 2 needs to make clear that it is an introduction to the explicit detailed references to Chapter 4. Cross references should be introduced, or the chapters should be moved so that they are not divorced from each other.

Section 2.2.1 (page 2-3) paragraph beginning “Information is defined as any type...” - The definition of *Information* given here varies subtly from the terminology section, and is not clearly consistent with dictionary definitions. It is not true to say that information is “always expressed” as data, although this IS true in an exchange.

Section 2.2.1 (page 2-3) paragraph beginning “Similarly the information stored within...” - A conceptual model and diagram might be useful at this point.

Section 2.2.1 (page 2-4) paragraph beginning “As a practical matter, software...” - It should be made clear here that Representation Information should be supplemented by software, not replaced by it.

Section 2.2.3 (page 2-6) paragraph beginning “It is necessary to distinguish...” The sentence: “It is necessary to distinguish between an Information Package that is preserved by an OAIS and the Information Packages that are submitted to and disseminated from, an OAIS” should read “It is necessary to distinguish between an Information Package that is preserved by an OAIS and the Information Packages that are submitted to **or** disseminated from, an OAIS”

Section 2.3 (page 2-8) Figure 2-4 - The labelling of the bottom right arcs, currently labelled “queries, results sets” may be better understood if labelled “search, locate”?

Updates to add missing concepts or strengthen weak concepts

Section 2 Introduction (p2-1) paragraph beginning “The explosion of computer processing power...” - The concept of an *active archive* in this paragraph needs further development. The definition of the term is not clear and this should be made explicit in chapter 1.7.2. If *active archive* refers to the process of digital curation then it needs to be understood, as digital curation processes will affect both the functions and workflows which are detailed later in the document. It is not clear whether *active archive* relates to the management of digital objects after ingest into the OAIS or whether the term also embraces the curation activities engaged in by the Producer and Management prior to ingest into the OAIS. A diagram explaining how the *active archive* relates to the OAIS would be helpful at this point.

Section 2.2.1 (page 2-3) paragraph beginning “A person or system...” - A more focused definition, of *Knowledge base*, which is more appropriate to the digital environment, needs to be provided. This is a critical concept which informs what Representation Information an OAIS implementation actually requires. It needs to be made clear that an OAIS implementation should define the expected Knowledge base of the *Designated Community* and that the *Knowledge Base* of the *Designated Community* may change over time.

Section 2.2.1 (page 2-3) paragraph beginning “Similarly, the information is stored...” - Is it really necessary for an OAIS to collect all the required Representation Information to decode the

Knowledge Base? An OAIS would need a massive input to ensure that all required Representation Information is collected and described. The paragraph needs extending to make clear whether an OAIS actually needs to collect all the relevant Representation Information, or whether it can reference its existence elsewhere and be responsible for ensuring that it remains available. A more technical definition of Representation Information in chapter 1.7.2 and the inclusion of examples and case studies would be helpful to ensure that the concept is fully understood.

Section 2.2.1 (page 2-3) paragraph beginning “Similarly, the information is stored...” An additional sentence could possibly be added to this paragraph after sentence ending “...Representation Information is understandable using the recipient’s Knowledge Base”. The additional sentence is: "Depending on the nature of the content, the recipient may need access to a substantially deeper, more specific and "expert" Knowledge Base.

Section 2.2.3 (page 2-7) paragraph beginning “The Submission Information Package...” - A Submission Information Package (SIP) is not necessarily either structured or complete when it is received by the OAIS. It should be reflected in the text that a degree of work may be required to prepare a SIP for ingest. Indeed a SIP may be made up of a number of SIPs which have been prepared by different producers. For instance, SIPs may arrive with limited or even no metadata and have to have this enhanced or added. This paragraph should introduce the idea of these more complex scenarios to let implementers know the sort of work which may be required and the planning processes involved in preparing a SIP and converting between a SIP and AIP. Some reference to *Producer-archive interface -- Methodology abstract standard* (ISO 20652:2006) and *Information and Documentation – Records management* (ISO 15489(2001)) would be helpful here to ensure implementers know that additional help is available.

Identification of any outdated material

Section 2.2.2 (page 2-6) paragraph beginning “The Packaging Information is that information...” - The example of Packaging Information given in this paragraph is out of date and misleading. It ties the concept of Packaging Information to something which wouldn’t be part of a data object. There needs to be some separation between Content Information and Packaging Information. The definition would benefit from some better examples.

Chapter 3

Updates needed for clarification

Section 3 (page 3-1) paragraph beginning “This subsection establishes...” - Make clear that the section establishes **minimum** responsibilities.

Section 3 (page 3-1) bullet 3 – After “...should be able to understand the information provided.” Add: “Define and if possible agree the Knowledge Base with the Designated Community.”

Updates to add missing concepts or strengthen weak concepts

Chapter 3 - This chapter could helpfully provide some additional information concerning responsibilities at the first stages of the ingest process, and explicitly reference the *Producer-archive interface -- Methodology abstract standard* (ISO 20652:2006) and *Information and documentation – Records management* (ISO 15489:2001). The requirement in Section 3.1 (page 3-1) bullet 1 to:

“Negotiate for ...appropriate information” would be an early activity. It could also helpfully provide some information concerning other possible responsibilities beyond the minimum specified here.

Section 3.1 (page 3-1) bullet 1 - The requirement to negotiate for information should be dropped on the grounds that most OAISs are places of deposit in which clear decisions have to be made concerning what is actually accepted. The sentence should be changed from “The OAIS must: Negotiate for and accept appropriate information from information Producers” to “The OAIS must: Accept or reject appropriate information from information Producers”

Section 3 (page 3-1) bullet 5 - This responsibility raises legal issues concerning the authenticity of information and who authenticity needs to satisfy. This responsibility should articulate with *Information and documentation – Records management* (ISO 15489:2001) and specify a degree of authenticity which will satisfy the Records management standard. Risk management should be added to this responsibility (see comment on chapter 4.1.1.5 below).

Chapter 4.1

Updates needed for clarification

Section 4.1 - The figures should be relocated. Figures 4-8 and 4-9 should be moved to the positions of Figures 4-2 and 4-3. This will give an overview of full Functional Model of the OAIS at the start of the discussion. The remainder of the figures should then be renumbered.

Section 4.1 (page 4-2) paragraph beginning “Preservation Planning: ...” - The abbreviation IP should be expanded.

Section 4.1.1.2 (page 4-5) Figure 4-2 - The [updated SIP] should be removed to simplify this diagram.

Section 4.1.1.6 (Page 4-12 to 4-14) - This whole chapter should be abstracted. It is too specific and prescriptive and veers too much towards implementation. Some good illustrative examples or use cases would be helpful especially of migration plans.

Section 4.1.1.6 (Page 4-12 to 4-14) paragraph beginning “The migration goals...” - The discussion of Migration Plans in this paragraph should be cross referenced to Chapter 5.

Updates to add missing concepts or strengthen weak concepts

Section 4.1.1.2 (page 4-5) - Some of the tasks described in this section would appear to be an activity covered by the *Producer-archive interface -- Methodology abstract standard* (ISO 20652:2006) and *Information and documentation – Records management* (ISO 15489:2001) It would help to have a section which summarises activities which are part of these standards, and directs users to them for further details. This would include negotiating for material (see comment on chapter 3 above); ensuring that submitted material conforms to a collecting policy and accepting or rejecting material as part of the Quality Assurance process (currently included in the Audit Submission function, page 4-11); and activities required to turn a SIP into an AIP. (See comment on Chapter 2.2.3 above). A diagram which explains activities in the early stages of ingest should be added.

Section 4.1.1.2 (page 4-5) paragraph beginning “The Quality Assurance function...” - The idea of a 'staging area' as a place where SIPs are stored is an implementation issue. This level of detail is not needed at this point in the document.

Section 4.1.1.4 (Page 4-7) - This section needs to make clear that an OAIS deals with structured information which can be interrogated and not necessarily data in a traditional database format.

Section 4.1.1.5 (page 4-10 to page 4-12) - A risk assessment and management function should be included as a function of administration, including the requirement to do a periodic review which assesses both the technology and the community. The concept of a trusted digital repository should be introduced as part of this process. Relevant diagrams should be included.

Section 4.1.1.5 (page 4-10 to page 4-12) - A sentence should be included, in this section, which makes clear that some of the functions detailed in this chapter will not apply to all implementations; for instance in the customer service function, billing is not actually required, but this is the place you would do it if your implementation required it.

Section 4.1.1.5 (Page 4-12) paragraph beginning “The Customer Service function...” - Delete sentence beginning “It will collect *billing information*”. This function is too specific to one model of implementation and not relevant to a large number of OAISs.

Section 4.1.1.7 (page 4-15) paragraph beginning “The Coordinate Access Activities...” - The need for an OAIS to co-ordinate requests should be added to this section.

Identification of any outdated material

Section 4.1.1.1 (page 4-3 to 4-5) - Common services are easily trapped in time and it does not seem to be the function of this document to describe the computing environment. This whole section would be better:

- left out altogether
- replaced with a simple list with reference to relevant standards
- described in Service- oriented architecture (SOA) terms

If it remains then diagrams explaining the Common Services and how they integrate with the OAIS should be provided.

Note: the IEEE POSIX Reference Model is withdrawn; it may (or may not) be replaced by ISO/IEC 9945-1 etc Information technology — Portable Operating System Interface (POSIX®).

Section 4.1.1.3 (page 4-6 to page 4-8) - This whole section should be withdrawn. This document cannot give the subject the level of treatment it requires and there are specific standards which deal with the issues. In addition parts of the chapter are now out of date as new technology and methodologies have been developed. It would be better to make a general statement and point towards other relevant standards as possible means of achieving an implementation.

One possible relevant standards is *Information technology -- Security techniques -- Code of practice for information security management* (ISO 17799:2005).

A possible replacement for Section 4.1.1.3 is:

"The OAIS assumes a highly reliable, highly available, scalable (as required) error-free storage layer, in which AIPs can be placed and from which they can later be retrieved in identical form. There are many ways to achieve this, depending on the scale and other requirements. A number of ISO (or other?) standards are applicable in achieving this, including X, Y, *Information technology -- Security techniques -- Code of practice for information security management* (ISO 17799:2005).

Long term reliable bit storage is NOT the same as long term preservation as defined by OAIS, but is a necessary part of it.

It is important to realise however that, although good practice in this area has been widespread in well-managed IT infrastructure services for many years, factors such as scale (total object size, individual object sizes, and total numbers of objects) and particularly stringent requirements for fixity and security will place special demands. The ways these demands are best met will vary rapidly with contemporary computing practice, and are beyond the scope of this document."

Section 4.1.1.7 (page 4-15) paragraph beginning "The Coordinate Access Activities..." - The concept of a single user interface is dated in the modern context of OAI harvesting etc. which is outwith the OAIS's control. Some examples which show the access possibilities for distributed access environment such as multiplicity of users, machine-to-machine and external and internal users would be helpful.

Chapter 4.2 and 4.3

Updates needed for clarification

Section 4.2.1.3 (page 4-20) paragraph beginning: "Used to draw out additional meaning" - definition of additional meaning is unclear; use of physical object example is also confusing.

Section 4.2.1.3.2 (page 4-21) Figure 4-11 "Other Representation Information" - This label needs more definition and some examples. Does it only refer to supplementary semantic information or does it only relate to software? Clarify if database software is semantic information or other information.

Section 4.2.1.4 (page 4-23) Figure 4-12 - What do the dots to the far right of the tree imply? - This is unclear and needs either deleted or explained.

Section 4.2.1.4.1 (page 4-27) paragraph beginning "As another example consider an electronic file..." - The final example about word processing document requires updating. Changes in that proprietary formats can also be open standards.

Section 4.2.1.4.2 (page 4-27) paragraph beginning "In addition to Content Information..." - The first sentence beginning "In addition to content information..." should be removed. It is unclear how these four categories map onto implementation and data structures. The text needs to be upfront in saying that these categories are simply for consideration to avoid people trying to shoehorn data into them. This is not a specification for data models. The word 'understanding' in the sentence beginning "In addition to Content Information..." should be expanded.

Paragraph beginning “Provenance Information...” throughout the digital preservation process, changes will be made to the data object; the standard should mention the need for audit trails? Provenance is where changes should be documented and authentication stated. Examples would be useful, particularly of packaging information.

The concept of packaging, wrappers *etc* is quite straightforward, but examples could demonstrate how; the CD-ROM example isn’t helpful.

Section 4.2.2.2-4.2.2.8 - These sections represents a poor attempt to explain hierarchical bundles and granularity. This could be much clearer (see below for detailed comments).

Section 4.2.2.3 (page 4-37) Figure 4-18 – “Other representation information” is not present in this Figure, although it appears in Figure 4-11 (page 4-22). Is relationship of fixity information/ audit trail /provenance clear?

Section 4.2.2.3 (page 4-37) Figure 4-18 - This diagram could be redrawn so that the PDI levels are all on one level, otherwise it is confusing. ‘Other Representation Information’ should be added in here if it is kept elsewhere. There should be an arrow from Content Information to PDI.

Identification of any outdated material

Section 4.2.1.3.2 (page 4-23) paragraph beginning “Since Access software will incorporate...” – The sentence beginning “The practical use of emulation...” needs to be deleted. Such a recommendation is out of place and possibly out of date.

Section 4.2.1.4.2 (page 4-27) paragraph beginning “As another example, consider...” – Sentence beginning “In common practice, this viewing...” - a new use of ‘understanding’ is used and is unclear; thus far has ‘understanding’ only applied to reference information?

Updates to add missing concepts or strengthen weak concepts

Section 4.2.1.4.3 (page 4-29) - This section on Packaging Information would benefit from some examples; perhaps about how this happens.

Section 4.2.2.2 (page 4-31) paragraph beginning “The Submission Information Package...” - Does the definition of a SIP need to be much earlier in the document?

Section 4.2.2.4 (page 4-37) - There is far too much detail about AIUs and AICs, and this ends up more confusing than helpful. It would suffice to simply say that an AIU can encapsulate any content information (single files, collections of items, collections of collections, etc) however it cannot contain any AIPs in the content information. An AIC is designed to collect together AIPs and so may require a different set of PDI to describe the collection of AIPs. An AIC can be implemented in encapsulated form or referentially as the virtual concatenation of various elements.

Section 4.3 (page 4-48) - Examples of Packaging Information are needed.

Section 4.3.2 (page 4-50) Last bullet point – Previous definition of SIP was to create an AIP; this seems contradictory and the example is also confused. Does this refer to a situation where no new AIP is created? This section requires clarification.

Chapter 5

Updates to add missing concepts or strengthen weak concepts

Section 5.1 (page 5-1) - The chapter requires a more abstract overview and a parallel section on emulation. The standard needs to make clear that there are two fundamental approaches to information preservation: either the digital object remains in its original form, and some derivation of the original access and use mechanisms are made available (generally referred to as Emulation), or the object is transformed into one that can be processed with contemporary access and use mechanisms (generally referred to as Migration). There are variants of these, but that fundamental distinction needs to be made.

This chapter also requires some discussion concerning significant properties and performance which do not seem to be tackled in the standard.

Updates needed for clarification

Section 5.1.4 (page 5-7) - This section is not clear enough that the original AIP should be preserved?

Section 5.2 (page 5-10) - The heading “Access Service Preservation” is not clear enough and doesn’t define the abstract concepts very well. Is it intended to be ‘Functionality Preservation’?

Section 5.2.2 (page 5-10) - This section is over prescriptive. Again this refers to the high-level intention of the model and here there is too much reference to a particular approach.

Sections 5.1.3.1; 5.1.3.2 and 5.1.3.3 (page 5-6) - Use cases and examples of the three types of migration would be better.

Chapter 6

Updates to add missing concepts or strengthen weak concepts

The introduction to Chapter 6 (page 6-1) should mention something about the impact of archive interoperability on the definition of designated community. This chapter should be updated, particular with regard to modular approaches and shared services here and at other points in the model. It should also be made clear that OAIS isn't automatically interoperable with another OAIS.

Section 6.1.4 (page 6-7) - Should mention registries, services and identifiers. It should address internal and external interoperability (internal interoperability being the use of external services for archive operations).