

Response to the CCSDS's DAI Working Group's call for corrections to the OAIS Draft for Public Examination

Compiled on behalf of the members of the Digital Curation Centre and the Digital Preservation Coalition by William Kilbride and Sarah Higgins, 11 June 2009

1.0	11/06/2009	Document initiated by WK by compiling comments received
2.0	12/06/09	Distributed as draft to DPC members and participants in previous workshops
3.0	15/06/2009	Edited by SH and submitted to CCSDS

1. Introduction to these comments

This document represents a response to the DAI Working Group of the Consultative Committee for Space Data Systems (CCSDS) recent request for the identification of errors in *Reference Model for an Open Archival Information System (OAIS) Pink Book (CCSDS 650.0-P-1.1, May 2009)*, in preparation of its submission to the International Organization for Standardization (ISO). These corrections have been compiled after consultation with members of the Digital Preservation Coalition and the Networks Associates members of the Digital Curation Centre.

The **Digital Preservation Coalition** is a not-for profit membership organisation whose primary objective is to raise awareness of the importance of the preservation of digital material and the attendant strategic, cultural and technological issues. Its vision is to make our digital memory accessible tomorrow.

The **Digital Curation Centre** exists to provide a national focus for research and development into curation issues and to promote expertise and good practice, both national and international, for the management of all research outputs in digital format

This document follows two previous submissions to the Working Group on 31 October 2006 and 8 December 2008, which offered substantive responses to the OAIS review. In reviewing the draft standard we note and welcome the many instances where specific recommendations have been adopted and incorporated into the standard. However we further note that certain of the recommendations made have not been addressed. Although the comments presented here identify errors as requested by the Working Group, we remain concerned that the limited scope of the revisions will reduce the current influence of the standard.

2. Generalised Corrections

This section highlights errors and inconsistencies which are general to the text and which therefore require editing across the document. Comments have been restricted to the identification of clear errors.

1. There is an omission in relation to required language for normative/non-normative parts of the standard - specifically definitions of 'must', 'shall', 'should' and related terms. The following extract from the CCSDS Publications Manual exemplifies the omission from the Draft presented here:

3.4.1.7.2 Nomenclature. The nomenclature subsection shall identify linguistic usages that apply in the document. For standards track documents, boilerplate text shall be placed in this subsection:

"The following conventions apply throughout this Recommended Standard:

a) the words 'shall' and 'must' imply a binding and verifiable specification;

b) the word 'should' implies an optional, but desirable, specification;

c) the word 'may' implies an optional specification;

d) the words 'is', 'are', and 'will' imply statements of fact.

It is unclear whether these are the meanings intended when these terms are deployed within the text. Each use should be evaluated.

2. The standard seems inconsistent in its approach, purporting in the early sections to cover "information" in the general sense, but in reality only covering digital data in any meaningful way.
3. There is an inconsistency throughout the document regarding the use of long term, long-term, and Long Term (as defined in the terminology).
4. There is an inconsistency with the use of bold type for terms in the terminology. Sometimes they are in bold type first time they are used, sometimes not.
5. Titles of figures need to be updated to represent the change of title of the entities to which they relate. Hence: 4.2 Functions of the Ingest entity; 4.3 Functions of the Archival Storage entity, etc.
6. The term 'bit stream' is not spelled consistently.
7. The term compact disc, compact disk, CD and CD-ROM should be made consistent.

3. Specific Corrections

This section provides specific corrections of typographical, lexical and grammatical errors within the document.

1. Section 1.3, second from last paragraph has an additional full stop at the end of the last sentence.
2. Section 1.5, bulleted list item 'standard(s) for accreditation of archives. This is erroneous: archives are certificated: bodies providing certification are accredited. Should read: 'standard(s) for certification of archives'.
3. Section 1.7.2, definition of AIP edition use of 'improved or upgraded' is overly subjective, should be 'changed'.
4. Section 1.7.2, definition of API should be provided.
5. Section 1.7.2, definition of Designated Community uses the phrase 'change/evolve over time'. This seems sloppy and imprecise. Surely 'change or evolve over time'.
6. Section 1.7.2, oversight: no definition of 'Functional Entity' is provided.
7. Section 1.7.2, definition of Information Package, last line. Capitalise 'Information, hence: '...Package Description Information used to ...'
8. Section 1.7.2, definition of Information Property is unclear. Are there missing or misused terms here?

9. Section 1.7.2, definition of Information Property Description is unclear. Are there missing or misused terms here?
10. Section 1.7.2, definition of Reversible Transformation is unclear. This uses the term representation which is not defined. The meaning here seems inconsistent with the usage elsewhere.
11. Section 1.7.2, definition of Succession Planning. This uses the term 'repository' which a term which has a variety of meanings to different communities. It should be replaced with 'OAIS'.
12. Section 2.2.1, paragraph beginning 'Similarly, the information...' second sentence contains a stray word 'East'. Should read ...'measured in degrees of latitude and longitude'.
13. Section 2.2.1, paragraph beginning 'In order for this ...' first sentence contains a split infinitive. Should read ... 'to identify clearly and to understand clearly' (not clear which verb the adverb qualifies).
14. Section 2.2.1, paragraph starting 'The OAIS reference model ...' second sentence. The word 'evolves' is not appropriate here: development of technology is not a natural process. Should read 'As digital technology develops...'
15. Section 2.2.2 bulleted list item 'Reference'. The phrase 'ISBN number' is tautological: should read 'Examples include an ISBN for a book, or ...'
16. Section 2.2.2, bulleted list item 'Fixity'. The phrase 'check sum' is incorrect: it should be 'checksum' as one word.
17. Section 2.2.2, bulleted list item 'Access Rights'. The Verb 'provides' should be pluralised. Should read 'Access Rights provide the terms of access.'
18. Section 2.2.2, paragraph starting 'The packaging information ...', last sentence. The acronym XFDU should be expanded and glossed.
19. Section 2.2.3, paragraph starting 'It is necessary...', second sentence. Stray word 'final'. Should read 'to meet OAIS preservation requirements'.
20. Section 2.2.3, paragraph starting 'The Submission Information ... first sentence. 'see related standards in section 1.5' Need to specify which standards.
21. Section 3.1, bulleted list item 'Determine, either by ...'. The phrase 'thereby defining its knowledge base' is erroneous. It is not the function of the OAIS to define the knowledge base of the designated community, merely to track and act accordingly. Consequently the phrase should be replaced with 'using elements of its knowledge base'.
22. Section 3.1, bulleted list item 'Follow documented policies ...'. Erroneous deletion of 'authoritative' within description of data.
23. Section 3.1, bulleted list item 'Follow documented policies ...'. The final parenthesis is a sentence in its own right and should finish with a full stop. Hence: '...approved strategy. There should be no ad-hoc deletions'.
24. Section 3.1, bulleted list item 'Make the preserved ...' delete the comma after 'Objects'

25. Section 3.2.1, paragraph that starts 'An organisation operating ...', first sentence. The verb 'aid' should be plural. Should read '...criteria that aid in determining ...'
26. Section 3.2.1, paragraph that starts 'An organisation operating ...', last sentence. the reference to 'moon rock' is surprising here. Should it be simply 'physical objects' as defined in 1.7.2.
27. Section 3.2.1, paragraph starting 'Negotiations can embrace ...' First sentence: make negotiation singular for better fit with sentence.
28. Section 3.2.3 heading conflates two concepts - ought it be entitled 'determines designated community' or 'determines consumer' The former is more in keeping with the sense of the text.
29. Section 3.2.3, paragraph that starts 'The possible evolution ...' the noun 'evolution' is not appropriate to what is being described. Should read 'Possible changes in the definition of the ...'
30. Section 3.2.4, paragraph beginning 'As another example ...' the final sentence. This has become muddled. Final sentence should read ... 'and should also logically be provided in the Preservation Description Information.'
31. Section 3.2.4, paragraph that starts 'Software is needed ...', last sentence. Delete the comma between 'ongoing' and 'testing', hence: '...a vigorous, ongoing testing and validation programme'.
32. Section 3.2.5, paragraph that starts 'It is essential ...', Full stop required and comma to be deleted in second line at point of insertion, and full stop required at the end of inserted sentence.
33. Section 3.2.5, paragraph that starts 'It is essential ...' inserted sentence. Delete 'strategy' and replace with 'policy'. Hence 'as part of an approved policy'.
34. Section 4.1, paragraph beginning 'In addition to...', last sentence. Line break to be removed between last line of sentence and full stop.
35. Section 4.1.1.4, paragraph beginning 'The administer database ...', first sentence. Delete unnecessary spaces around the phrase 'for storing'.
36. Section 4.1.1.5, paragraph beginning 'The manage system ...', first sentence. A split infinitive in line two. Should read 'archive system to monitor continuously the ...'
37. Section 4.1.1.5, paragraph beginning 'The Establish standards ...', last sentence has two full stops.
38. Section 4.1.1.5, paragraph beginning 'The Audit Submission ...' second sentence has two full stops.
39. Section 4.1.1.6, paragraph beginning 'The develop preservation ...', first sentence has an extra comma between 'risks' and 'to'.
40. Section 4.1.1.6, paragraph beginning 'The develop packaging ...', first and second sentences use the abbreviation 'IP'. This should be expanded to avoid confusion with 'Intellectual property' a more commonly understood usage.

41. Section 4.2, paragraph beginning 'As discussed in ...', second last sentence. Capitalise 'A' in 'Annex'.
42. Section 4.2.1.3.1, paragraph beginning 'The representational information ...', last sentence has two full stops.
43. Section 4.2.1.3.1, paragraph beginning 'This figure also ...', first sentence is an extended run on sentence which is hard to follow. Must be broken up into meaningful chunks.
44. Section 4.2.1.3.1, paragraph beginning 'Structure, Semantic and ...' this could be more elegant; Should read: 'Structure Information, Semantic Information and Other Representation Information are both sub-types and components of Representation Information'.
45. Section 4.2.1.3.2, paragraph beginning 'To preserve the ...', last sentence. Split infinitive: should be: 'to convey fully'.
46. Section 4.3.2, paragraph beginning 'Once the SIP ...', first sentence. The pronoun 'their' should be singular, hence 'Once the SIP is within the OASIS, its form and content ...'
47. Section 5.1.3.4, paragraph beginning 'Digital migrations that ...' last sentence mistaken renders 'AIP' and 'API'.
48. Section 6, last paragraph, last sentence. There's an unclear reference to a section currently marked as '0' which needs to be connected to a section
49. Section 6.1, paragraph beginning 'In general one ...', line two, verb is malformed in clause '...from some level of interoperability to be desirable..'. Delete 'to ' and insert 'may'. Should read 'some level of interoperability may be desirable...'
50. Section 6.1.4, paragraph starting 'Figure6-4', first sentence contains a MS word internal reference error.