

ISO/IEC JTC 1/SC 34/JWG 7

Joint JTC 1/SC 34-TC 46/SC 4-IEC/TC 100/TA 10 WG: EPUB

Convenorship: KATS (Korea, Republic of)

Document type:	Disposition of Comments Report
Title:	Disposition of comments for SC34 N2483 and N2485 Summary of Voting on ISO/IEC PDTS 22424-2 EPUB 3 Preservation - Part 2: Metadata requirements
Status:	This document is a disposition of comments for mainly two summaries of voting came from JTC 1/SC 34 and ISO TC 46/SC 4. All comments are combined into single document for efficient resolution process.
Date of document:	2018-05-21
Source:	Project editor (Juha Hakala)
Expected action:	INFO
No. of pages:	8
Email of convenor:	samoh21@gmail.com , zzosang@gmail.com
Committee URL:	https://isotc.iso.org/livelink/livelink/open/jtc1sc34jwg7

0	1	2	3	4	5	6	7	8	9
#	NB ¹	Line number (e.g. 17)	Clause/ Subclause (e.g. 3.1)	Paragraph/ Figure/ Table/ (e.g. Table 1)	Type of commen ²	Comment (justification for change) by the NB	Proposed change by the NB	Proposed Disposition of Comments	Disposition of Comments
1	GB 01 (SC34)				Ge	The scope of the Draft Technical Specification is insufficiently clear in the drafts as presented. The scope statements are insufficiently aligned with the contents of each document.		Note: resolver per GB01 at the CRM of 22424-1	Note: resolver per GB01 at the CRM of 22424-1
2	GB 02 (SC34)				Ge	The drafts as presented do not make a clear distinction between EPUB profile requirements and packaging needs.	The Technical Specification should provide separate profiles for EPUB and METS, preferably in separate Parts.	Note: resolver per GB02 at the CRM of 22424-1	Note: resolver per GB02 at the CRM of 22424-1
3	GB 03 (SC34)				Ge	The Draft Technical Specification contains too much discussion of general digital preservation issues.	The text should focus on the issues and challenges of EPUB preservation.	Note: resolver per GB03 at the CRM of 22424-1	Note: resolver per GB03 at the CRM of 22424-1
4	GB 04 (SC34)				Ge	The Draft Technical Specification relies too heavily on theoretical arguments that do not always align well with the experience of practitioners of EPUB preservation. As an example of this, the text makes false assumptions about who is responsible for preparing METS SIP packages: this is not always the producer of the e-book, but frequently the preserving institution.		Note: resolver per GB04 at the CRM of 22424-1	Note: resolver per GB04 at the CRM of 22424-1
5	GB 05 (SC34)				Ge	The proposed approach to the encoding of “unarchivable” data is probably out of scope, as it depends upon institutional policies.		Note: resolver per GB05 at the CRM of 22424-1	Note: resolver per GB05 at the CRM of 22424-1
6	GB 06 (SC34)				Ge	The attribution of issues documented in Annex A of Part 1 is unclear: material that originally appeared in a British Library report is interspersed with comments from other sources.	All sources should be clearly attributed.	Note: resolver per GB06 at the CRM of 22424-1	Note: resolver per GB06 at the CRM of 22424-1
7	GB 07 (SC34)				Ge	The text needs significant editorial work to remove repetitions, unhelpful elaborations, broken links,		Note: resolver per GB07 at the CRM of	Note: resolver per GB07 at the CRM of

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)

2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

0	1	2	3	4	5	6	7	8	9
#	NB ¹	Line number (e.g. 17)	Clause/ Subclause (e.g. 3.1)	Paragraph/ Figure/ Table/ (e.g. Table 1)	Type of commen ²	Comment (justification for change) by the NB	Proposed change by the NB	Proposed Disposition of Comments	Disposition of Comments
						ungrammatical phrases etc.		22424-1	22424-1
8	CA 01 (TC46)		02	Normative References	Ge	Consider whether other standards and specification should be listed as normative references. In several places in this specification, elements from these standards are described as being obligatory. It is necessary to give credit in this specification when concepts, practices and schemae have been copied from, adapted from those standards. It is also necessary to state which versions of those standards are being used. ISO 14721 A significant portion of this technical specification arises from the OAIS standard ISO 14721:2012. METS Metadata Encoding Technical Standard PREMIS Dublin Core, as manifested in ISO 15836.	Add ISO 14721:2012, METS, PREMIS and Dublin Core as manifested in ISO 15836 as normative references. State the versions of METS, PREMIS and Dublin Core that are used..	OK.	Accepted. As proposed.
9	CA 02 (TC46)		03	Terms and Definitions	Ge	It is recognized that terms and definitions from ISO 14721 have not been added to the ISO OBP. This makes it necessary to add citations to each specific term and definition. This assures the reader that the working group was working with the current definitions from the current version of 14721.	Add a citation to ISO 14721:2012 to all of the terms and definitions copied from the published standard. Doing this will facilitate harmonization efforts within the ISO family in the future.	Note: resolver per CA02 at the CRM of 22424-1	Note: resolver per CA02 at the CRM of 22424-1
10	US 03 (TC46)		06.04		Ed	It is unlikely that most publishers can/will abide by requirement that an ISBN is not used as a ackage identifier. While we agree with the remise of the argument in this document, in truth f implementation the ISBN is and will continue to be used.	Recommend the use of ISBN	Usage of ISBN as package identifier can cause problems and therefore it cannot be recommended. But in the revised Part 2 ISBN may be used as package identifier if it is certain that	Acknowledged. No action. <i>Note: see proposed disposition of comments.</i>

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

0	1	2	3	4	5	6	7	8	9
#	NB ¹	Line number (e.g. 17)	Clause/ Subclause (e.g. 3.1)	Paragraph/ Figure/ Table/ (e.g. Table 1)	Type of commen ²	Comment (justification for change) by the NB	Proposed change by the NB	Proposed Disposition of Comments	Disposition of Comments
								this does not cause problems.	
11	US 04 (TC46)		06.07, pg 18		Ge	In most preservation systems, original MD records are preserved when updates are sent.		OK. The specification contains this text now, and says that new metadata records shall get new identifiers.	Accepted. As amended. <i>Note: see proposed disposition of comments.</i>
12	US 05 (TC46)		06.08		Ge	While not against all those dates, the most important date to be sent is the traditional “publication date” of the item to be preserved. MD creation date, EPUB creation date – they are all interesting but not the most important.	Ensure publication date is preserved.	OK.	Accepted. As amended.
13	US 06 (TC46)		06.09		Ge	We strongly recommend that publication date also be a required MD element.	Specify the publication date be a required element.	See above.	Resolved per US05.
14	US 07 (TC46)		07		Te	We recommend that if the EPUB3 publication contains one or more links to external MD records, that those always be downloaded and included in the SIP or as part of the AIP ingestion process, rather than this being optional (the problem is, how is the archive to know, definitively, which MD is correct and most robust without downloading it?)	Specify that one or more links to external MD records be downloaded and included in the SIP.	The specification requires that any metadata or media file which is external in SIP is harvested during ingest and included in AIP.	Accepted. As amended. <i>Note: see proposed disposition of comments.</i>
15	CA 04 (TC46)		07.01.1		Ed	The citation to footnote 23 was not given in superscript.	Give "23" in superscript.	Multiple footnotes has lost their superscript status during the editing of the file. The superscript status will be restored before the TS is published.	Accepted. As proposed.
16	US 09		07.03		Te	There is no stipulation that the METS and EPUB structural container	Make a stipulation that METS and EPUB structural container descriptions not	OK.	Accepted. As

1 MB = Member body / NC = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)

2 Type of comment: ge = general te = technical ed = editorial

0	1	2	3	4	5	6	7	8	9
#	NB ¹	Line number (e.g. 17)	Clause/ Subclause (e.g. 3.1)	Paragraph/ Figure/ Table/ (e.g. Table 1)	Type of commen ²	Comment (justification for change) by the NB	Proposed change by the NB	Proposed Disposition of Comments	Disposition of Comments
	(TC46)					description not contradict one another (should they both have them), should this be addressed?	contradict.		amended.
17	US 10 (TC46)		07.04		Ge	"If the original version of the document is also archived, preservation metadata SHOULD specify the differences between the two renderings of the document. This metadata can be useful when a customer is deciding which version of the publication would serve his needs better." --> The above paragraph presumes the archive knows far more about the content and structure of the preserved content and how different versions relate to one another than is plausible. There are two primary reasons why one might have multiple preservation copies of an object: 1) the creator sent it to the archive multiple times -- the first time when created and then again when it is updated, and 2) the archive might migrate content to a new format or have improved tools for extracting descriptive MD. In either situation, there is no reason to expect the archive to have enough information to specify the *differences* between the versions -- outside of typical preservation MD facts (used these tools on this date for content received on that date, etc.).		This section discusses SIP, so it does not assume anything about the archive. The original wording did not make this sufficiently clear. New version of the text: If the original (unarchivable) version of the publication is also submitted to the OAIS archive, producer-generated preservation metadata in the SIP SHOULD specify the differences between the two renderings of the publication. Such metadata is useful when a customer is deciding which version of the publication would serve his needs better. If and when the publication is migrated during preservation, similar metadata about quirks in the format migration SHOULD	Accepted. As amended. <i>Note: see proposed disposition of comments.</i>

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)

2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

0	1	2	3	4	5	6	7	8	9
#	NB ¹	Line number (e.g. 17)	Clause/ Subclause (e.g. 3.1)	Paragraph/ Figure/ Table/ (e.g. Table 1)	Type of commen ²	Comment (justification for change) by the NB	Proposed change by the NB	Proposed Disposition of Comments	Disposition of Comments
								also be created and stored in the new AIP.	
18	US 11 (TC46)		07.04		Ge	validation --> it is a good ideal to expect/require producers to validate content and in 10 years of working for a preservation service, I have rarely seen it happen. In our experience, a preservation service should expect to receive invalid content and have a robust methodology for dealing with such either during the ingest process or after the content has been preserved.		Since producers rarely validate their publications, this requirement has been changed from SHOULD to MAY	Accepted. As amended. <i>Note: see proposed disposition of comments.</i>
19	US 12 (TC46)		08		Ge	"A SIP SHALL contain a manifest file specifying the content of the package." --> This is an optimistic view of the current capability of many producers/publishers. Much of the content Portico receives does not come with a manifest. Perhaps working with EPUB will force publishers to create a manifest?		Since producers rarely provide this information the requirement has been changed to SHOULD	Accepted. As amended. <i>Note: see proposed disposition of comments.</i>
20	US 13 (TC46)		9, pg 35		Te	Checksums - Portico receives a tremendous amount of content every day and very little of it is provided to us with checksums included. I would recommend revising this to be a preference, not a requirement, else wise there may be too many publishers who simply won't do anything because they cannot comply with the rigor of the guidelines.	Change Checksums to be a preference and not a requirement.	This requirement, although essential, was changed to SHOULD. A footnote was added that the producers should have tools with which they can easily create SIPs which meet the archiving requirements.	Accepted. As amended. <i>Note: see proposed disposition of comments.</i>
21	US 14 (TC46)		Introduction		Ge	Within the preservation community, emulation is considered to be a viable option for some content		The relevant section in the Introduction was modified to make it more	Accepted. As amended. <i>Note: see proposed</i>

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)

2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

0	1	2	3	4	5	6	7	8	9
#	NB ¹	Line number (e.g. 17)	Clause/ Subclause (e.g. 3.1)	Paragraph/ Figure/ Table/ (e.g. Table 1)	Type of comment ²	Comment (justification for change) by the NB	Proposed change by the NB	Proposed Disposition of Comments	Disposition of Comments
								<p>positive towards emulation:</p> <p>Migration does not always produce a satisfactory digital copy of the original document. Sometimes it cannot be applied at all; programs cannot be migrated without access to and good understanding of the source code. In such cases long term preservation is possible only if the OAIS archive responsible is able to emulate either the publication's original hardware or software environment.</p> <p>Within the preservation community, emulation is considered to be a viable option for some content, although there is no full understanding on how well it will function in the long</p>	<i>disposition of comments.</i>

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)

2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

0	1	2	3	4	5	6	7	8	9
#	NB ¹	Line number (e.g. 17)	Clause/ Subclause (e.g. 3.1)	Paragraph/ Figure/ Table/ (e.g. Table 1)	Type of commen ²	Comment (justification for change) by the NB	Proposed change by the NB	Proposed Disposition of Comments	Disposition of Comments
								term.	
22	US 15 (TC46)		Introduction		Ge	Portico has preserved over 700,000 books and none of them have come with a METS record. Of most importance, we have found, is stressing to producers/publishers that their content be consistent – whatever formats they choose, just use them consistently week-to-week.	Recommend that METS not be required.	METS or some other container standard is required.	Acknowledged. No action. <i>Note: see proposed disposition of comments.</i>
23	US 16 (TC46)		Introduction		Ed	There are robust standards for metadata about software and the entire hardware/software stack needed to emulate content.	Consider identifying the metadata standards in the bibliography.	This version of the standard does not encompass emulation.	Acknowledged. No action. <i>Note: see proposed disposition of comments.</i>
24	US 17 (TC46)		Introduction		Ed	The footnote references have become inline with the sentences, rather than made small and on the top (this seems to be true throughout the document)	Fix the footnotes throughout.	OK.	Accepted. As amended.
25	US 18 (TC46)		Introduction		Ge	If an AIP contains both the original and subsequent versions of an object, it should not need to track the differences in the MD (indeed, this could be very difficult to do)		PREMIS preservation metadata can capture such differences as migration events.	Acknowledged. No action. <i>Note: see proposed disposition of comments.</i>
26	US 19 (TC46)		Introduction 3, 5, 6, 7, 8, 9, Annex A, B		Ed	The words should, should not and shall have been given in uppercase throughout the document. The emphasis on conformance with metadata requirements is appreciated.	Please consider using lowercase should, should not and shall in conformance with standard English practice.	ISO practice must be followed in this case.	Accepted. As amended.
27	W3C 01		Pg 22		Te	There are many mentions of remote resources but limited acknowledgement of external metadata, as discussed in the metadata section of 3.1 [1]. It seems that this would work both for and against		The section mentioned in the comments has been modified to make clear that no matter	Accepted. As amended. <i>Note: see proposed</i>

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)

2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

0	1	2	3	4	5	6	7	8	9
#	NB ¹	Line number (e.g. 17)	Clause/ Subclause (e.g. 3.1)	Paragraph/ Figure/ Table/ (e.g. Table 1)	Type of commen ²	Comment (justification for change) by the NB	Proposed change by the NB	Proposed Disposition of Comments	Disposition of Comments
						archiving. Of course, any necessary metadata would need to be in the package. The EPUB spec requires limited fields in the package, but there are many other fields that an author or publisher might view as necessary. Dc:creator is not a required field, for example. However, the fact that EPUB codifies the use of link means that there is a natural home for the archiving metadata. This is mentioned in passing on page 17 of part 2, but it is probably worth a little more than one sentence. In fact, in part 2, page 22, it says “The EPUB metadata link vocabulary ¹⁹ supports MARC 21, MODS and ONIX”. 3.1 expanded the scope of this.		whether metadata is embedded in SIP or just linked to it, AIP shall contain all relevant metadata, and submission agreement shall clarify how this is guaranteed.	<i>disposition of comments.</i>
28	W3C 02		Pg30		Te	the sample navigation document shows only a list of tables. I suspect that they intended to show a more basic TOC.		It is not clear how the document should be edited based on this comment.	Accepted. As amended. <i>Note: see proposed disposition of comments.</i>

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)

2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial