ISO/IEC 22424:####(X)

ISO IEC JTC 1/SC 34/JWG 7

Date: 2025-04-05

Digital Publishing — EPUB 3 archiving for preservation (EPUB/A)

CD stage

Warning for WDs and CDs

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Contents

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity, or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by ISO IEC Joint Technical Committee 1, Subcommittee 34, Joint Working Group 7.

This first edition replaces earlier documents ISO/IEC TS 22424-1 and ISO/IEC TS 22424-2.

The main changes are as follows:

—   ISO/IEC TS 22424-1 and ISO/IEC TS 22424-2 facilitated the long-term preservation of EPUB publications by, respectively, laying out the principles and metadata requirements. When first written these documents were envisioned as a stepping stone towards a detailed specification. That detailed specification is contained within the current document.

A list of all parts in the ISO/IEC 22424 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

EPUB 3 provides a means of representing, packaging, and encoding structured and semantically enhanced Web content — including HTML, CSS, SVG and other resources — for distribution in a single-file container.

EPUB 3 is generally regarded as a suitable format for digital archiving because it is open/transparent, a standard, generally forward/backward compatible, there is a degree of protection against file corruption, the frequency of version releases, and it is generally interoperable.

EPUB 3 is widely deployed in book publishing and if the archival version of this is standardized it has the potential to become widely used to preserve human knowledge.

This document formalizes a set of requirements for the creation EPUB 3 publications suitable for long-term digital preservation. It disallows the use of features that may jeopardize preservation, for example, while requiring the use of features that improve the suitability.

Where features are restricted, the document endeavours to explain why these features may jeopardize digital preservation. Similarly, it adds justification for those features that are required. The goal is not to define requirements in a vacuum but to enable stakeholders who have adopted EPUB to understand key concepts in digital preservation (e.g., why it is important and how, at a high level, it is done).

EPUB 3 was originally developed by the International Digital Publishing Forum (IDPF), and the version of EPUB 3 submitted to ISO in ISO/IEC 23736 is the equivalent of EPUB 3.0.1, which was produced in 2013. Since that time, IDPF has merged into the World Wide Web Consortium (W3C) and W3C has taken over development of EPUB 3. The W3C's EPUB 3 Working Group produced the latest version of EPUB 3 – EPUB 3.3 – in May 2023. The new standard was developed in accordance with the process requirements of W3C, which are much stricter than were those of the IDPF, so the new version contains many important improvements and clarifications.

EPUB 3.3 is expected to go through the joint ISO and W3C process to become an ISO standard, but this process is not expected to be completed before publication of EPUB/A.

For this reason, this document refers to both the current ISO standard and the latest W3C standard. We have attempted to explain throughout any relevant difference between the versions throughout. For clarity we refer to the current ISO standard as ISO/IEC 23736 and the latest W3C standard as EPUB 3.3. Content that conformed to EPUB 3.0.1 when it was the current version of EPUB 3 will conform to 3.3 now that it is the latest version (so long as you are not using the deprecated features in 3.3).

1 Digital Publishing — EPUB 3 preservation (EPUB/A)

2 1 Scope

3 This document specifies content conformance and packaging requirements for the long-term
4 preservation of EPUB publications. It represents a more restrictive form of EPUB publication than
5 allowed by ISO/IEC 23736 and EPUB 3.3.

6 Markup authoring is not in scope, except where it directly affects the preservability of EPUB/A
7 publications.

8 Preservation of EPUB 2 and EPUB 3.1 publications is not in scope.

9 2 Normative references

10 The following documents are referred to in the text in such a way that some or all their content constitutes
11 requirements of this document. For dated references, only the edition cited applies. For undated
12 references, the latest edition of the referenced document (including any amendments) applies.

13 ISO 15836-2, *Information and documentation — The Dublin Core metadata element set — Part 2: DCMI*
14 *Properties and classes*

15 EPUB 3.3. W3C Recommendation. 25 May 2023. Matt Garrish, et al. Available at
16 https://www.w3.org/TR/epub-33/

17 EPUB Accessibility 1.1. W3C Recommendation. 25 May 2023. Matt Garrish, et al. Available at
18 https://www.w3.org/TR/epub-a11y-11/

19 ISO/IEC 23736-1 *Information technology — Digital publishing — EPUB 3.0.1 — Part 1: Overview*
20
21 ISO/IEC 23736-2 *Information technology — Digital publishing — EPUB 3.0.1 — Part 2: Publications*
22
23 ISO/IEC 23736-3 *Information technology — Digital publishing — EPUB 3.0.1 —Part 3: Content*
24 *documents*
25
26 ISO/IEC 23736-4 *Information technology — Digital publishing — EPUB 3.0.1 — Part 4: Open container*
27 *format*
28
29 ISO/IEC 23736-5 *Information technology — Digital publishing — EPUB 3.0.1 — Part 5: Media overlays*
30
31 ISO/IEC 23736-6 *Information technology — Digital publishing — EPUB 3.0.1 — Part 6:*
32 *Canonical fragment identifiers*
33
34 ISO 23761, *Information technology — Digital publishing —EPUB accessibility — Conformance and*
35 *discoverability requirements for EPUB publications*

36 3 Terms and definitions

37 For the purposes of this document, the terms and definitions given in ISO/IEC 23736, EPUB 3.3, and the
38 following apply.

39 ISO and IEC maintain terminology databases for use in standardization at the following addresses:

40 — ISO Online browsing platform: available at https://www.iso.org/obp
41 — IEC Electropedia: available at https://www.electropedia.org/

42
43 **3.1**
44 **EPUB/A publication**
45 an EPUB publication that conforms to the requirements of this standard.

46 ## 4 Relationship to EPUB 3

47 Whenever the EPUB 3 standard has been updated it has been assigned a new minor number (e.g., 3.0,
48 3.0.1, 3.1, 3.2). These minor revision numbers do not represent new formats, however. Each represents
49 a linear progression in the evolution of EPUB 3, and each is expected to provide backwards compatibility
50 with the preceding versions.

51 The only exception to this rule is EPUB 3.1. This version of the standard intentionally introduced breaking
52 changes to move the standard closer to open web principles (e.g., it fully removed EPUB 3-defined
53 features that are not compatible with web content rendering engines). EPUB 3.1 was subsequently
54 abandoned in favour of a return to backwards compatibility, making EPUB 3.2 a direct successor to EPUB
55 3.0.1 – although retaining many of the improvements of EPUB 3.1. Thus, although it is technically possible
56 to create EPUB 3.1-conforming publications, no publishers do in practice. For this reason, EPUB
57 publications that conform to EPUB 3.1 are not conforming EPUB/A publications.

58 ### 4.1 Relationship between ISO/IEC 23736 and EPUB 3.3

59 Because of the progressive evolution of EPUB 3, EPUB publications that conform to the format as it was
60 defined in ISO/IEC 23736 will predominantly conform to the format as defined in EPUB 3.3. The only
61 EPUB publications that will not conform are those that use features that have subsequently been
62 deprecated or where serious issues have been corrected (e.g., for security), but these have proven to be
63 a minority.

64 The major changes since the publication of ISO/IEC 23736 were to better align EPUB 3 to adapt to
65 changes to the web standards that underpin it. EPUB 3.3 does not reference dated versions of HTML, SVG
66 or Unicode, for example, and no longer defines its own support profile of CSS. This allows EPUB creators
67 to use new features of these technologies as they are published. It also means that the content of EPUB
68 publications that conform to EPUB 3.3 may be more feature rich than that of ISO/IEC 23736, but this will
69 not impact on the preservability of either.

70 There is, in fact, no way to readily identify that an EPUB publication was made to conform to ISO/IEC
71 23736, and it is unlikely that EPUB publications are still produced to this standard. Validation tools are
72 kept up to date with the latest revision of the standard, so it is generally only possible to verify that EPUB
73 publications conform to EPUB 3.3.

74 Where ISO/IEC 23736 and EPUB 3.3 most significantly diverge is in their structure, but this also should
75 not affect their preservability. ISO/IEC 23736 splits the EPUB 3 format into separate parts, where each
76 part defines a specific technology (e.g., one part defines the package document, another how content
77 documents are authored, and another how to package a publication). It lacks a clear "EPUB 3"
78 specification, with ISO/IEC 23736-2[2] generally considered the primary entry point for the standard.
79 ISO/IEC 23736 also combined authoring and reading system developer requirements in each document,
80 making it harder for each audience to find the information they needed.

81 EPUB 3.3 takes what is often called a monolithic approach to standard writing. It combines all the
82 authoring requirements of ISO/IEC 23736 parts 2[2], 3[3], 4[4], and 5[5]. The reading system requirements
83 from those parts are combined into EPUB Reading Systems 3.3[9]. ISO/IEC 23736-6[6] is no longer
84 considered part of the core of EPUB 3 – it was published as a W3C Note and is no longer being actively
85 maintained. The overview of the EPUB 3 format defined in ISO/IEC 23736-1[1] continues to be updated
86 as a W3C Note but this document only provides a non-normative overview of the format.

87 For more detailed information what has changed since the publication of ISO/IEC 23736, it is
88 recommended to review EPUB 3.2 Changes[8] and the revision history in EPUB 3.3.

89 **4.2   Relationship between ISO/IEC 23761 and EPUB Accessibility 1.1**

90 EPUB publications that meet the accessibility requirements of ISO/IEC 23761 will also conform to the
91 requirements of EPUB Accessibility 1.1. The only difference is in the conformance string used to identify
92 conformance.

93 ISO/IEC 23761 uses URLs based on the domain of the International Digital Publishing Forum (IDPF). The
94 IDPF merged with the World Wide Web Consortium (W3C) in 2017 and maintenance of EPUB 3 was
95 transferred to the latter organization, making this domain obsolete.

96 EPUB Accessibility 1.1 uses organization-neutral text-based conformance strings that focus on the
97 version of the standard the content conforms to, as well as the version of the W3C's Web Content
98 Accessibility Guidelines (WCAG).

99 EPUB publications that conform to EPUB Accessibility 1,1 will not always conform to ISO/IEC 23761,
100 however. The updated version of the specification does not require an accessibility summary in the
101 metadata, for example, so publications that omit this metadata will be non-conformant.

102 NOTE   A potential update to fix this incompatibility (i.e., to make accessibility summaries only
103 recommended in both versions) is under consideration by ISO and W3C.

104 EPUB Accessibility 1.1 also accounts for ongoing updates to WCAG 2. EPUB publications that conform to
105 ISO/IEC 23761 can only claim conformance to WCAG 2.0 (ISO/IEC 40500[7]).

106 **5   Publication conformance**

107 An EPUB/A publication:

108 • shall meet all the requirements for EPUB publications as defined in section 2 of EPUB 3.3 with the
109   following additional conditions:
110   o it shall meet the requirements for publication resources defined in section 6.
111   o it shall meet the requirements for the EPUB container defined in section 7.
112   o it shall meet the requirements for the package document defined in section 8.
113 • should meet the recommendations for EPUB publications as defined in EPUB 3.3 as well as the
114   additional recommendations in this document, however failing to meet a recommendation does
115   not invalidate an EPUB/A publication.

116 The requirements for EPUB publications defined in EPUB 3.3 ensure that the content is packaged and
117 formatted properly, and that sufficient information is provided to consistently render the publication and
118 its content. Failing to meet these requirements can leave a publication unrenderable in both obvious (the
119 publication completely fails to load) and subtle ways (reading system behavior could differ, or failures
120 might not be detected until certain content is accessed).

121 This document adds additional restrictions on EPUB publications as set out in the following sections.
122 These sections do not repeat the general requirements defined in EPUB 3.3, so it is necessary to read the
123 documents together to fully understand the requirements of EPUB/A publications. Each section of this
124 document begins by addressing the key differences and then explains why these differences are
125 necessary for long-term preservation.

126 NOTE   EPUBCheck is the de facto validator for all EPUB publications. It is kept current to the latest
127 version of the standard, so the only way to validate against ISO/IEC 23736 is to locate an archived version
128 of the tool from the time of its publication. Such versions are likely to provide incomplete validation,
129 however, as recent fixes to EPUBCheck often address issues running back to EPUB 3.0. Except in rare
130 cases, EPUB publications that conformed to ISO/IEC 23736 will conform to the latest version of EPUB 3.

6   Publication resources

6.1   General

An EPUB/A publication typically consists of many resources – XHTML content documents, images, audio, video, scripts, CSS style sheets, etc. The EPUB 3.3 specification deals with the requirements for these resources separately from how they are packaged in the EPUB container, declared in the package document, or used in EPUB content documents.

The requirements for resources in EPUB/A publications are generally the same as those for EPUB publications, but this section introduces some additional restrictions to improve the long-term preservation of EPUB/A publications.

NOTE   For more information about publication resources, refer to section 3 of EPUB 3.3.

6.2   Core media types

EPUB/A publications shall contain only core media type resources.

EPUB 3 designates a select set of resources as core media types. This designation is applied to resources once they achieve broad support on the web. Practically, this means that reading systems are required to support the formats – provided they support the related functionality (e.g., a reading system without a viewport is not expected to support image formats). All other formats are termed foreign resources and reading system support for them is not guaranteed. Requiring only core media types in EPUB/A publications ensures their long-term renderability.

The list of core media types is defined in section 3.2 of EPUB 3.3.

NOTE The list of core media types in EPUB 3.3 now includes WOFF2 and SFNT fonts, OPUS audio files, and WebP images. These were not valid without fallbacks in ISO/IEC 23736.

6.3   Exempt resources

The use of exempt resources in EPUB/A publications is discouraged, although they cannot always be avoided.

EPUB 3 classifies resources referenced from certain elements as exempt from the core media type requirements. These resources are not guaranteed to be supported by reading systems, but there is no requirement to provide a fallback for them.

For video resources, the use of the H.264 and VP8 codecs is recommended in EPUB 3.3 as reading systems that support video are likely to support at least one of these. This recommendation should also be followed for EPUB/A publications.

EPUB 3 also classifies resources in the EPUB container that are not referenced from spine resources or embedded in EPUB content documents as exempt. This exemption is to allow resources, such as data sets, to travel in the container regardless of their format. As these resources are not used in rendering, they may be included in EPUB/A publications.

NOTE   For the full list of exemptions, refer to section 3.4 of EPUB 3.3.

6.4   Resource fallbacks

EPUB/A publications:

- shall not use EPUB 3's manifest fallbacks mechanism.

- may use HTML's intrinsic fallback mechanisms.

Although EPUB and its underlying technologies, such as HTML, provide fallback mechanisms for foreign resources, these mechanisms are not always well supported. EPUB's manifest fallbacks, in particular, have minimal support. EPUB creators generally avoid foreign resources for this reason, and it is a strong reason why foreign resources are not allowed in EPUB/A publications.

174 Although EPUB/A publications only support core media types, that does not mean that there is no need
175 for fallbacks. The HTML `picture` element's `source` elements can be used to allow the selection of JPEGs
176 of different resolution, for example. This is the reason why intrinsic fallbacks remain supported.

177 While manifest fallbacks were designed to provide fallbacks for core media types in HTML documents
178 (e.g., to supply an alternative image), this feature is even less well supported than for providing fallbacks
179 for resources in the package document spine. It also makes EPUB diverge from the standard web
180 rendering model, and is another reason why the feature is not supported for long-term preservation.

6.5 XML conformance

182 XML-based media types in EPUB/A publications shall be UTF-8 encoded.

183 Although EPUB 3 allows the use of either UTF-8 or UTF-16 for encoding XML-based resources in an EPUB
184 publication, UTF-8 has become the default encoding for web content. Although UTF-16 is supported on
185 the web and in EPUB, it is actively discouraged due to security concerns.

186 NOTE   ISO/IEC 23736 did not prefer one encoding over the other.

7   Open container format

7.1   General

189 The open container format (OCF) defines the virtual file system for an EPUB 3 publication and the ZIP
190 container that packages it. These two parts are called the OCF abstract container (the file system) and the
191 OCF ZIP container (more commonly called the EPUB container).

192 Both the abstract and ZIP containers have specific rules for their production. Unlike a traditional ZIP file,
193 for example, the EPUB container defines specific rules such as required files and zipping order. The
194 abstract container, likewise, has rules on file naming and where files can be located.

195 All EPUB publications must be zipped in conforming EPUB containers for distribution. Similarly, all
196 EPUB/A publications will also be preserved following the same rules for the abstract and ZIP containers,
197 but with additional restrictions to ensure long-term preservation.

198 NOTE   For more information about the Open container format, refer to section 4 of EPUB 3.3.

7.2   Resource location

200 All EPUB/A publication resources shall be located in its EPUB container.

201 EPUB allows specific large data resources to be hosted outside of the EPUB container; specifically, audio,
202 video, fonts, and data files for use by scripts. These are called remote resources in the EPUB standard.

203 Although hosting resources remotely minimizes the initial size of the EPUB container, allowing users to
204 download and begin reading more quickly, it is problematic for long-term preservation as it exposes
205 publications to link rot (i.e., the resources no longer being available, or available at the URLs they were
206 originally referenced from). The loss of these resources can materially impact the ability to read the
207 publication.

208 Providing a copy of all remote resources with an EPUB/A publication does not solve the referencing
209 problem. The EPUB/A publication will continue to reference the resources at their remote locations.
210 Audio and video files may be viewable separately from the publication, but font and data resources will
211 never work as intended unless they are rehosted on the web and all references in the EPUB container
212 updated to point to them.

213 NOTE   ISO/IEC 23736 did not allow EPUB creators to host data files outside of the EPUB container. This is a recent
214 addition.

7.3   Encryption

216 EPUB/A publication resources shall not be encrypted.

217 Although resources in the EPUB container can be encrypted using digital rights management schemes,
218 the application of DRM to an EPUB publication seriously complicates long-term preservation of EPUB/a
219 publications. Even if a DRM scheme is publicly documented, decrypting the content requires the exchange
220 and comparison of public and private key pairs. The long-term availability of the servers involved in these
221 exchanges presents a threat to the ability to decrypt and read EPUB publications. Moreover, the
222 preservation of the public keys necessary to decrypt the content – separate from the entity that controls
223 the DRM scheme – can present legal issues.

224 7.4   Font obfuscation

225 EPUB/A publications may contain font resources obfuscated using EPUB's font obfuscation algorithm.

226 EPUB 3 defines an algorithm for obfuscating font resources to prevent trivial copying of the resources
227 into other publications. The obfuscation provided by this algorithm is easily undone by reversing its steps,
228 so it does not pose a barrier to access.

229 NOTE   This algorithm is not the same as applying a digital rights management scheme to the resources,
230 as defined in 7.3 Encryption, which would make their future use unreliable.

231 7.5   Multiple renditions

232 EPUB/A publications shall not include multiple renditions until such time as the technology becomes a
233 stable W3C Recommendation or is replaced by another technology.

234 Although EPUB 3 allows more than one publication in an EPUB container, by allowing more than one
235 package document to be specified in the container.xml file, in practice this feature is not used and
236 not supported by reading systems.

237 The W3C EPUB 3 Multiple-Renditions 1.1 note[10] attempts to define how multiple publications can be
238 packaged together, but it is only an informational document at the time of writing.

239 8   Package document

240 8.1   General

241 The package document is an XML document that defines the central processing and rendering
242 requirements for EPUB publications. It consists of three key elements:

243 • metadata – The metadata element provides bibliographic and rendering metadata for EPUB
244   reading systems. Bibliographic metadata, such as title and author(s), is used in bookshelves and
245   reading interfaces, while rendering metadata, such as the layout and flow direction, optimizes the
246   publication for display.
247 • manifest – The manifest element contains an enumeration of all the resources in the EPUB
248   publication. It includes information about each resource, such as where each is located, what type
249   of resource each is, processing hints for the file, whether there is a fallback resource, and whether
250   the resource can be synchronized with audio playback.
251 • spine – The spine element defines the default reading order for the EPUB publication. The spine
252   allows reading systems to move users, when they reach the end of one document, automatically
253   to the next, regardless of whether readers are reading the publication linearly from beginning to
254   end or, for example, jumping between chapters.

255 The package document also defines a fourth element for defining collections of resources, but this feature
256 is largely ignored due to a lack of implementation.

257 This section reviews additional EPUB/A requirements for these package document elements necessary
258 for long-term preservation.

259 NOTE   For more information about the package document, refer to section 5 of EPUB 3.3.

8.2 Metadata

EPUB/A publications require the following metadata which may be supplied either in Dublin Core or DC Terms together:

- Creator (`dc:creator`)
- Creation Date or Start Date/End Date (`dc:date)>`)
- Publisher/producer/distributor (`dc:publisher`)
- A standard unique manifestation identifier, such as an ISBN or DOI (`dc:identifier`)
- Rights metadata (`dc:rights`)

The requirements listed here are in addition to those already required by EPUB 3.3/ISO 23736,.

If there is embedded media, there should also be associated metadata for each object. It should at least identify the format and possibly additional attributions and rights if they differ from the container document.

8.3 Manifest

EPUB/A publications add the following restrictions on the package document manifest:

- Each resource shall have the correct media type.
- The "`remote-resources`" and "`switch`" values shall not be used with the `properties` attribute.

EPUB 3 requires that all publication resources be listed in the package document's `manifest` element, with each resource defined by a child `item` element. In addition to the file location, the manifest entries also indicate the media type of each resource. For EPUB/A publications, the media types are required to be core media types (refer to section 6.1).

Although the EPUB 3 specification requires a media type be specified for each resource, it does not require the correct media type be specified. Although this media type is only a hint to reading systems, it is important to have a realiable source for the nature of each resource, which is why EPUB/A makes this a requirement. Validation cannot capture whether media types have been correctly set, so this information must be manually inspected.

The manifest entries also allow reading systems to determine various properties of the content, such as which resource is the navigation document and whether content documents contain scripting. This information is contained in the `properties` attribute. As EPUB/A publications do not allow remote hosting of resources, the "`remote-resources`" value is not allowed., Similarly, as content switching is deprecated in EPUB 3.3, that is why its value is also disallowed.

8.4 Spine

EPUB/A publications shall identify all non-linear content in the spine.

The `spine` element defines the default reading order for reading systems (i.e., the order in which to render each resource as the user progresses through the publication). It also indicates to reading systems whether the items are linear or non-linear (i.e., whether the content the resource contains is considered part of the primary reading order or is secondary content). How reading systems handle linear and non-linear content is not standardized, however. Reading systems can present non-linear content where it is inserted in the spine, suppress it and present only after the user has read the primary content, or suppress it and only allow users to reach it by following links in the content. Despite this inconsistency, identifying non-linear content ensures this information about its purpose is preserved. Spine items not marked as non-linear are linear by default, so there is no corresponding need to identify linear content.

8.5 Collections

EPUB/A publications should not include features defined by `collection` elements until such time as their respective specifications become recognized standards.

306 The package document also supports a named the `collection` element for assembling resources together
307 into a logical grouping. This element can be repeated as many times as necessary after the spine. When
308 ISO/IEC 23736 was published, the idea was that this element could be used to extend EPUB 3 for previews,
309 indexes, dictionaries, and similar content types. Although several extension specifications were produced
310 using this element, they have not found support in reading systems and have been abandoned.

8.6 Legacy features

312 EPUB/A publications may include features of the package document marked as legacy in EPUB 3.

313 Legacy features are only for backwards compatibility with EPUB 2 reading systems and are not used in
314 the rendering of EPUB 3 publications. Their inclusion does not affect long-term preservation.

315 NOTE   For more information about legacy features, refer to section 5.9 of EPUB 3.3.

9   EPUB content documents

9.1   General

318 An EPUB content document is an XHTML or SVG document used in rendering an EPUB/A publication.
319 These documents define the presentation of the content that is presented to users when they read a
320 publication.

321 EPUB 3 defines profiles of XHTML and SVG that include some restrictions on these technologies, as well
322 as allowing some additional features. EPUB 3 has moved away from adding non-standard features,
323 however, such as the custom CSS profile and EPUB-namespaced elements defined in ISO/IEC 23736-3[3].

324 This document is not intended to define content-level restrictions on what XHTML or SVG elements can
325 be use in an EPUB/A publications, but there are content-level features that can impact the ability to
326 preserve publications for the long term. The restrictions in this section are intended only to avoid these
327 complications.

9.2   CSS

329 CSS style sheets in an EPUB/A publication shall be encoded using UTF-8.

330 This requirement is the same as the one imposed on XML-based resourced in 6.4 XML conformance.

331 NOTE   ISO/IEC 23736 did not prefer one encoding over the other.

9.3   Scripting and forms

333 The EPUB content documents in an EPUB/A publication:

334 • shall not include scripts that retrieve resources or information from outside the EPUB container
335   (e.g., using JavaScript APIs such as XmlHttpRequest and Fetch).

336 • shall not include HTML forms that submit data to external servers.

337 Similar to hosting resources outside the container, JavaScript and HTML forms provide EPUB creators
338 the ability to send and receive data from external servers. The servers these features communicate with
339 cannot be relied on to be available in the future, however, hence these restrictions.

10 Accessibility

341 EPUB/A publications shall meet the discoverability metadata requirements of ISO/IEC 23761.

342 Accessibility metadata helps users determine the usability of an EPUB publication. ISO/IEC 23761
343 identifies required and recommended metadata that can be used to inform users about the accessibility
344 features of an EPUB/A publication, any hazards it might present, and the sensory and perceptual faculties
345 necessary to read the content.

346 Even if an EPUB publication is not widely accessible, providing this information helps users determine
347 whether they might still be able to read the publication.

348 NOTE 1 EPUB Accessibility 1.1 slightly loosened the metadata reporting requirements from ISO/IEC 23761. In the
349 newer version, an accessibility summary is only recommended, not required. The change accounts for a summary
350 not being necessary when it would simply restate the conformance claim.

351 EPUB creators are strongly encouraged to ensure their EPUB/A publications meet all the requirements
352 of ISO/IEC 23761, as this is a recommendation of both ISO/IEC 23736 and EPUB 3.3. Meeting these
353 conformance requirements will help make EPUB/A publications accessible to the broadest spectrum of
354 users.

355 EPUB/A publications that fully meet accessibility conformance requirements can be differentiated from
356 those that do not by the presence of a `dcterms:conformsTo` property. When the value of this property
357 matches one of the strings specified in ISO/IEC 23761 or EPUB Accessibility 1.1, the publisher is stating
358 conformance to that standard and indicating which version and level of the W3C Web Content
359 Accessibility Guidelines (WCAG) the publication attains.

360 NOTE 2 EPUB Accessibility 1.1 differs from ISO/IEC 23761 in allowing for a greater range of WCAG 2 conformance.
361 ISO/IEC 23761 only allows conformance against WCAG 2.0 (ISO/IEC 40500[7]), while EPUB Accessibility 1.1 allows
362 conformance claims against any subsequent version of WCAG 2.X. The conformance identifier string also changed
363 across versions so that it is no longer relies on a URL associated with the former maintenance organization of EPUB.

11 Deprecated features

365 EPUB/A publications shall not include features marked as deprecated in EPUB 3.3.

366 Reading system support for these features is poor, so they do not work widely, and developers are not
367 expected to add support.

368 NOTE   The following features of EPUB 3 have been deprecated since the publication of ISO/IEC 23736: bindings
369 and new collection types in the EPUB package document, content switching and triggers in EPUB content documents,
370 and viewport declarations in fixed-layout documents. In the package document metadata, using specific types for
371 linked metadata records has also been deprecated, as has specifying metadata authorities.

12 Conformance

373 EPUB/A publications that meet the requirements of Clause 5 shall indicate their conformance in the
374 package document metadata by including a `conformsTo` property in accordance with ISO 15836-2. The
375 value of this property shall be "`EPUB/A 1.0`".

376 Adding a conformance statement to the package document metadata ensures that EPUB/A publications
377 can be differentiated from non-conforming EPUB publications regardless of how they are preserved.

378 The following example shows the tagging for a conformance claim.

```
<metadata>
   …
 <meta property="dcterms:conformsTo">EPUB/A 1.0</meta>

   …
</metadata>
```

13 Archive location

385 The EPUB/A files for scholarly and academic books, educational material, and even some trade titles form
386 an important part of our shared cultural and intellectual heritage. Safeguarding this heritage – in the form
387 of archived copies of those books and e-books – is the remit of national libraries, and for digital material,
388 of 'dark archives' committed to long-term preservation. Copies lodged with these preservation services

389 are secured and inaccessible and may be made available under controlled circumstances following some
390 catastrophic event in the distant future.

391 It is good practice to communicate where EPUB/A files are preserved. This can be done using ONIX or
392 MARC metadata.

393

394 13.1 ONIX Metadata

395 Preservation provider metadata may be included in ONIX records for books.

396 The location of an archived copy can be specified using the `<Website>` composite, usually within the
397 `<Publisher>` context:

398

```
399     <Publisher>
400         <PublishingRole>01</PublishingRole>                          <!-- publisher -->
401         <PublisherName>Wolters Kluwer</PublisherName>
402         <Website>
403             <WebsiteRole>18</WebsiteRole>                  <!-- publisher's B2C website -->
404             <WebsiteLink>https://www.wolterskluwer.com</WebsiteLink>
405         </Website>
406         <Website>
407             <WebsiteRole>48</WebsiteRole>       <!-- long-term preservation location -->
408             <WebsiteLink>https://clockss.org</WebsiteLink>
409         </Website>
410     </Publisher>
```

411

412 The example above specifies both the publisher's own (B2C) website and the website of a third-party
413 preservation service (a publisher's own dark archive would use role code 47, and digital preservation
414 within a National Library archive can be specified using role code 51). The URL in `<WebsiteLink>`
415 identifies the particular preservation service or archive. An extended preservation URL may point to
416 metadata to confirm the preservation status of the book – this is recommended where the particular
417 archive's preservation metadata is publicly accessible.

418 In addition, the publisher can nominate a role-based contact for queries relating to legal deposit and long-
419 term preservation in the `<ProductContact>` composite:

420

```
421     <ProductContact>
422         <ProductContactRole>08</ProductContactRole>.         <!-- deposits contact -->
423         <ProductContactName>Production Dept, My Publisher Inc</ProductContactName>
424         <EmailAddress>preservation@mypublisher.com</EmailAddress>
425     </ProductContact>
```

426

427 In this case, the contact is a generic role-based e-mail address, but it could be a named person, and it need
428 not be a contact at the publisher – it could be a contact at the distributor, for example. Role code 08
429 specifies a contact for purposes of CIP, legal deposit and long-term preservation queries.

430   For more information, see page 4 of this EDItEUR application note.

13.2 MARC Metadata

432   For information about how digital preservation is handled in MARC records, please refer to the OCLC
433   BibFormats document, https://www.oclc.org/bibformats/en/8xx/857.html.

Bibliography

[1]   ISO/IEC 23736-1   *Information technology — Digital publishing — EPUB 3.0.1 — Part 1: Overview*

[2]   ISO/IEC 23736-2 *Information technology — Digital publishing — EPUB 3.0.1 — Part 2: Publications*

[3]   ISO/IEC 23736-3 *Information technology — Digital publishing — EPUB 3.0.1 —Part 3: Content documents*

[4]   ISO/IEC 23736-4 *Information technology — Digital publishing — EPUB 3.0.1 — Part 4: Open container format*

[5]   ISO/IEC 23736-5 *Information technology — Digital publishing — EPUB 3.0.1 — Part 5: Media overlays*

[6]   ISO/IEC 23736-6 *Information technology — Digital publishing — EPUB 3.0.1 — Part 6: Canonical fragment identifiers*

[7]   ISO/IEC 40500 *Information technology — W3C Web Content Accessibility Guidelines (WCAG) 2.0*

[8] EPUB Changes 3.2. W3C Final Community Group Specification. 8 May 2019. Matt Garrish, et al. Available at https://www.w3.org/publishing/epub32/epub-changes.html

[9] EPUB Reading Systems 3.3. W3C Recommendation. 25 May 2023. Matt Garrish, et al. Available at https://www.w3.org/TR/epub-rs-33/

[10] EPUB 3 Multiple-Rendition Publications 1.1. W3C Note. 21 December 2023. Matt Garrish. Available at https://www.w3.org/TR/epub-multi-rend-11/