ISO #####:####(X)

ISO IEC JTC 1/SC 34/JWG 7

Date: 2024-04-08

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

WD 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.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

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.

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 of EPUB 3 publications suitable for long-term digital preservation. It disallows the use of features that may jeopardize preservation while requiring the use of features that improve preservability.

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 the 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 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 the 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 any relevant difference between the versions throughout this document. 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).

This document also describes how the location of archived EPUB/a files can be communicated.

Digital Publishing — EPUB 3 preservation (EPUB/A)

1 Scope

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

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

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

2 Normative references

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

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

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

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

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

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

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

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

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

ISO 23761, *Information technology — Digital publishing —EPUB accessibility — Conformance and discoverability requirements for EPUB publications*

3 Terms and definitions

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

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

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

3.1
EPUB/A publication
an EPUB publication that conforms to the requirements of this standard.

4 Relationship to EPUB 3

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

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

1.1 Relationship between ISO/IEC 23736 and EPUB 3.3

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

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

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

Where ISO/IEC 23736 and EPUB 3.3 most significantly diverge is in their structure, but this also should not affect their preservability. ISO/IEC 23736 is like an atomization of the EPUB 3 format, where each part defines a specific technology of EPUB 3. It lacks a clear “EPUB 3” specification, with ISO/IEC 23736-2[2] generally considered the primary entry point for the standard. ISO/IEC 23736 also combined authoring and reading system developer requirements in each document, making it harder for each audience to find the information they needed.

EPUB 3.3 takes what is often called a monolithic approach to standard writing. It combines all the authoring requirements of ISO/IEC 23736 parts 2[2], 3[3], 4[4], and 5[5]. The reading system requirements from those parts are combined into EPUB Reading Systems 3.3[8]. ISO/IEC 23736-6[6] is no longer considered part of the core of EPUB 3 – it was published as a W3C Note and is no longer being actively maintained. The overview of the EPUB 3 format defined in ISO/IEC 23736-1[1] continues to be updated as a W3C Note but this document only provides a non-normative overview of the format.

For more detailed information what has changed since the publication of ISO/IEC 23736, it is recommended to review EPUB 3.2 Changes[7] and the revision history in EPUB 3.3.

2 Publication conformance

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

For these reasons, an EPUB/A publication:

- shall meet all the requirements for EPUB publications as defined in section 2 of EPUB 3.3 with the following additional conditions:
  - it shall meet the requirements for publication resources defined in section 6.
  - it shall meet the requirements for the EPUB container defined in section 7.
- it shall meet the requirements for the package document defined in section 8 and should meet the recommendations for EPUB publications as defined in EPUB 3.3 as well as the additional recommendations in this document, but failing to meet a recommendation does not invalidate an EPUB/A publication.

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

3 Publication resources

3.1 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.

3.2 Core media types

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.

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.

To ensure the long-term renderability of EPUB/A publications, they shall contain only core media type resources.

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, for example, 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.

Fallback mechanisms may be used to provide options between core media type resources. For example, the HTML picture element's source elements can be used to allow the selection of JPEGs of different resolutions.

6.2 Exempt resources

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.

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

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.

6.3 Character encoding

Although EPUB 3 allows the use of either UTF-8 or UTF-16 for encoding resources in an EPUB publication, EPUB/A publications shall only use UTF-8 encoding.

UTF-8 has become the default encoding for web content. Although UTF-16 is supported on the web and in EPUB, it is actively discouraged due to security concerns.

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

7 Open container format

7.2 General

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

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

All EPUB publications must be zipped in conforming EPUB containers for distribution, so all EPUB/A publications will also be preserved following the same rules for the abstract and ZIP containers. This section details further restrictions that EPUB/A places on the EPUB container to ensure long-term preservation.

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

7.3 Resource location

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

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

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

For these reasons, publication resources shall be located in the EPUB container.

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

7.4 Encryption

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

For these reasons, resources in the EPUB container shall not be encrypted.

7.5 Font obfuscation

EPUB 3 defines an algorithm for obfuscating font resources to prevent trivial copying of the resources into other publications. The obfuscation provided by this algorithm is easily undone by reversing its steps, so it does not pose a barrier to access. Consequently, EPUB/A publications may contain font resources obfuscated using this algorithm.

NOTE: This algorithm is not the same as applying a digital rights management scheme to the resources, as defined in 5.3 Encryption, which would make their future use unreliable.

7.6 Multiple renditions

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

The W3C EPUB 3 Multiple-Renditions 1.1 note[9] attempts to define how multiple publications can be packaged together, but it is only an informational document at the time of writing. EPUB/A publications must not include multiple renditions until such time as the note becomes a stable W3C Recommendation or is replaced by another technology.

8 Package document

8.2 General

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

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

This section reviews additional requirements for the package document necessary for long-term preservation.

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

8.3 Metadata

Both EPUB 3 and ISO/IEC 23736 require the metadata to include a title (dc:title), language (dc:language), unique identifier (dc:identifier) and last modification date (dcterms:modified). EPUB/A publications require the following additional metadata:

- Information about any compression methods or any reading systems for which the file is optimized
- Creator
- Creation Date or Start Date/End Date
- Publisher/producer/distributor
- A standard unique manifestation/work identifier (e.g. ISBN or ISSN)
- An item or instance identifier, if available (e.g. a DOI)
- Metadata format and its versions
- Schema identifiers for metadata records (e.g. Dublin Core or PREMIS)
- Administrative metadata (e.g. nature and formats of embedded media)
- Technical metadata (e.g. file formats and versions, digital signatures and checksums)
- Rights metadata
- Structural metadata

8.4 Manifest

EPUB 3 requires that all publication resources be listed in the package document's `manifest` element, with each resource defined by an `item` element in the manifest. It is the second element in the package document after the metadata. 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. For EPUB/A publications, each resource shall have the correct media type. 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 shall not be used with this attribute. In addition, as content switching is deprecated in EPUB 3.3, the "`switch`" value also shall not be used.

8.5 Spine

The third element of the package document is the spine. It 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).

The spine 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, EPUB/A publications shall identify all non-linear content to ensure that the information about its purpose is preserved (all spine items not marked as non-linear are linear by default, so there is no corresponding need to identify linear content).

8.6 Collections

The package document defines a `collection` element for assembling resources together into a logical grouping. This element can be repeated as many times as necessary after the spine. When ISO/IEC 23736 was published, the idea was that this element could be used to extend EPUB 3 for previews, indexes, dictionaries, and similar content types. Although a number of extension specifications were produced using this element, they have not found support in reading systems and have been abandoned. Until such time as these specifications become recognized standards, EPUB/A publications should not include the features they define.

8.7 Legacy features

EPUB/A publications may include features of the package document marked as legacy in EPUB 3. As these features are only for backwards compatibility with EPUB 2 reading systems and are not used in the rendering of EPUB 3 publications, their inclusion does not affect long-term preservation.

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

9 EPUB content documents

9.2 General

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

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

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

9.3 Scripting and forms

Similar to hosting resources outside the container, JavaScript and HTML forms provide EPUB creators the ability to send and receive data from external servers. As the servers these features communicate with cannot be relied on to be available in the future, EPUB content documents:

- shall not include scripts that retrieve resources or information from outside the EPUB container (e.g., using JavaScript APIs such as XmlHttpRequest and Fetch).
- shall not include HTML forms that submit data to external servers.

10 Accessibility

To help users determine the accessibility of an EPUB publication, ISO/IEC 23761 includes a section on providing discoverability metadata. This metadata informs users about the accessibility features of the publication, any hazards it might present, and the sensory and perception faculties necessary to read the content. Even if an EPUB publication is not fully accessible, providing this information allows users to determine whether it is usable by them or not.

To ensure that the accessibility of EPUB/A publications is always determinable, they shall meet the discoverability metadata requirements of ISO/IEC 23761.

EPUB creators are strongly encouraged to ensure their EPUB/A publications meet all the requirements of ISO/IEC 23761. Following its guidelines ensures EPUB/A publications will be available to the broadest spectrum of users.

11 Deprecated features

EPUB/A publications shall not include features marked as deprecated in EPUB 3.3. Reading system support for these features is poor, so they do not work widely, and developers are not expected to add support.

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

12 Reporting where the EPUB/a is archived

The EPUB/a files for scholarly and academic books, educational material, and even some trade titles form an important part of our shared cultural and intellectual heritage. Safeguarding this heritage – in the form of archived copies of those books and e-books – is the remit of national libraries, and for digital material, of 'dark archives' committed to long-term preservation. Copies lodged with these preservation services are secured and inaccessible, and may be made available under controlled circumstances following some catastrophic event in the distant future.

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

12.2 ONIX Metadata

Preservation provider metadata can be included in ONIX records for books.

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

```
<Publisher>
    <PublishingRole>01</PublishingRole>                          <!-- publisher -->
    <PublisherName>Wouters Kluwer</PublisherName>
    <Website>
        <WebsiteRole>18</WebsiteRole>              <!-- publisher's B2C website -->
        <WebsiteLink>https://www.wolterskluwer.com</WebsiteLink>
    </Website>
    <Website>
        <WebsiteRole>48</WebsiteRole>      <!-- long-term preservation location -->
        <WebsiteLink>https://clockss.org</WebsiteLink>
    </Website>
</Publisher>
```

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

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

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



```
</ProductContact>
```

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

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

12.3 MARC Metadata

For information about how digital preservation is handled in MARC records, please refer to the OCLC 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] 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

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

[9] 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/

