
**Digital publishing — EPUB3
preservation —**

**Part 2:
Metadata requirements**

**iTeh STANDARD PREVIEW
(standards.iteh.ai)**

[ISO/IEC TS 22424-2:2020](https://standards.iteh.ai/catalog/standards/sist/6593bc04-42f9-4fed-aa39-0583d68d439d/iso-iec-ts-22424-2-2020)

<https://standards.iteh.ai/catalog/standards/sist/6593bc04-42f9-4fed-aa39-0583d68d439d/iso-iec-ts-22424-2-2020>



iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC TS 22424-2:2020
<https://standards.iteh.ai/catalog/standards/sist/6593bc04-42f9-4fed-aa39-0583d68d439d/iso-iec-ts-22424-2-2020>



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

	Page
Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Abbreviated terms	2
5 Syntax	2
6 Packaging metadata	4
6.1 General.....	4
6.2 Package creator / submitter information.....	4
6.3 Package status.....	5
6.4 Package identifier.....	5
6.5 Work and publication identifiers.....	6
6.6 Core media type resource identifiers.....	8
6.7 Foreign resource identifiers.....	9
6.8 Identifiers for metadata records.....	10
6.9 Dates.....	11
6.9.1 General.....	11
6.9.2 Creation date of a submission information package.....	12
6.9.3 Modification date of a submission information package.....	12
6.9.4 Creation/modification date of an EPUB publication.....	12
6.9.5 Creation/modification of a metadata record.....	13
6.10 Metadata format and its versions.....	13
7 Administrative metadata	15
7.1 General.....	15
7.2 Technical metadata.....	16
7.2.1 File formats and their versions.....	16
7.2.2 Digital signatures and checksums.....	19
7.3 Rights metadata.....	20
7.3.1 General.....	20
7.3.2 Preservation related rights.....	21
7.4 Structural metadata.....	22
7.5 Preservation metadata.....	24
8 Structure of submission information packages	26
9 Content of submission information packages	27
Annex A (informative) Digital signature	29
Annex B (informative) Events	31
Bibliography	35

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

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 document 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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see <http://patents.iec.ch>).

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 Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 34, *Document description and processing languages*.

A list of all parts in the ISO/IEC TS 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

This document facilitates the long-term preservation of EPUB publications by specifying metadata elements which are required or recommended for long-term preservation (such as identifiers) and the ways in which the EPUB publication and related metadata can be packaged. EPUB versions 3 and 3.0.1 are covered; if necessary, the EPUB version applicable is specified.

Long-term preservation in general requires two things:

- making the object such as EPUB publication fit for preservation – including features to be used and feature to avoid;
- packaging the object (and any metadata related to it) together with any additional data such as other versions of the object and other documentation into an Open Archival Information System (OAIS) submission information package (SIP).

ISO/IEC TS 22424-1 concentrates on the archivability of EPUB documents.

The background to this document comes from the Open Archival Information System, which is described in ISO/IEC TS 22424-1.

When a submission information package (SIP) is formed, mandatory preservation metadata need to be present in the package. Depending on the agreements made between the producer and the archive, metadata elements are stored either in the container document or the EPUB publication itself, or both. Usually an archive would expect to find all relevant metadata in the container, unless the submission agreement allows embedding of metadata into EPUB publications.

This document does not require any changes to be made to the current or future EPUB standards. However, when an EPUB publication is created or modified for submission to an archive, there are some EPUB features that should be used and others that should be avoided. ISO/IEC TS 22424-1 describes how the EPUB format should be applied. This document concentrates on mandatory and recommended metadata elements needed for the long-term preservation of EPUB publications and their METS encoding. ISO/IEC TS 22424-1 recommends the usage of METS but allows also other container standards; this document concentrates on preservation metadata and its METS encoding in SIPs. Future editions of these documents may specify other encodings such as BITS (Book Interchange Tag Suite)¹⁾.

In order to guarantee access to documents, OAIS archives may migrate documents into new file formats when the original formats are no longer supported by commonly used rendering tools. If the document to be migrated is an e-book in an outdated EPUB format, migration can be made to a more modern version of EPUB or, at least in principle, to another e-book format.

Generally, migration into another file format should be straightforward if the current and new format are compatible and there are efficient and reliable migration tools available. If the target format is a more modern version of the current format, compatibility should not be a problem. But if a format is rich, migration tools may not be able to render all the properties of a resource.

This document applies to EPUB versions 3 and 3.0.1. Earlier versions (EPUB 2 and 2.0.1) are not covered. Since there are no implementations of version 3.1, it is not covered in this document either. EPUB 3.2 was published in May 2019²⁾. It will be taken into account in the next edition of this document.

This document does not cover issues related to migration between EPUB versions or from EPUB to other e-book formats. Migration to other formats is often lossy; this applies to e-book formats as well, since there are EPUB features which are not supported in other e-book formats, and vice versa. Moreover, even if the same feature is supported, technical implementations can be incompatible. For instance, if an EPUB 3 publication using fixed layout is migrated to Amazon's KF8 format, preserving fixed layout properties requires special attention since there are significant technical differences between these formats in how this feature has been implemented.

1) <https://www.loc.gov/preservation/digital/formats/fdd/fdd000453.shtml>

2) <https://w3c.github.io/publ-epub-revision/epub32/spec/epub-spec.html>

Sometimes migration 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 program's original hardware or software environment.

Within the preservation community, emulation is considered to be a viable option for some content. For the time being there is no full understanding on how emulation will function in the long-term, but this may change with emulation as a service approach coming to the market.

Metadata requirements in this document are based on the migration of file formats. Emulation is not covered (just a single example of emulation-related preservation metadata is given), although emulation is likely to be the best preservation method for fixed layout EPUB publications and interactive EPUB publications. Preservation metadata requirements for emulation-based preservation strategy may be added into a future version of this document.

Supporting emulation might require just information about appropriate tools in the submission agreement or in the related documentation. A more sustainable approach is to include a description of the emulation environment (hardware and/or software) in the premis:object section of the PREMIS metadata record in the SIP. During ingest this information is copied into the archival information package (AIP). If migration is used, hardware and software environments needed for rendering the versions of the document in the AIP can be specified separately as access environments.

Ambition level of migration may vary. Usually it is to preserve the intellectual content, since retaining also the original look and feel of preserved documents is considered to be too demanding. If semantics and layout are interlinked, it is important to keep also the original EPUB publication in order to facilitate preservation of the semantics via emulation-based access to the original content.

Migration both requires and produces preservation metadata. For instance, staff in the archives has to figure out which tools can be used to carry out the migration and what weak points they may have. The intention of the preservation community is to maintain this information in format libraries such as PRONOM³). When a new AIP is created after a migration, the package should contain both the old and the new representation of the migrated document and preservation metadata describing the migration event and the possible differences between the document versions⁴). Depending on their needs and archived resources archive users can then make a choice between the original, which is authentic but possibly difficult to render, and the migrated document, which should be easy to use but less authentic. In practice, finding access software to outdated versions of preserved documents may be difficult. The OAIS archive, on the other hand, can migrate the original document again when better tools can be used, or if there are significant issues in migrated documents.

Metadata elements that need to be included in SIPs are a priori essential for digital preservation. For instance, if there is no digital signature present and a secure transfer channel has not been used, it is impossible to guarantee the information entering the archive has not changed during transfer or that it is coming from a correct source. Moreover, if the data has already been tampered with before it enters the archive, all subsequent preservation actions may be useless.

This document does not specify generic conformance requirements for EPUB publications, but may make some restrictions to the use of EPUB specifications. The generic conformance requirements made in the EPUB Contents Documents Specification apply to EPUB publications in SIPs as well.

ISO/IEC TS 22424-1 defined a set of requirements for archivable EPUB publications. Please consult ISO/IEC TS 22424-1 for more information.

3) <http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>

4) This document is only concerned with those metadata elements which are to be included in SIPs. Preservation metadata needed in AIPs (which describes the preservation related events such as migration) is beyond the scope.

Digital publishing — EPUB3 preservation —

Part 2: Metadata requirements

1 Scope

The ISO/IEC TS 22424 series supports long-term preservation of EPUB publications via a dual strategy. This document makes EPUB compliant with current practices of Open Archival Information Systems (OAIS) archives and technical requirements of repository systems. The former tend to rely on OAIS in their operations; the latter prefer to ingest electronic documents only in containers conforming to standards such as METS (Metadata Encoding and Transmission Standard).

ISO/IEC TS 22424-1 considers EPUB features from a long-term preservation point of view.

2 Normative references

The following documents are referred to in the text in such a way that some or all of 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 8601 (all parts), *Date and time — Representations for information interchange*

ISO/IEC TS 22424-1, *Digital publishing — EPUB3 preservation — Part 1: Principles*

METS *Metadata Encoding & Transmission Standard, Version 1.12.1*. [online]. Library of Congress, 2019. Available from: <https://www.loc.gov/standards/mets/>

PREMIS *PREMIS Data Dictionary for Preservation Metadata, Version 3.0*. [online]. Library of Congress, 2015. Available from <http://www.loc.gov/standards/premis/>

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC TS 22424-1 and the following apply.

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

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

3.1 data dictionary

organized and constructed (electronic data base) compilation of descriptions of data concepts that provides a consistent means for documenting, storing and retrieving the syntactical form (i.e. representational form) and the meaning and connotation of each data concept

Note 1 to entry: PREMIS is a data dictionary. PREMIS Data Dictionary for Preservation Metadata (<https://www.loc.gov/standards/premis/>) is a leading metadata specification for metadata needed for long-term preservation.

[SOURCE: ISO 24531:2013, 4.14, modified — Note 1 to entry has been added.]

3.2

structural metadata

metadata that indicates how compound objects are put together, for example how the pages of a document are arranged to form chapters

Note 1 to entry: The definition is adapted from Reference [14].

4 Abbreviated terms

AIP	archival information package
DIP	dissemination information package
DRM	digital rights management
OAIS	Open Archival Information System
PDI	preservation description information
SIP	submission information package

5 Syntax

This document provides examples of how metadata elements should be expressed using either

- 1) Metadata Encoding and Transmission Standard (METS⁵) version 1.12.1 and PREMIS Data Dictionary for Preservation Metadata (PREMIS⁶) version 3.0, and/or
- 2) EPUB version 3.0 and 3.0.1

for encoding SIPs. Other container standards may be added to the future editions of this document.

This dual approach was chosen because there are different options available for a producer to turn existing EPUB publications into SIPs:

- 1) All metadata (mandatory and otherwise) may be embedded in the EPUB publication.
- 2) Mandatory metadata is copied from EPUB document to the METS container if and when it is already present, or created and placed in the METS container (recommended approach).
- 3) Option 2, but a container standard other than METS is used.

The first option looks appealing because that way it would be relatively easy to create EPUB publications suitable for long-term preservation, especially if the mandatory metadata elements are already present (and if the EPUB publication itself does not have features unsuitable for preservation).

Unfortunately this approach has some issues:

- Commonly used repository systems expect information packages based on container standards such as METS. Current versions of these applications may not be able to process SIPs which contain only an EPUB publication.
- Depending on the mandatory metadata required, it may not be possible to include all preservation metadata into EPUB publication.

5) <http://www.loc.gov/standards/mets/>
6) <http://www.loc.gov/standards/premis/>

- If there is no container document, it may be difficult to send multiple EPUB publications in a single SIP, or partial updates (for instance, only descriptive metadata about a publication that has already been archived).

Options 2 and 3 are based on the idea that there are two independent specifications, the core EPUB specification (currently version 3.2), and a container specification (this document). This allows the two communities (EPUB and digital archivists) to cooperate without putting unnecessary constraints on each other. Both specifications are independent from one another, which makes it easier to manage them.

From a technical point of view, the main strength of the second option is that METS containers are almost universally accepted in long-term preservation applications. One reason for the popularity of the standard is that it is flexible – it is possible to embed any descriptive or administrative metadata into a METS document. Whatever mandatory metadata will be agreed upon by the producer and the OAIS archive, METS can be used as a container.

The option of using some other container standard than METS or EPUB is not examined in this document. METS is used due to its technical features and popularity among long-term preservation application vendors as well as libraries, archives, and museums. If and when other options emerge in the future, it is possible to extend this document to support other container standards as well.

The main weakness of METS approach is that currently very few publishers support it. Unless production processes change radically, a common solution will be to submit e-books in EPUB format as such, with accompanying ONIX metadata. In this approach, the producer (which can be the OAIS archive) creates the METS SIP during pre-ingest, using the data and metadata delivered by the publisher. The publisher does not need to know METS, but EPUB documents themselves and the accompanying metadata should meet the requirements made in the submission agreement.

This document requires that each SIP shall have a METS document with mandatory descriptive and administrative metadata elements embedded, using e.g. Dublin Core (ISO 15836-1) and PREMIS formats. The use of a separate, METS based preservation layer enables the current long-term preservation applications to ingest EPUB publications. Producers and OAIS archives may also choose other approaches, such as embedding all metadata in EPUB publications or using another container standard. Whichever strategy is chosen, it should be planned out carefully.

In the hybrid approach, some descriptive and administrative metadata needed during ingest may not be copied from the EPUB document to the METS document. In order to use this metadata, the OAIS archive shall have reading systems or other applications which are able to render EPUB publications and extract the relevant metadata from them.

This document does not require copying of EPUB structural metadata to METS documents. Therefore, the structural metadata in METS is simple, only specifying the location of EPUB publication or publications in the SIP but not their internal structure. EPUB reading systems would not be able to use the structural metadata in a METS document, because they utilize structural metadata in the EPUB spine element when publications are rendered.

In order to eliminate uncertainty concerning the syntax and semantics of SIPs, submission agreements shall specify a METS profile or profiles which can be used to facilitate packaging of EPUB publications. This document can be used as a basis for these profiles. The profile can be part of the submission agreement, or linked to it. The latter approach was chosen in the Finnish Digital Library initiative; the benefit is that submission agreements will be relatively simple because technical details are stated in the document “Metadata requirements and preparing content for digital preservation”⁷⁾. Finnish Digital Library initiative has published also a separate document titled “File formats”⁸⁾, which lists the file formats suitable for ingest and preservation. Unfortunately, this document does not contain guidelines on how these file formats should be applied. EPUB is an example of a file format which is in principle archivable, but in practice can be used in a way which may makes long-term preservation challenging. The purpose of ISO/IEC TS 22424-1 is to provide guidelines for creation of archivable EPUB publications.

7) <http://digitalpreservation.fi/files/Metadata-1.7.1-en.pdf>

8) <http://digitalpreservation.fi/files/File-Formats-1.7.0-en.pdf>

Specifications, such as the ones created in Finnish Digital Library initiative, shall be sufficiently detailed; for instance, they shall specify all mandatory metadata elements and all archivable or ingestible file formats. Otherwise SIPs may lack crucial data, or contain files that cannot be processed. Of course even this may not be sufficient; in addition to only saying that MXF, TIFF and EPUB are archivable formats, it is also necessary to specify what type of MXF videos, TIFF images and EPUB publications are acceptable. Digital archiving projects like the National Digital Library in Finland do not necessarily have a mandate or resources for such work; that is why specifications like this one for EPUB, AS-07 for archivable MXF⁹⁾ and TI/A¹⁰⁾ for archival of TIFF images are needed.

If just listing all the archivable file formats is not enough, it is also insufficient to provide just a list of mandatory preservation metadata elements. Element specific guidelines are often necessary. For instance, it is not enough to just say that SIPs must contain identifiers for EPUB publications. Producer and OAIIS archive shall also agree on what needs to be identified (for instance, EPUB publications, their component parts, metadata records), which identifiers (ISBNs, DOIs, URNs, etc.) are accepted and – just to give an EPUB specific example on identifier usage – whether EPUB release identifiers are acceptable. Metadata is crucial in digital archiving, because it affects all the steps in the preservation process – ingest, archival, and dissemination. When a producer and an OAIIS archive decide on which identifiers to use, this may have an impact not only on SIPs, but also on archival information packages (AIPs) and dissemination information packages (DIPs) the archive will be able to create.

If the SIP does not meet the requirements, usually the ingest process fails and OAIIS archive asks the provider to fix the problem. But submission agreement can specify other approaches; for instance, if the provider does not have sufficient technical skills, the OAIIS archive or a third party could take care of fixing technical problems in submitted EPUB documents may be submission agreement. It might even be possible to ignore certain minor issues during ingest, although even minor problems may endanger long term preservation.

Sometimes it is not possible or practical to create SIPs which meet all the requirements. For instance, an SIP may contain the same resource both in the original (non-archivable) and archivable formats. In such case, METS encoding should indicate that the original file is not validated during ingest. Omission of mandatory metadata element(s) should be agreed upon between the producer and the OAIIS archive in advance, in order to avoid ingest failures.

6 Packaging metadata

6.1 General

This clause covers mainly metadata about the SIP (container) which is usually submitted using METS elements and attributes.

NOTE It is not possible to make a clear division between descriptive and administrative metadata. For instance package creator information is normally just administrative metadata. But if the package creator has modified the EPUB publication to make sure that SIP meets the requirements of the submission agreement, the creator could have performed tasks which normally belong to the editor of the publication. The name of the editor is regarded as descriptive metadata.

6.2 Package creator / submitter information

Both the name of the original creator of the package and the name of the submitting organization shall be included in the METS header, if the submitting organization has made any changes to the package. If the submitting organization has not modified the content, the creator name is sufficient.

If a secure transmission channel is used and it allows identification of the submitting organization, submitter information may be omitted.

9) http://www.digitizationguidelines.gov/guidelines/MXF_app_spec.html

10) <http://www.preforma-project.eu/dpf-manager.html>

Creator / submitter identifier should be included, if the name alone does not uniquely identify the organization. The identifier should be an ISNI or another standard identifier. The identifier system in use shall be indicated.

Examples

SIP creator:

```
<mets:metsHdr CREATEDATE="2017-07-15T12:00:00" RECORDSTATUS="NEW">
  <mets:agent ROLE="CREATOR" TYPE="ORGANIZATION">
    <mets:name> National library of Finland </mets:name>
  [...]
</mets:metsHdr>
```

SIP submitter:

```
<mets:metsHdr CREATEDATE="2018-02-11T08:00:00" RECORDSTATUS="NEW">
  <mets:agent ROLE="PRESERVATION" TYPE="ORGANIZATION">
    <mets:name> Kansalliskirjasto </mets:name>
    <mets:note> ISNI 0000 0001 2033 7602 </mets:note>
  [...]
</mets:metsHdr>
```

6.3 Package status

The METS header RECORDSTATUS attribute value "REPLACEMENT" should be used to indicate the status of the package if the package is resubmitted. If the attribute is not present, its value is assumed to be "NEW".

Example

Modified SIP to replace one sent earlier:

```
<mets:metsHdr CREATEDATE="2018-01-10T17:12:55" RECORDSTATUS="REPLACEMENT">
  [...]
</mets:metsHdr>
```

6.4 Package identifier

Every SIP shall have a unique identifier. The submission agreement shall specify the identifier type or types used (for instance, UUID).

In practice, some producers may prefer to use alternative methods, such as time stamp added to the file name. Such arrangements shall be specified in the submission agreements.

SIPs themselves are not preserved after the ingest process is finished, but the SIP identifier may be preserved both in the repository system and in producer's production systems, if there is a possibility the SIP identifier could be needed later on.

There are two encoding options, the first one of which is mandatory:

- 1) An identifier shall be located in the root element of the METS document using the OBJID attribute, which identifies the METS object as a whole.
- 2) An SIP identifier may also be expressed in a PREMIS metadata record, if it is intended as a persistent identifier.

If a private identifier system is used, the name of the creator of the package (if the creator is not the producer) may be part of the identifier. This makes it possible to identify the creator, and the OAIS archive is able to contact that organization directly – instead of the producer – if there are technical problems during the ingestion process.

Elements within the METS document may be identified using ID attribute, which uses the XML ID data type for identifiers. Therefore the first character of the ID attribute value must be a letter. OBJID attribute uses data type string and has no restrictions on the first character.

Examples

Package identifier in the root of a METS document:

```
<mets:mets OBJID="urn:uuid:A1B0D67E-2E81-4DF5-9E67-A64CBE366809"  
xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd">  
[...]  
</mets:mets>
```

Publication identifier used as a package identifier in a Dublin Core record embedded in an EPUB publication:

```
<dc:identifier id="pub-id">urn:uuid:A1B0D67E-2E81-4DF5-9E67-A64CBE366809</dc:identifier>  
<meta refines="#pub-id" property="identifier-type" scheme="xsd:string">uuid</meta>
```

Publication identifier shall not be used as package identifiers. An SIP can contain multiple EPUB publications; one EPUB publication can be submitted in multiple SIPs and even if an SIP contains just one publication it may be necessary to re-send the SIP with other package identifier.

Example

```
<metadata xmlns:dc="http://purl.org/dc/elements/1.1/">  
  <dc:identifier id="pub-id">urn:uuid:A1B0D67E-2E81-4DF5-9E67-A64CBE366809</  
dc:identifier>  
  <meta property="dcterms:modified">2011-01-01T12:00:00Z</meta>  
  ...  
</metadata>
```

results in the Package ID:

```
urn:uuid:A1B0D67E-2E81-4DF5-9E67-A64CBE3668092011-01-01T12:00:00Zfed-aa39-  
0583d68d439d/iso-iec-ts-22424-2-2020
```

6.5 Work and publication identifiers

According to the EPUB specification, each EPUB publication shall have a (globally) unique identifier. However, revised publications do not need to have a new standard identifier if only minor changes have been made, such as metadata updates or errata fixes. In such cases, usage of release identifiers (which consist of e.g. ISBN and the publication date) is recommended in the EPUB specification, but not mandatory. This approach is similar to the one in the ISBN standard.

In order to facilitate long-term preservation, each rendition and version of an EPUB publication submitted to an OAIS archive shall have an identifier, and the submission agreement or other guidelines shall specify the identifier systems allowed. If the archive's repository system cannot process EPUB release identifiers (for instance because the system assumes each e-book has its own ISBN or other standard identifier), release identifiers assigned by the publisher should be replaced with identifiers the OAIS archive is able to use during pre-ingest by the producer.

Identifiers belonging to the different manifestations of a work should be included in the metadata records describing these manifestations. In addition, a work identifier may be used in order to facilitate interlinking of manifestations of a work.

NOTE 1 ISBN is universally used for identification of books, but there is no widely used identifier system for textual works. ISTC has not been successful, and following the closure of the ISTC International Centre the future of the identifier is uncertain.

Component parts of EPUB publications shall have separate identifiers if they are submitted as independent publications. For instance, if each chapter of an e-book is submitted as a separate EPUB publication, they shall have their own identifiers even if all chapters (EPUB publications) are sent in the same SIP.

Producers may provide identifiers to fragments within publications, such as paragraphs or sentences within a text. If such identifiers or other methods are used to provide links between for instance a text and an audio version of the text within an EPUB 3 document, the OASIS archive shall maintain the links even when the text or audio file is migrated into a new format, if submission agreement requires that such functionality is retained and if the new audio format allows such linking.

The submission agreement shall specify the encoding of publication identifiers. There are at least two encoding options, of which one shall be selected in the agreement:

- 1) An identifier is included in the Dublin Core metadata embedded in the EPUB publication, as specified in the EPUB publications specification. The EPUB specification requires that the identifier of a publication is provided in the Dublin Core element identifier. The EPUB META element may be used to indicate the identifier type, using an authorized list such as ONIX Code List 5¹¹⁾.
- 2) Identifier is expressed using the <premis:objectIdentifier> element. If so, PREMIS encoding shall specify the identifier type.

NOTE 2 It is possible that digital preservation systems are unable to handle EPUB release identifiers. For instance, their duplicate-check algorithms could expect standard identifiers such as ISBNs for books. It is possible to build a digital preservation system capable of using EPUB release identifiers, but as of this writing no such systems exist in the library sector.

If the repository system cannot process EPUB release identifiers and ISBN or other standard identifiers cannot be applied, it is possible to use custom made identifier systems to identify renditions. The submission agreement shall specify such systems. They should be used with caution, since if and when archived data is transferred to a new OASIS archive, non-standard identifier systems can become a problem.

Producers should have guidelines on how to use identifiers. For instance, if ISBN cannot be used to identify all submitted books, alternative solutions should be clearly specified.

Examples

<https://standards.iteh.ai/catalog/standards/sist/6593bc04-42f9-4fed-aa39-0583d68d439d/iso-iec-ts-22424-2-2020>

Identifiers in a Dublin Core record (including an identifier for the resource itself and its source):

```
<dc:identifier id="isbn-id">urn:isbn:9780101010101</dc:identifier>
  <meta refines="#isbn-id" property="identifier-type" scheme="onix:codelist5">15</meta>

  <dc:source id="src-id">urn:isbn:9780375704024</dc:source>
  <meta refines="#src-id" property="identifier-type" scheme="onix:codelist5">22</meta>
```

NOTE There may be 1-n sources.

Identifier in PREMIS record:

```
<premis:objectIdentifier>
  <premis:objectIdentifierType>urn</premis:objectIdentifierType>
  <premis:objectIdentifierValue>
    URN:ISBN:978-952-222-272-5
  </premis:objectIdentifierValue>
</premis:objectIdentifier>

<mets:fileSec>
  <mets:file ID="filee01" OWNERID="URN:ISBN:978-952-222-272-5" ...>
    [...]
  </mets:file>
</mets:fileSec>
```

Release identifier (unique identifier and the last modification date):

```
<metadata xmlns:dc="http://purl.org/dc/elements/1.1/">
  <dc:identifier id="pub-id">urn:isbn:951-0-18434-9</dc:identifier>
  <meta property="dcterms:modified">2016-05-03T12:00:00Z</meta>
  ...
```

11) <http://www.stison.com/onix/codelists/onix-codelist-5.htm>