
**Digital publishing — EPUB3
preservation —**

**Part 1:
Principles**

Publications numériques — EPUB3 preservation —

Partie 1: Principes

iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

ISO/IEC TS 22424-1:2020

<https://standards.iteh.ai/catalog/standards/iso/1007a1ed-40b8-4be5-9178-ac5789e198d5/iso-iec-ts-22424-1-2020>



Reference number
ISO/IEC TS 22424-1:2020(E)

© ISO/IEC 2020

iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

ISO/IEC TS 22424-1:2020

<https://standards.iteh.ai/catalog/standards/iso/1007a1ed-40b8-4be5-9178-ac5789e198d5/iso-iec-ts-22424-1-2020>



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	9
5 Packaging standards	9
6 Construction of OAIS information packages	11
6.1 Overview	11
6.2 General principles	12
6.2.1 EPUB publications shall be sent to a repository system as well-formed and complete submission information packages (SIPs)	12
6.2.2 Regardless of its type or format, it shall be possible to include any data or metadata in SIPs	14
6.2.3 It should be possible to transfer SIPs by any means, methods, or tools from the submitting organization to the repository system	16
6.2.4 The archive shall have a way to verify the identity of the submitting organization/person, no matter how the information packages are transferred	16
6.2.5 There is no 1:1 relation between OAIS information packages	16
6.2.6 A SIP may contain 0-n EPUB 3 publications, and one EPUB 3 publication may be submitted to the repository system in 1-n SIPs	16
6.2.7 The information package type (in this case, SIP) shall be indicated	16
6.2.8 SIP packaging method shall not restrict the application of any preservation method	17
6.2.9 The packaging method shall not limit the size of the SIP	17
6.3 Identification of information packages and their content	17
6.3.1 It shall be possible to identify any SIP uniquely both during and after the ingest process	17
6.3.2 Information objects (EPUB publications, PREMIS preservation metadata record, etc.) within SIPs shall be identified uniquely and persistently	17
6.3.3 EPUB Fragment Identifiers should not be used in EPUB publications sent to a repository system, unless the submission agreement explicitly allows their use	18
6.4 Structure of information packages	18
6.5 Generic Information package metadata	19
6.5.1 Metadata in information packages shall be based on standards	19
6.5.2 Metadata should allow (automatic) validation of the structure and content of SIPs in terms of integrity, fixity, and syntax	19
6.5.3 It shall be possible to edit metadata in information packages	19
Annex A (informative) EPUB and digital preservation: issues and recommendations	20
Bibliography	24

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

0.1 General

This document facilitates the long-term preservation of EPUB publications by specifying in general level EPUB features which are mandatory for long-term preservation (such as font embedding) and features which should be avoided if possible.

This document can be seen as a stepping stone towards a detailed specification which would be related to EPUB in the same way as PDF/A, specified in ISO 19005-1 to ISO 19005-3, is related to the Portable Document Format (PDF). If and when the EPUB community develops detailed guidelines for the production of archivable EPUB publications, this document could be used as one of the starting points.

Long-term preservation in general requires two things:

- making the object such as EPUB publication fit for preservation – including features to be used and features 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).

Packaging is covered in ISO/IEC TS 22424-2.

0.2 EPUB

The EPUB standard

defines a distribution and interchange format for digital publications and documents. The EPUB® format 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.^[17]

EPUB format was developed by the International Digital Publishing Forum, IDPF, which merged with the World Wide Web Consortium, W3C, in January 2017. Ongoing technical development of the standard, related extension specifications and ancillary deliverables are the responsibility of the W3C EPUB 3 Community Group¹⁾, which published its charter in February 2017. According to the charter,

work on any future major revision of EPUB, e.g. an EPUB 4, is initially out of scope on the presumption that this will be taken up by a new W3C WG as a W3C [Recommendation Track](https://www.w3.org/publishing/track) activity. The EPUB 3 CG will coordinate its work with such new WG, and meanwhile with the existing W3C [Digital Publishing Interest Group](https://www.w3.org/publishing/interest) (DPUB IG).^[23]

The International Digital Publishing Forum, IDPF, has ceased operations as a membership organization in January 2017, and its website²⁾ is now an archive. The latest version of the standard and information about future EPUB developments is available at the Publishing@W3C webpage, <https://www.w3.org/publishing/>.

The specification at hand covers EPUB 3 versions up to EPUB 3.0.1³⁾. EPUB 3.1⁴⁾ was the first major revision of EPUB 3.0.1, but there are no implementations of version 3.1 and therefore it is not covered in this document. The most widely used version of the standard is still 3.0.1. EPUB 3.2, was published in May 2019⁵⁾. Unlike 3.1, it is fully backwards compatible with 3.0.1. It will be covered in the next edition of this document.

1) <https://www.w3.org/publishing/groups/epub3-cg/>

2) <http://idpf.org/>

3) <http://idpf.org/epub/301>

4) <https://www.w3.org/Submission/epub31/>

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