INTERNATIONAL STANDARD

ISO/IEC 23736-6

First edition
2020-02

Information technology — Digital publishing — EPUB 3.0.1 —

Part 6: Canonical fragment identifiers

Technologies de l'information — Publications numériques — EPUB 3.0.1 —

Partie 6: Identificateurs de fragment canoniques

Reference number
ISO/IEC 23736-6:2020(E)

Documents
Document Sharing Hub
© ISO/IEC 2020

Full document details: ISO IEC 23736-6-2020

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

© ISO/IEC 2020 – All rights reserved

Documents
Document Sharing Hub

Full document details: ISO IEC 23736-6-2020

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 (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 the World Wide Web Consortium (W3C) (as EPUB Canonical Fragment Identifiers 1.1) and drafted in accordance with its editorial rules. It was adopted, under the JTC 1 PAS procedure, by Joint Technical Committee ISO/IEC JTC 1, *Information technology*.

A list of all parts in the ISO/IEC 23736 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.

© ISO/IEC 2020 – All rights reserved

Documents
Document Sharing Hub
Full document details: ISO IEC 23736-6-2020

Documents
Document Sharing Hub

Full document details: ISO IEC 23736-6-2020

EPUB Canonical Fragment Identifiers 1.1

Recommended Specification 5 January 2017

This version

http://www.idpf.org/epub/linking/cfi/epub-cfi-20170105.html

Latest version

http://www.idpf.org/epub/linking/cfi/epub-cfi.html

Previous version

http://www.idpf.org/epub/linking/cfi/epub-cfi-20161130.html

Previous recommendation

http://www.idpf.org/epub/linking/cfi/epub-cfi-20140628.html

Document history

Changes to this document
Issues addressed in this revision
Report an issue
Errata

Editors

Peter Sorotokin, Adobe

Garth Conboy, Google Inc.

Brady Duga, Google Inc.

John Rivlin, Google Inc.

Don Beaver, Apple Inc.

Kevin Ballard, Apple Inc.

Alastair Fettes, Apple Inc.

Daniel Weck, DAISY Consortium

Copyright © 2011 International Digital Publishing Forum™

All rights reserved. This work is protected under Title 17 of the United States Code. Reproduction and dissemination of this work with changes is prohibited except with the written permission of the International Digital Publishing Forum (IDPF).

EPUB is a registered trademark of the International Digital Publishing Forum.

Status of this Document

This section describes the status of this document at the time of its publication. Other documents might supersede this document.

AVA Documents
Document Sharing Hub

Full document details: ISO IEC 23736-6-2020

This document was produced by the EPUB Working Group under the EPUB Working Group Charter approved on 8 July 2015.

This document has been reviewed by the IDPF membership and is endorsed by the IDPF Board as a Recommended Specification. This document is considered stable and can be referenced from other specifications and documents.

Feedback on this document can be provided to the EPUB Working Group's mailing list or issue tracker.

This document is governed by the IDPF Policies and Procedures.

Table of Contents

- 1. Overview
  - 1.1. Purpose and Scope
  - 1.2. Terminology
  - 1.3. Typographic Conventions
  - 1.4. Conformance Statements
- 2. EPUB CFI Definition
  - 2.1. Introduction
  - 2.2. Syntax
  - 2.3. Character Escaping
- 3. EPUB CFI Processing
  - 3.1. Path Resolution
    - 3.1.1. Step Reference to Child Element or Character Data (/)
    - 3.1.2. XML ID Assertion ([)
    - 3.1.3. Step Indirection (!)
    - 3.1.4. Character Offset (:)
    - 3.1.5. Temporal Offset (~)
    - 3.1.6. Spatial Offset (@)
    - 3.1.7. Temporal-Spatial Offset (~ + @)
    - 3.1.8. Text Location Assertion ([)
    - 3.1.9. Side Bias ([+ ;s=)
    - 3.1.10. Examples
  - 3.2. Sorting Rules
  - 3.3. Intra-Publication CFIs
  - 3.4. Simple Ranges
  - 3.5. Intended Target Location Correction
- 4. Extending EPUB CFIs
- References

› 1 Overview

› 1.1 Purpose and Scope

This section is informative

This specification, EPUB Canonical Fragment Identifier (epubcfi), defines a standardized method for referencing arbitrary content within an EPUB® Publication through the use of fragment identifiers.

The Web has proven that the concept of hyperlinking is tremendously powerful, but EPUB Publications have been denied much of the benefit that hyperlinking makes possible

Documents
Document Sharing Hub

© ISO/IEC 2020 – All rights reserved
Full document details: ISO IEC 23736-6-2020

because of the lack of a standardized scheme to link into them. Although proprietary schemes have been developed and implemented for individual Reading Systems, without a commonly-understood syntax there has been no way to achieve cross-platform interoperability. The functionality that can see significant benefit from breaking down this barrier, however, is varied: from reading location maintenance to annotation attachment to navigation, the ability to point into any Publication opens a whole new dimension not previously available to developers and Authors.

This specification attempts to rectify this situation by defining an arbitrary structural reference that can uniquely identify any location, or simple range of locations, in an EPUB Publication: the EPUB CFI. The following considerations have strongly influenced the design and scope of this scheme:

- The mechanism used to reference content should be interoperable: references to a reading position created by one Reading System should be usable by another.
- Document references to EPUB content should be enabled in the same way that existing hyperlinks enable references throughout the Web.
- Each location in an EPUB file should be able to be identified without the need to modify the document.
- All fragment identifiers that reference the same logical location should be equal when compared.
- Comparison operations, including tests for sorting and comparison, should be able to be performed without accessing the referenced files.
- Simple manipulations should be possible without access to the original files (e.g., given a reference deep in a file, it should be possible to generate a reference to the start of the file).
- Identifier resolution should be reasonably efficient (e.g., processing of the first chapter is not necessary to resolve a fragment identifier that points to the last chapter).
- References should be able to recover their target locations through parser variations and document revisions.
- Expression of simple, contiguous ranges should be supported.
- An extensible mechanism to accommodate future reference recovery heuristics should be provided.

In the case of both Standard EPUB CFIs and Intra-Publication EPUB CFI, this specification conforms with the guidelines expressed by W3C in Section 6. Best Practices for Fragid Structures [FragIDBestPractices].

In other words, both standard CFI URIs (e.g., "`book.epub#epubcfi(…)`", referred media type "`application/epub+zip`") and intra-publication CFI URIs (e.g., "`package.opf#epubcfi(…)`", referred media type "`application/oebps-package+xml`") make use of a fragment identifier syntax that does not overlap with existing schemes in the context of the aforementioned media types' suffix registrations (i.e., "`-xml`" and "`-zip`").

› 1.2 Terminology

Documents
Document Sharing Hub

© ISO/IEC 2020 – All rights reserved

Full document details: ISO IEC 23736-6-2020

Please refer to [EPUB 3.1] for definitions of EPUB-specific terminology used in this document.

Standard EPUB CFI

A publication-level EPUB CFI links into an EPUB Publication. The path preceding the EPUB CFI references the location of the EPUB Publication.

Intra-Publication EPUB CFI

An intra-publication EPUB CFI allows one Content Document to reference another within the same Rendition of an EPUB Publication. The path preceding the EPUB CFI references the current Rendition's Package Document.

Refer to Intra-Publication CFIs for more information.

› 1.3 Typographic Conventions

The following typographic conventions are used in this specification:

`markup`

All markup (elements, attributes, properties), code (JavaScript, pseudo-code), machine-readable values (string, characters, media types) and file names are in red monospace font.

`markup_link`

Links to markup and code definitions are in underlined red monospace font.

`http://www.idpf.org/`

URIs are in navy blue monospace font.

hyperlink

Hyperlinks are underlined and blue.

[reference]

Normative and informative references are enclosed in square brackets.

Term

Terms defined in the Terminology are in capital case.

Term Link

Links to term definitions have a dotted blue underline.

> Normative element, attribute and property definitions are in blue boxes.

```
Informative markup examples are in light gray boxes.
```

AVA Documents
Document Sharing Hub

Full document details: ISO IEC 23736-6-2020
© ISO/IEC 2020 – All rights reserved