PD ISO/IEC TS 30135-6:2014

BSI Standards Publication

# Information technology — Digital publishing — EPUB3

## Part 6: EPUB Canonical Fragment Identifier

bsi.

...making excellence a habit.™

Document Sharing Hub

Full document details: PD ISO/IEC TS 30135-6:2014

## National foreword

This Published Document is the UK implementation of ISO/IEC TS 30135-6:2014.

The user's attention is drawn to the fact that the UK committee voted against this document at the enquiry stage. While considering the specification of EPUB to be technically sound, the committee had a number of concerns with the editorial quality of the specification, specifically the grouping of terms and definitions, the use of cross-references between parts, and the typographic layout of the document. The UK committee were concerned that this could present obstacles to interpretation by any user who had not participated in the development of EPUB and was unfamiliar with assumptions and interpretations made by other users.

The UK participation in its preparation was entrusted to Technical Committee ICT/-/1, Information systems co-ordination.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2015.

Published by BSI Standards Limited 2015

ISBN 978 0 580 82304 6

ICS 35.240.30

**Compliance with a British Standard cannot confer immunity from legal obligations.**

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 31 January 2015.

### Amendments/corrigenda issued since publication

| Date | Text affected |
|---|---|

Documents
Document Sharing Hub

Full document details: PD ISO/IEC TS 30135-6:2014

TECHNICAL SPECIFICATION

# ISO/IEC TS 30135-6

First edition
2014-11-15

# Information technology — Digital publishing — EPUB3 —

## Part 6: EPUB Canonical Fragment Identifier

*Technologies de l'information — Publications numériques — EPUB3 —*

*Partie 6: Identificateurs de fragment canoniques EPUB*

Reference number
ISO/IEC TS 30135-6:2014(E)

Documents
Document Sharing Hub

© ISO/IEC 2014

Full document details: PD ISO/IEC TS 30135-6:2014

**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2014

All rights reserved. Unless otherwise specified, 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
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland



© ISO/IEC 2014 – All rights reserved

Documents
Document Sharing Hub

Full document details: PD ISO/IEC TS 30135-6:2014

## 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. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In other circumstances, particularly when there is an urgent market requirement for such documents, the joint technical committee may decide to publish an ISO/IEC Technical Specification (ISO/IEC TS), which represents an agreement between the members of the joint technical committee and is accepted for publication if it is approved by 2/3 of the members of the committee casting a vote.

An ISO/IEC TS is reviewed after three years in order to decide whether it will be confirmed for a further three years, revised to become an International Standard, or withdrawn. If the ISO/IEC TS is confirmed, it is reviewed again after a further three years, at which time it must either be transformed into an International Standard or be withdrawn.

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.

ISO/IEC TS 30135 series were prepared by Korean Agency for Technology and Standards (as KS X 6070 series) with International Digital Publishing Forum and were adopted, under a special “fast-track procedure”, by Joint Technical Committee ISO/IEC JTC 1, Information technology, in parallel with its approval by the national bodies of ISO and IEC.

ISO/IEC TS 30135 consists of the following parts, under the general title *Information technology — Document description and processing languages — EPUB 3:*

- *Part 1: Overview*
- *Part 2: Publications*
- *Part 3: Content Documents*
- *Part 4: Open Container Format*
- *Part 5: Media Overlay*
- *Part 6: Canonical Fragment Identifier*
- *Part 7: Fixed-Layout Documents*

Documents
Document Sharing Hub

Full document details: PD ISO/IEC TS 30135-6:2014

# EPUB Canonical Fragment Identifier (epubcfi) Specification

## Recommended Specification 11 October 2011

### This version

http://www.idpf.org/epub/linking/cfi/epub-cfi-20111011.html

### Latest version

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

### Previous version

http://www.idpf.org/epub/linking/cfi/epub-cfi-20110908.html

*A diff of changes from the previous draft is available at this link.*

*Please refer to the errata for this document, which may include some normative corrections.*

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.

## 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

## Table of Contents

1. Overview
    1.1. Purpose and Scope
    1.2. Terminology
    1.3. 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 Node (/)
        3.1.2. XML ID Assertion ([)
        3.1.3. Step Indirection (!)

Documents
Document Sharing Hub

Full document details: PD ISO/IEC TS 30135-6:2014

3.1.4. Terminating Step – Character Offset (:)
3.1.5. Terminating Step – Temporal Offset (~)
3.1.6. Terminating Step – Spatial Offset (@)
3.1.7. Terminating Step – 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 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 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 a 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 required 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.

Documents
Document Sharing Hub