

---

---

**Information technology — Digital  
publishing — EPUB 3.0.1 —**

**Part 5:  
Media overlays**

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

*Partie 5: Superposition de médias*





**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](mailto:copyright@iso.org)  
Website: [www.iso.org](http://www.iso.org)

Published in Switzerland

## 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](http://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](http://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](http://www.iso.org/iso/foreword.html).

This document was prepared by the World Wide Web Consortium (W3C) (as EPUB Media Overlays 3.0.1) 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](http://www.iso.org/members.html).



# EPUB Media Overlays 3.0.1



## Recommended Specification 26 June 2014

### THIS VERSION

<http://www.idpf.org/epub/301/spec/epub-mediaoverlays-20140626.html>

### LATEST VERSION

<http://www.idpf.org/epub3/latest/mediaoverlays>

### PREVIOUS VERSION

<http://www.idpf.org/epub/301/spec/epub-mediaoverlays-20140228.html>

A [diff of changes](#) from the previous version is also available.

Please refer to the [errata](#) for this document, which may include some normative corrections.

Copyright © 2010-2013 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

Marisa DeMeglio, DAISY Consortium

Daniel Weck, DAISY Consortium

### TABLE OF CONTENTS

#### [1. Overview](#)

- [1.1. Purpose and Scope](#)
- [1.2. Relationship to Other Specifications](#)
- [1.3. Terminology](#)
- [1.4. Typographic Conventions](#)
- [1.5. Conformance Statements](#)
- [1.6. Namespace prefix mappings](#)

#### [2. Media Overlay Document Definition](#)

- [2.1. Introduction](#)
- [2.2. Content Conformance](#)
- [2.3. Reading System Conformance](#)
- [2.4. Media Overlay Document Definition](#)
  - [2.4.1. The \*\*smil\*\* Element](#)
  - [2.4.2. The \*\*head\*\* Element](#)
  - [2.4.3. The \*\*metadata\*\* Element](#)
  - [2.4.4. The \*\*body\*\* Element](#)
  - [2.4.5. The \*\*seq\*\* Element](#)
  - [2.4.6. The \*\*par\*\* Element](#)
  - [2.4.7. The \*\*text\*\* Element](#)
  - [2.4.8. The \*\*audio\*\* Element](#)

### [3. Creating Media Overlays](#)

#### [3.1. Overview](#)

#### [3.2. Relationship to the EPUB Content Document](#)

##### [3.2.1. Structure](#)

##### [3.2.2. Granularity](#)

##### [3.2.3. Embedded Audio and Video](#)

##### [3.2.4. Text-to-Speech](#)

#### [3.3. Semantic Inflection](#)

#### [3.4. Associating Style Information](#)

#### [3.5. Packaging](#)

##### [3.5.1. Including Media Overlays](#)

##### [3.5.2. Media Overlays Metadata Vocabulary](#)

### [4. Playback Behaviors](#)

#### [4.1. Loading the Media Overlay](#)

#### [4.2. Basic Playback](#)

##### [4.2.1. Timing and Synchronization](#)

##### [4.2.2. Rendering Audio](#)

##### [4.2.3. Rendering EPUB Content Document Elements](#)

#### [4.3. Interacting with the EPUB Content Document](#)

##### [4.3.1. Navigation](#)

##### [4.3.2. Embedded Audio and Video](#)

##### [4.3.3. Text-to-Speech](#)

#### [4.4. Skippability and Escapability](#)

##### [4.4.1. Skippability](#)

##### [4.4.2. Escapability](#)

### [A. Media Overlays Schema](#)

### [B. Examples of Clock Values](#)

### [C. Acknowledgements and Contributors](#)

### [References](#)

## › 1 Overview

---

### › 1.1 Purpose and Scope

#### **This section is informative**

This specification, EPUB Media Overlays 3.0.1, defines a usage of [\[SMIL\]](#) (Synchronized Multimedia Integration Language), the [Package Document](#), the [EPUB® Style Sheet](#), and the [EPUB Content Document](#) for representation of audio synchronized with the EPUB Content Document.

This specification is one of a family of related specifications that compose EPUB 3, the third major revision of an interchange and delivery format for digital publications based on XML and Web Standards. It is meant to be read and understood in concert with the other specifications that make up EPUB 3:

- The EPUB 3 Overview [\[EPUB3Overview\]](#), which provides an informative overview of EPUB and a roadmap to the rest of the EPUB 3 documents. The Overview should be read first.
- EPUB Publications 3.0.1 [\[Publications301\]](#), which defines the semantics and overarching conformance requirements for each [Rendition](#) of an [EPUB Publication](#).
- EPUB Content Documents 3.0.1 [\[ContentDocs301\]](#), which defines profiles of XHTML, SVG and CSS for use in the context of [EPUB Publications](#).

- EPUB Open Container Format (OCF) 3.0.1 [OCF301], which defines a file format and processing model for encapsulating a set of related resources into a single-file (ZIP) [EPUB Container](#).

## › 1.2 Relationship to Other Specifications

### This section is informative

This specification relies on a subset of [\[SMIL\]](#), from which the EPUB Media Overlays elements and attributes defined in [Media Overlay Document Definition](#) are derived.

## › 1.3 Terminology

### EPUB Publication

A collection of one or more [Renditions](#) conforming to this specification and its [sibling specifications](#), packaged in an [EPUB Container](#).

An EPUB Publication typically represents a single intellectual or artistic work, but this specification and its [sibling specifications](#) do not circumscribe the nature of the content.

### Rendition

A logical document entity consisting of a set of interrelated [resources](#) representing one rendering of an [EPUB Publication](#).

### Publication Resource

A resource that contains content or instructions that contribute to the logic and rendering of at least one [Rendition](#) of an [EPUB Publication](#). In the absence of this resource, the EPUB Publication might not render as intended by the Author. Examples of Publication Resources include a Rendition's [Package Document](#), [EPUB Content Document](#), [EPUB Style Sheets](#), audio, video, images, embedded fonts and scripts.

With the exception of the Package Document itself, the Publication Resources required to render a Rendition are listed in that Rendition's [manifest \[Publications301\]](#) and bundled in the [EPUB Container](#) file (unless specified otherwise in [Publication Resource Locations \[Publications301\]](#)).

Examples of resources that are not Publication Resources include those identified by the Package Document [link \[Publications301\]](#) element and those identified in outbound hyperlinks that resolve outside the [EPUB Container](#) (e.g., referenced from an [\[HTML5\]](#) [a](#) element [href](#) attribute).

### EPUB Content Document

A [Publication Resource](#) that conforms to one of the EPUB Content Document definitions ([XHTML](#) or [SVG](#)).

An EPUB Content Document is a [Core Media Type](#), and may therefore be included in the EPUB Publication without the provision of [fallbacks \[Publications301\]](#).

### XHTML Content Document

An [EPUB Content Document](#) conforming to the profile of [\[HTML5\]](#) defined in [XHTML Content Documents \[ContentDocs301\]](#) .

XHTML Content Documents use the [XHTML syntax](#) of [\[HTML5\]](#).

### SVG Content Document

An [EPUB Content Document](#) conforming to the constraints expressed in [SVG Content Documents \[ContentDocs301\]](#) .

### EPUB Navigation Document

A specialization of the [XHTML Content Document](#), containing human- and machine-readable global navigation information, conforming to the constraints expressed in [EPUB Navigation Documents \[ContentDocs301\]](#) .

### Core Media Type

A set of [Publication Resource](#) types for which no fallback is required. Refer to [Publication Resources \[Publications301\]](#) for more information.

### Package Document

A [Publication Resource](#) carrying bibliographical and structural metadata about a given [Rendition](#) of an [EPUB Publication](#), as defined in [Package Documents \[Publications301\]](#) .

### Manifest

A list of all [Publication Resources](#) that constitute the given [Rendition](#) of a [EPUB Publication](#).

Refer to [manifest \[Publications301\]](#) for more information.

### Spine

An ordered list of [Publication Resources](#), [typically](#) [EPUB Content Documents](#), representing the default reading order of the given [Rendition](#) of an [EPUB Publication](#).

Refer to [spine \[Publications301\]](#) for more information.

### Media Overlay Document

An XML document that associates the [XHTML Content Document](#) with pre-recorded audio narration in order to provide a synchronized playback experience, as defined in this specification.

### Text-to-Speech (TTS)

The rendering of the textual content of an [EPUB Publication](#) as artificial human speech using a synthesized voice.

### EPUB Style Sheet (or Style Sheet)

A CSS Style Sheet conforming to the CSS profile defined in [EPUB Style Sheets \[ContentDocs301\]](#) .

### Viewport

The region of an [EPUB Reading System](#) in which the content of an [EPUB Publication](#) is rendered visually to a [User](#).

### CSS Viewport

A Viewport capable of displaying CSS-styled content.

### EPUB Container (or Container)

The ZIP-based packaging and distribution format for EPUB Publications defined in [OCF301].

### Author

The person(s) or organization responsible for the creation of an EPUB Publication, which is not necessarily the creator of the content and resources it contains.

### User

An individual that consumes an EPUB Publication using an EPUB Reading System.

### EPUB Reading System (or Reading System)

A system that processes EPUB Publications for presentation to a User in a manner conformant with this specification and its sibling specifications .

## › 1.4 Typographic Conventions

The following typographic conventions are used in this specification:

### **markup**

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

### **markup**

Links to markup and code definitions are underlined and in red-orange monospace font. Only the first instance in each section is linked.

### **http://www.idpf.org/**

URIs are in navy blue monospace font.

### **hyperlink**

Hyperlinks are underlined and in blue.

### **[reference]**

Normative and informative references are enclosed in square brackets.

### **Term**

Terms defined in the Terminology are in capital case.

### **Term**

Links to term definitions have a dotted blue underline. Only the first instance in each section is linked.

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

Informative markup examples are in white boxes.

NOTE

Informative notes are in yellow boxes with a "Note" header.

CAUTION

Informative cautionary note are in red boxes with a "Caution" header.

## › 1.5 Conformance Statements

The keywords **MUST**, **MUST NOT**, **REQUIRED**, **SHALL**, **SHALL NOT**, **SHOULD**, **SHOULD NOT**, **RECOMMENDED**, **MAY**, and **OPTIONAL** in this document are to be interpreted as described in [\[RFC2119\]](#).

All sections of this specification are normative except where identified by the informative status label "This section is informative". The application of informative status to sections and appendices applies to all child content and subsections they may contain.

All examples in this specification are informative.

## › 1.6 Namespace prefix mappings

For convenience, the following namespace prefix mappings [\[XMLNS\]](#) are used throughout this specification:

| prefix      | namespace URI                                                           |
|-------------|-------------------------------------------------------------------------|
| <b>epub</b> | <a href="http://www.idpf.org/2007/ops">http://www.idpf.org/2007/ops</a> |

## › 2 Media Overlay Document Definition

---

### › 2.1 Introduction

#### **This section is informative**

Books featuring synchronized audio narration are found in mainstream e-books, educational tools and e-books formatted for persons with print disabilities. In EPUB 3, these types of books are created by using Media Overlay Documents to describe the timing for the pre-recorded audio narration and how

it relates to the EPUB Content Document markup. The file format for Media Overlays is defined as a subset of [SMIL](#), a W3C recommendation for representing synchronized multimedia information in XML.

The Media Overlays feature is designed to be transparent to [EPUB Reading Systems](#) that do not support the feature. The inclusion of Media Overlays in a [Rendition](#) of an EPUB Publication has no impact on the ability of Media Overlay-unaware Reading Systems to render that Rendition as though the Media Overlays are not present.

Although future versions of this specification may incorporate support for video media (e.g., synchronized text/sign-language books), this version supports only synchronizing audio media with the EPUB Content Document.

## › 2.2 Content Conformance

A [Media Overlay Document](#) **MUST** meet all of the following criteria:

### Document Properties

- › It **MUST** meet the conformance constraints for XML documents defined in [XML Conformance \[Publications301\]](#) .
- › It **MUST** be valid to the Media Overlays schema as defined in [Appendix A, Media Overlays Schema](#) and conform to all content conformance constraints expressed in [Media Overlay Document Definition](#).
- › It **MUST** be authored to reflect the structure of the [EPUB Content Document](#) with which it is associated, as stated in [Structure](#) .
- › [Authors](#) **SHOULD** avoid using scripts to control audio and video embedded in the EPUB Content Document, as stated in [Embedded Audio and Video](#).
- › It **SHOULD** use semantic markup where appropriate, as described in [Semantic Inflection](#).
- › It **MUST** be packaged with the [EPUB Publication](#) as shown in [Packaging](#).

### File Properties

- › The Media Overlay Document filename **SHOULD** use the file extension **.smil**.

## › 2.3 Reading System Conformance

[EPUB Reading System](#) support for Media Overlays is **OPTIONAL**. A Reading System that supports Media Overlays **MUST** meet the following criteria:

- › It **MUST** process the Media Overlay Document in conformance with all Reading System conformance constraints expressed in [Media Overlay Document Definition](#).
- › It **MUST** support [XHTML Content Documents](#), and it **MAY** support [SVG Content Documents](#).
- › It **MUST** render Media Overlay elements as described in [Basic Playback](#).

- › It **MUST** allow User navigation while a Media Overlay is being played, as discussed in [Navigation](#).
- › It **MUST** adhere to rules regarding referenced audio and video embedded in the EPUB Content Document, as stated in [Embedded Audio and Video](#).
- › Text-to-Speech (TTS)-capable Reading Systems **SHOULD** conform to [Reading System Text-to-Speech Conformance Requirements \[Publications301\]](#) .
- › It **SHOULD** offer the skippability and escapability features described in [Skippability and Escapability](#).

A Reading System that does not support Media Overlays **MUST** meet the following criteria:

- › It **MUST** ignore both the **media-overlay** attribute on manifest item elements and the manifest item elements where the **media-type** attribute value equals **application/smil+xml**.

## › 2.4 Media Overlay Document Definition

All elements [XML] defined in this section are in the <http://www.w3.org/ns/SMIL> namespace [XMLNS] unless otherwise specified.

### › 2.4.1 The **smil** Element

The **smil** element **MUST** be the root element of all Media Overlay Documents.

#### Element Name

**smil**

#### Usage

The **smil** element is the root element of the Media Overlay Document.

#### Attributes

**version** [required]

Specifies the version number of the [SMIL] specification to which the Media Overlay adheres.

This attribute **MUST** have the value "3.0" to indicate compliance with this version of the specification.

**id** [optional]

The ID [XML] of this element, which **MUST** be unique within the document scope.

**epub:prefix** [optional]

Declares additional metadata vocabulary prefixes.

Refer to [Semantic Inflection](#) for more information.

## Content Model

In this order: **head** [optional], **body** [required]

### › 2.4.2 The **head** Element

The **head** element is the container for metadata in the Media Overlay Document, and consists of zero or one child **metadata** element.

#### Element Name

**head**

#### Usage

The **head** element is the optional first child of the **smil** element.

#### Attributes

None.

#### Content Model

**metadata** [0 or 1].

As this specification defines no metadata properties that must occur in the Media Overlay Document, the **head** element is optional.

### › 2.4.3 The **metadata** Element

The **metadata** element represents metadata for the Media Overlay Document. The **metadata** element is an extension point that allows the inclusion of metadata from any metainformation structuring language.

#### Element Name

**metadata**

#### Usage

As a child of the **head** element.

#### Attributes

None.

#### Content Model

[0 or more] elements from any namespace.

This specification defines no metadata properties that **MUST** occur in the Media Overlay Document; the **metadata** element is provided for custom metadata requirements.

## › 2.4.4 The **body** Element

The **body** element is the starting point for the presentation contained in the Media Overlay Document. It contains the main sequence of **par** and **seq** elements.

### Element Name

**body**

### Usage

The **body** element is the required second child of the **smil** element.

### Attributes

**epub:type** [optional]

An expression of the structural semantics of the corresponding element in the [EPUB Content Document](#).

The value is a whitespace separated list of [property \[Publications301\]](#) types. Refer to [Semantic Inflection](#) for more information.

**id** [optional]

The ID [\[XML\]](#) of this element, which **MUST** be unique within the document scope.

**epub:textref** [optional]

The relative IRI reference [\[RFC3987\]](#) of the corresponding EPUB Content Document, including a fragment identifier that references the specific element as per the [\[XPTRSH\]](#).

### Content Model

In any order: **seq** [0 or more] Or **par** [0 or more]

At least one **par** or **seq** is required.

## › 2.4.5 The **seq** Element

The **seq** element contains media objects which are to be rendered sequentially.

### Element Name

**seq**

### Usage

One or more **seq** elements **MAY** occur as children of the **body** element and of the **seq** element.

### Attributes

**epub:type** [optional]

An expression of the structural semantics of the corresponding element in the [EPUB Content Document](#).

The value is a whitespace separated list of [property](#) [Publications301] types. Refer to [Semantic Inflection](#) for more information.

**id** [optional]

The ID [XML] of this element, which **MUST** be unique within the document scope.

**epub:textref** [required]

The relative IRI reference [RFC3987] of the corresponding EPUB Content Document, including a fragment identifier that references the specific element as per the [XPTRSH].

### Content Model

In any order: **seq** [0 or more] Or **par** [0 or more].

At least one **par** or **seq** is required.

## › 2.4.6 The **par** Element

The **par** element contains media objects which are to be rendered in parallel.

### Element Name

**par**

### Usage

One or more **par** elements **MAY** occur as children of the **body** and **seq** elements.

### Attributes

**epub:type** [optional]

An expression of the structural semantics of the corresponding element in the [EPUB Content Document](#).

The value is a whitespace separated list of [property](#) [Publications301] types. Refer to [Semantic Inflection](#) for more information.

**id** [optional]

The ID [XML] of this element, which **MUST** be unique within the document scope.

## Content Model

In any order: **text** [required] and **audio** [optional]

The **audio** element is optional only if its sibling **text** element refers to audio or video media (see [Embedded Audio and Video](#)), or to textual content intended for rendering via [Text-to-Speech \(TTS\)](#).

› 2.4.7 The **text** Element

The **text** element references an element in the [EPUB Content Document](#). A **text** element typically refers to a textual element, but can also refer to other EPUB Content Document media elements (see [Embedded Audio and Video](#)).

## Element Name

**text**

## Usage

As a required child of the **par** element.

## Attributes

**src** [required]

The relative IRI reference [\[RFC3987\]](#) of the corresponding EPUB Content Document, including a fragment identifier that references the specific element as per the [\[XPTRSH\]](#).

**id** [optional]

The ID [\[XML\]](#) of this element, which **MUST** be unique within the document scope.

## Content Model

Empty.

› 2.4.8 The **audio** Element

The **audio** element represents a clip of audio media.

## Element Name

**audio**

## Usage

A required child of the **par** element unless its sibling **text** element refers to audio or video media, in which case it is optional (see [Embedded Audio and Video](#)).

## Attributes

**id** [optional]

The ID [XML] of this element, which **MUST** be unique within the document scope.

**src** [required]

The relative or absolute IRI reference [RFC3987] of an audio file. The audio file **MUST** be one of the audio formats listed in the [Core Media Types \[Publications301\]](#) table.

**clipBegin** [optional]

A clock value that specifies the offset into the physical media corresponding to the start point of an audio clip.

Clock values are a subset of [SMIL clock values](#), defined in [SMIL]. See [Appendix B, Examples of Clock Values](#).

**clipEnd** [optional]

A clock value that specifies the offset into the physical media corresponding to the end point of an audio clip.

Clock values are a subset of [SMIL clock values](#), defined in [SMIL]. See [Appendix B, Examples of Clock Values](#).

The chronological offset of the terminating position **MUST** be after the starting offset specified in the **clipBegin** attribute.

## Content Model

Empty.

## › 3 Creating Media Overlays

### › 3.1 Overview

#### **This section is informative**

A pre-recorded narration of a publication can be represented as a series of audio clips, each corresponding to part of the [EPUB Content Document](#). A single audio clip, for example, typically represents a single phrase or paragraph, but infers no order relative to the other clips or to the text of a document. Media Overlays solve this problem of synchronization by tying the structured audio narration to its corresponding text (or other media) in the EPUB Content Document using [SMIL](#) markup. Media Overlays are, in fact, a simplified subset of SMIL 3.0 that allow the playback sequence of these clips to be defined.

The SMIL elements primarily used for structuring Media Overlays are **body** (used for the main sequence), **seq** (sequence) and **par** (parallel). (Refer to [Media Overlay Document Definition](#) for more information on these and other SMIL elements.)

The **par** element is the basic building block of an Overlay and corresponds to a phrase in the EPUB Content Document. The element provides two key pieces of information for synchronizing content: 1) the audio clip containing the narration for the phrase; and 2) a pointer to the associated EPUB Content Document fragment. The **par** element uses two media element children to represent this information: an **audio** element and a **text** element. Since **par** elements render their children in parallel, the audio clip and EPUB Content Document fragment are played at the same time, resulting in a synchronized presentation.

The **text** element **src** attribute references the associated phrase, sentence, or other segment of the EPUB Content Document by its IRI reference. The **audio** element **src** attribute similarly references the location of the corresponding audio clip, and adds the optional **clipBegin** and **clipEnd** attributes to indicate a specific offset within the clip.

*The following example shows the Media Overlays markup for a single phrase or sentence.*

```
<par>
  <text src="chapter1.xhtml#sentence1"/>
  <audio src="chapter1_audio.mp3" clipBegin="23s" clipEnd="30s"/>
</par>
```

**par** elements are placed together sequentially to form a series of phrases or sentences. Not every element of the EPUB Content Document will have a corresponding **par** element in the Media Overlay, only those relevant to the audio narration.

*The following example shows a basic Media Overlay Document containing a sequence of phrases. The **body** element acts as the main sequence for the whole document.*

```
<smil xmlns="http://www.w3.org/ns/SMIL"
  version="3.0">
  <body>
    <par id="par1">
      <text src="chapter1.xhtml#sentence1"/>
      <audio src="chapter1_audio.mp3" clipBegin="0s"
clipEnd="10s"/>
    </par>
    <par id="par2">
      <text src="chapter1.xhtml#sentence2"/>
      <audio src="chapter1_audio.mp3" clipBegin="10s"
clipEnd="20s"/>
    </par>
    <par id="par3">
      <text src="chapter1.xhtml#sentence3"/>
      <audio src="chapter1_audio.mp3" clipBegin="20s"
clipEnd="30s"/>
    </par>
  </body>
</smil>
```

**par** elements can also be added to **seq** elements to define more complex structures such as parts and chapters (see [Structure](#)).

## › 3.2 Relationship to the EPUB Content Document

**NOTE** In this section, the EPUB Content Document is assumed to be an XHTML Content Document. While Media Overlays can be used with SVG Content Documents, playback behavior might not be consistent and therefore interoperability is not guaranteed.

### › 3.2.1 Structure

The ordering of the Media Overlay elements **MUST** match the default reading order of the EPUB Content Document. The **par** element represents phrases, and the **seq** element (sequence) represents nested EPUB Content Document containers such as sections, asides, headers, and footnotes. **seq** children **MUST** be other **seq** or **par** elements. Each **seq** element **MUST** contain an **epub:textref** attribute which references the corresponding EPUB Content Document element by IRI reference.

*The following example shows a Media Overlay Document with nested **seq** elements, representing a chapter with both a section header and a sidebar, which itself has a nested figure.*

```
<smil xmlns="http://www.w3.org/ns/SMIL"
      xmlns:epub="http://www.idpf.org/2007/ops"
      version="3.0">
  <body>

    <!-- a chapter -->
    <seq id="id1" epub:textref="chapter1.xhtml#sectionstart"
epub:type="chapter">

      <!-- the section title -->
      <par id="id2">
        <text src="chapter1.xhtml#section1_title"/>
        <audio src="chapter1_audio.mp3" clipBegin="0:23:23.84"
clipEnd="0:23:34.221"/>
      </par>

      <!-- some sentences in the chapter -->
      <par id="id3">
        <text src="chapter1.xhtml#text1"/>
        <audio src="chapter1_audio.mp3" clipBegin="0:23:34.221"
clipEnd="0:23:59.003"/>
      </par>
      <par id="id4">
        <text src="chapter1.xhtml#text2"/>
        <audio src="chapter1_audio.mp3" clipBegin="0:23:59.003"
clipEnd="0:24:15.000"/>
      </par>

      <!-- an informational sidebar -->
      <seq id="id5" epub:textref="chapter1.xhtml#sidebar"
epub:type="sidebar">
        <par id="id6">
          <text src="chapter1.xhtml#sidebartitle"/>
          <audio src="chapter1_audio.mp3"
clipBegin="0:24:15.000" clipEnd="0:24:18.123"/>
        </par>

        <!-- a figure within the sidebar -->
```

```

        <seq id="id7" epub:textref="chapter1.xhtml#figure">
            <par id="id8">
                <text src="chapter1.xhtml#photo"/>
                <audio src="chapter1_audio.mp3"
clipBegin="0:24:18.123" clipEnd="0:24:28.764"/>
            </par>
            <par id="id9">
                <text src="chapter1.xhtml#caption"/>
                <audio src="chapter1_audio.mp3"
clipBegin="0:24:28.764" clipEnd="0:24:50.010"/>
            </par>
        </seq>

        <!-- some sentences in the sidebar -->
        <par id="id10">
            <text src="chapter1.xhtml#sidebartext1"/>
            <audio src="chapter1_audio.mp3"
clipBegin="0:24:50.010" clipEnd="0:25:28.530"/>
        </par>
        <par id="id11">
            <text src="chapter1.xhtml#sidebartext2"/>
            <audio src="chapter1_audio.mp3"
clipBegin="0:25:28.530" clipEnd="0:25:45.515"/>
        </par>
    </seq>

    <!-- more sentences in the chapter (outside the sidebar) -->
    <par id="id12">
        <text src="chapter1.xhtml#text3"/>
        <audio src="chapter1_audio.mp3" clipBegin="0:25:45.515"
clipEnd="0:26:30.203"/>
    </par>
    <par id="id13">
        <text src="chapter1.xhtml#text4"/>
        <audio src="chapter1_audio.mp3" clipBegin="0:26:30.203"
clipEnd="0:27:15.000"/>
    </par>

</seq>
</body>
</smil>

```

**NOTE**

The reason for grouping structures like sidebars, section headers, figures, tables, and footnotes in a **seq** element is so that their start and end positions can be identified during playback. Reading Systems can then offer playback options tailored to the layout of the given Rendition, such as jumping past a long sidebar, turning off rendering of page break announcements (see [Skippability and Escapability](#)), or customizing the reading mode to suit structures such as tables.

The following example shows the EPUB Content Document that corresponds to the [previous Media Overlay example](#).

```

<html xmlns="http://www.w3.org/1999/xhtml"
xmlns:epub="http://www.idpf.org/2007/ops"

```

```

    xml:lang="en"
    lang="en">
<head>
    <title>Media Overlays Example of EPUB Content Document</title>
</head>
<body id="sec1">
    <section id="sectionstart" epub:type="chapter">
        <h1 id="section1_title">The Section Title</h1>
        <p id="text1">The first phrase of the main text body.</p>
        <p id="text2">The second phrase of the main text body.</p>
        <aside id="sidebar" epub:type="sidebar">
            <h2 id="sidebartitle">The Sidebar Title</h2>
            <figure id="figure">
                
                <figcaption id="caption">The photo
caption</figcaption>
            </figure>
            <p id="sidebartext1">A phrase in the sidebar.</p>
            <p id="sidebartext2">Another phrase in the sidebar</p>
        </aside>
        <p id="text3">The third phrase of the main text body.</p>
        <p id="text4">The fourth phrase of the main text body.</p>
    </section>
</body>
</html>

```

### › 3.2.2 Granularity

#### This section is informative

Media Overlay **text** elements' **src** attributes refer to [EPUB Content Document](#) elements by their IDs [\[XML\]](#). The granularity level of the Media Overlay therefore depends on how the EPUB Content Document is marked up. If the finest level of markup is at the paragraph level, then that is the finest possible level at which Media Overlay synchronization can be authored. Likewise, if sub-paragraph markup is available, such as [\[HTML5\]](#) **span** elements representing phrases or sentences, then finer granularity is possible in the Media Overlay. Finer granularity gives [Users](#) more precise results for synchronized playback when navigating by word or phrase and when searching the text, but increases the file size of the Media Overlay Documents.

### › 3.2.3 Embedded Audio and Video

Any [EPUB Content Document](#) associated with a Media Overlay **MAY** contain embedded media such as video, audio, and images. The Media Overlay **text** element **MAY** be used in such instances to reference the embedded media by its ID [\[XML\]](#) value.

When a **text** element references embedded media that contains audio, no **audio** sibling element is required, though one is allowed.

Authors **SHOULD** avoid using scripts to control playback of referenced embedded EPUB Content Document media, as this may conflict with Media Overlays playback behavior.

### › 3.2.4 Text-to-Speech

This specification allows the use of [Text-to-Speech \(TTS\)](#) in addition to pre-recorded audio clips. When a Media Overlay **text** element with no **audio** sibling element references an element within the target [EPUB Content Document](#), the contents of that referenced element **MUST** be appropriate for rendering via TTS. For example, it could be a textual EPUB Content Document element or contain a text fallback.

### › 3.3 Semantic Inflection

In order to express semantic inflections, the [epub:type attribute \[ContentDocs301\]](#) **MAY** be attached to Media Overlay **par** , **seq** , and **body** elements.

Values for the Media Overlay **epub:type** attribute are constrained identically to the **epub:type** attribute in [EPUB Content Documents](#). Refer to [XHTML Semantic Inflection \[ContentDocs301\]](#) for details.

The **epub:type** attribute facilitates Reading System behavior appropriate for the semantic type(s) indicated. Examples of these behaviors are [Skippability and Escapability](#) and [Table Reading Mode](#).

*The following example shows the semantic markup for a Media Overlay containing a sidebar.*

```
<smil xmlns="http://www.w3.org/ns/SMIL"
      xmlns:epub="http://www.idpf.org/2007/ops"
      version="3.0">
  <body>
    <seq id="id1" epub:textref="chapter1.xhtml#sidebar"
epub:type="sidebar">
      <par id="id2">
        <text src="chapter1.xhtml#sidebartitle"/>
        <audio src="chapter1_audio.mp3" clipBegin="0:24:15.000"
clipEnd="0:24:18.123"/>
      </par>
      <par id="id3">
        <text src="chapter1.xhtml#sidebartext1"/>
        <audio src="chapter1_audio.mp3" clipBegin="0:24:18.123"
clipEnd="0:24:38.530"/>
      </par>
      <par id="id4">
        <text src="chapter1.xhtml#sidebartext2"/>
        <audio src="chapter1_audio.mp3" clipBegin="0:24:38.530"
clipEnd="0:25:00.515"/>
      </par>
    </seq>
  </body>
</smil>
```

This specification adopts the vocabulary association mechanisms defined in [Vocabulary Association \[ContentDocs301\]](#) unmodified. Terms from the [default vocabulary \[ContentDocs301\]](#) **MUST** be used unprefixed in Overlay Documents.

### › 3.4 Associating Style Information

Visual rendering information for the currently-playing EPUB Content Document element **MAY** be expressed in the EPUB Style Sheet using author-defined classes. These author-defined class names **SHOULD** be declared in the Package Document metadata, using the metadata properties **active-class** and **playback-active-class**. The class names are then discoverable by Reading Systems.

This example demonstrates how authors may associate style information with the currently-playing EPUB Content Document.

#### NOTE

Although this example uses the class names **-epub-media-overlay-active** and **-epub-media-overlay-playing**, any class names are permitted. The class names chosen may be used along with any supported CSS features.

*The author-defined CSS class names, declared using the metadata properties [active-class](#) and [playback-active-class](#) in the Package Document:*

```
<meta property="media:active-class">-epub-media-overlay-active</meta>
<meta property="media:playback-active-class">-epub-media-overlay-
playing</meta>
```

*The EPUB Style Sheet containing the author-defined class names:*

```
/* emphasize the active element */
.-epub-media-overlay-active {
    background-color: yellow;
    color: black !important;
}

/* fade out the inactive text */
html.-epub-media-overlay-playing * {
    color: gray;
}
```

*The relevant EPUB Content Document excerpt:*

```
<html>
...
<span id="txt1">This is the first phrase.</span>
<span id="txt2">This is the second phrase.</span>
<span id="txt3">This is the third phrase.</span>
...
</html>
```

In this example, the Reading System would apply the author-defined **-epub-media-overlay-active** class to each text element in the EPUB Content Document as it became active during playback. Conversely, the class name is removed when the element is no longer active. The User would see each EPUB Content Document element styled with a yellow background for the duration of that element's playback.

The Reading System would also apply the author-defined `-epub-media-overlay-playing` class to the document element of the EPUB Content Document when Media Overlays playback begins. The class name is removed when playback stops. In the case of an XHTML Content Document, the class name would be applied to the `html` element. In the case of an SVG Content Document, it would be applied to the `svg` element. The User would see all the inactive text elements turn gray during Media Overlays playback. When playback stopped, the elements' colors would return to their defaults.

## › 3.5 Packaging

### › 3.5.1 Including Media Overlays

Manifest [item elements](#) [Publications301] in the [Package Document](#) **MAY** specify a Media Overlay via the `media-overlay` attribute. Media Overlays are themselves manifest items and **MUST** be referred to by their IDs [XML].

*The following example shows how to include Media Overlays in the manifest of a Package Document.*

```
<manifest>
  <item id="ch1"
        href="chapter1.xhtml"
        media-type="application/xhtml+xml"
        media-overlay="ch1_audio"/>
  <item id="ch1_audio"
        href="chapter1_audio.smil"
        media-type="application/smil+xml"/>
</manifest>
```

Manifest items which refer to Media Overlays **MUST** have the media-type `application/smil+xml` as specified in [Core Media Types](#) [Publications301] .

The `media-overlay` attribute **MUST** be attached to manifest `item` elements that reference [EPUB Content Documents](#) only.

A single Media Overlay file **MAY** refer to more than one EPUB Content Document, but an EPUB Content Document **MUST NOT** be referenced by more than one Media Overlay file.

Each EPUB Content Document manifest `item` is **NOT REQUIRED** to have a Media Overlay associated with it. If an EPUB Content Document is wholly or partially referenced by a Media Overlay, then its manifest `item` entry **MUST** indicate this via the `media-overlay` attribute.

This is a forwards-compatible addition: 2.0 Reading Systems **MAY** safely ignore the `media-overlay` attribute and process documents in their normal fashion.

### › 3.5.2 Media Overlays Metadata Vocabulary

The following tables both define a set of properties for use in [Package Document](#) metadata and constitute a referenceable vocabulary.

The base IRI for referencing this vocabulary is <http://www.idpf.org/epub/vocab/overlays/#>.

## NOTE

The prefix [media](#) is reserved by [Publications301] for the inclusion of these properties in package metadata.

**active-class**

|                   |                                                                                                |
|-------------------|------------------------------------------------------------------------------------------------|
| Description:      | Author-defined CSS class name to apply to the currently-playing EPUB Content Document element. |
| Allowed value(s): | <b>xsd:string</b>                                                                              |
| Cardinality:      | zero or one                                                                                    |
| Example:          | <code>&lt;meta property="media:active-class"&gt;-epub-media-overlay-active&lt;/meta&gt;</code> |

**duration**

|                   |                                                                                                                                                                                                                                  |
|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description:      | The duration of the entire presentation or of a specific Media Overlay. The specified durations account for the audio clips known at authoring time, and so exclude live streaming from external resources and speech synthesis. |
| Allowed value(s): | A clock value.<br><br>Clock values are a subset of <a href="#">SMIL clock values</a> , defined in [SMIL]. See <a href="#">Appendix B, Examples of Clock Values</a> .                                                             |
| Cardinality:      | Exactly one for a given <a href="#">Rendition</a> and for each Media Overlay.                                                                                                                                                    |
| Example:          | <code>&lt;meta property="media:duration"&gt;1:36:20&lt;/meta&gt;</code>                                                                                                                                                          |

**narrator**

|                   |                                                                             |
|-------------------|-----------------------------------------------------------------------------|
| Description:      | Name of the narrator.                                                       |
| Allowed value(s): | <b>xsd:string</b>                                                           |
| Cardinality:      | zero or more                                                                |
| Example:          | <code>&lt;meta property="media:narrator"&gt;Joe Speaker&lt;/meta&gt;</code> |

**playback-active-class**

|                   |                                                                                                                 |
|-------------------|-----------------------------------------------------------------------------------------------------------------|
| Description:      | Author-defined CSS class name to apply to the EPUB Content Document's document element when playback is active. |
| Allowed value(s): | <b>xsd:string</b>                                                                                               |
| Cardinality:      | zero or one                                                                                                     |
| Example:          | <code>&lt;meta property="media:playback-active-class"&gt;-epub-media-overlay-playing&lt;/meta&gt;</code>        |

The Package Document **MUST** include the **duration** of each Media Overlay as well as of the entire Rendition. The Package Document **MAY** include **narrator** information, as well, in particular when each Media Overlay has its own narrator or there is one narrator specified for the entire Rendition. The

Package Document **MAY** also include an author-defined CSS class name to be applied to the currently-playing EPUB Content Document element.

When a **meta** element is specific to a single Media Overlay Document, the **refines** attribute is used to reference which one. A **meta** element without a **refines** attribute is considered to be about the entire Rendition. The **active-class** and **playback-active-class** properties **MUST NOT** be used in conjunction with a **refines** attribute, as it is always considered to apply to the entire Rendition.

*The following example shows a Package Document with metadata about Media Overlays.*

```
<package>
  <metadata>
    ...
    <meta property="media:duration"
refines="#ch1_audio">0:32:29</meta>
    <meta property="media:duration"
refines="#ch2_audio">0:34:02</meta>
    <meta property="media:duration"
refines="#ch3_audio">0:29:49</meta>
    <meta property="media:duration">1:36:20</meta>
    <meta property="media:narrator">Joe Speaker</meta>
    <meta property="media:active-class">-epub-media-overlay-
active</meta>
    <meta property="media:playback-active-class">-epub-media-
overlay-playing</meta>
    ...
  </metadata>
  ...
</package>
```

## › 4 Playback Behaviors

---

### › 4.1 Loading the Media Overlay

When an EPUB Reading System loads a Package Document, it **MUST** refer to the manifest item elements' **media-overlay** attributes to discover the corresponding Media Overlays for EPUB Content Documents. Playback **MUST** start at the Media Overlay element which corresponds to the desired EPUB Content Document starting point. Note that the start of an EPUB Content Document **MAY** correspond to an element at the start or in the middle of a Media Overlay. When the Media Overlay Document has finished playing, the Reading System **SHOULD** load the next EPUB Content Document (as specified in the Package Document spine) and also load its corresponding Media Overlay Document, provided that one is given.

### › 4.2 Basic Playback

#### › 4.2.1 Timing and Synchronization

Reading Systems **MUST** render immediate children of the **body** element in a sequence. A **seq** element's children **MUST** be rendered in sequence, and playback completes when the last child has finished playing. A **par** element's children **MUST** be rendered in parallel (with each starting at the same time), and playback completes when all the children have finished playing. When the **body** element's last child has finished playing, playback of the Media Overlay Document is done.

#### › 4.2.2 Rendering Audio

When presented with a Media Overlay **audio** element, Reading Systems **MUST** play the audio resource referenced by the **src** attribute, starting at the clip offset time given by the **clipBegin** attribute and ending at the clip offset time given by the **clipEnd** attribute. The following rules **MUST** be observed:

- If **clipBegin** is not specified, its value is assumed to be "0".
- If **clipEnd** is not specified, its value is assumed to be the full duration of the physical media.
- If **clipEnd** exceeds the full duration of the physical media, then its value is assumed to be the full duration of the physical media.

User-controllable audio playback options should include timescale modification, in which the playback rate is altered without distorting the pitch. The suggested range is half-speed to double-speed.

#### › 4.2.3 Rendering EPUB Content Document Elements

When presented with a Media Overlay **text** element, Reading Systems **SHOULD** ensure the EPUB Content Document element referenced by the **src** attribute is visible in the Viewport. During Media Overlays playback, Reading Systems with a CSS Viewport **SHOULD** add the class names given by the metadata properties **active-class** and **playback-active-class** to the appropriate elements in the EPUB Content Document. Conversely, the class names **SHOULD** be removed when the playback state changes, as described in Associating Style Information.

The **active-class** and **playback-active-class** metadata properties are optional, and if omitted, Reading System behavior is implementation-specific.

### › 4.3 Interacting with the EPUB Content Document

#### › 4.3.1 Navigation

Because the Media Overlay is closely linked to the EPUB Content Document, it is very easy for Reading Systems to locate a position in the EPUB Content Document based on the current position in the Media Overlay playback. If the User pauses synchronized playback and navigates to a different part of the EPUB Publication, synchronized playback **MUST** resume at that point. For example, if a specific page number in the EPUB Content Document is the desired location, then this same point is located in the Media Overlay and playback started there.

This same approach allows for synchronizing the Media Overlay playback with User selection of a navigation points in the EPUB Navigation Document. The Reading System loads the Media Overlay

for that file and finds the correct point for starting playback based on the ID [XML] of the navigation point target.

## NOTE

A Media Overlay Document may also be associated directly with a Navigation Document in order to provide synchronized playback of its contents, regardless of whether the [XHTML Content Document](#) in which it resides is included in the [spine](#). The Reading System should keep playback of the Navigation Document's Media Overlay synchronized with the User's current position in the EPUB Content Document.

## › NOTE

Media Overlay Document elements may be associated with EPUB Content Document structures such as tables. Reading Systems should ensure that Media Overlay playback remains synchronized with User navigation of table rows and cells. The Reading System may also play the corresponding table header preceding the contents of the cell.

## › 4.3.2 Embedded Audio and Video

An [EPUB Content Document](#) with which a Media Overlay is associated **MAY** itself contain embedded video and audio media, which **MAY** be pointed to by Media Overlay elements. Unlike text and images, video and audio media has an intrinsic duration. Consequently, when a Reading System renders the synchronization described by a Media Overlay, the default playback behavior of audio and video media embedded within the associated EPUB Content Document **MUST** be overridden.

Note that the rules below apply only to *referenced* [\[HTML5\] video](#) or [audio](#) elements within the associated EPUB Content Document. That is to say, the rules apply to only those elements pointed to by [text](#) elements within the Media Overlay (i.e., via the [src](#) attribute). Embedded media that is not referenced by Media Overlay elements is not subject to these rules.

- All referenced audio and video media embedded within an EPUB Content Document **MUST** have their public playback interface deactivated (typically: play/pause control, time slider, volume level, etc.). This behavior is required to avoid interference between the scheduled playback sequence defined by the Media Overlay, and the arbitrary playback behavior due to [User](#) interaction or script execution. As a result, when the Reading System is in playback mode, it **SHOULD**:
  - Hide the individual video/audio UI controls from the page, which overrides the default behavior defined by the [\[HTML5\] controls](#) attribute.
  - Prevent scripts embedded within the EPUB Content Document from invoking the JavaScript audio/video playback API (i.e., authored as part of the default behavior). It is **RECOMMENDED** that content producers avoid publishing embedded scripts dedicated to controlling the playback of embedded audio/video media. The published Media Overlay can then retain full control of the synchronized presentation without any risk of interference from script-enabled custom behaviors.
- All referenced audio and video media embedded within an EPUB Content Document **MUST** be initialized to their "stopped" state, and be ready to be played from the zero-position within their content stream (possibly displaying the image specified using the [\[HTML5\] poster](#) attribute). This requirement overrides the default behavior defined by the [\[HTML5\] autoplay](#) attribute.

- When an EPUB Content Document element becomes active, the [EPUB Style Sheet](#) visual highlighting rules apply regardless of the content type referred to by that element's `src` attribute (e.g., the CSS class name defined by the `active-class` metadata property **SHOULD** be applied to visible video and audio player controls within the host EPUB Content Document).
- In addition to the default behavior of Media Overlay activation for textual fragments and images, audio and video playback **MUST** be started and stopped according to the duration implied by the authored Media Overlay synchronization (as per the standard [\[SMIL\]](#) timing model). There are two possible scenarios:
  - When a Media Overlay `text` element has no `audio` sibling within its `par` parent container, the referenced EPUB Content Document audio or video media **MUST** play until it ends, at which point the `text` element's lifespan terminates. In this case, the implicit duration of the `text` element (and by inference, of the parent `par` container) is that of the referenced audio or video clip.
  - When a Media Overlay `text` element has an `audio` sibling within its `par` parent container, the playback duration of the referenced EPUB Content Document audio or video media **MUST** be constrained by the duration of the `audio` sibling. In this case, the actual duration of the parent `par` container is that of the child audio clip, regardless of the duration of the video or audio media pointed to by the `text` element. This behavior may result in embedded video or audio media ending playback prematurely (before reaching its full duration), or ending before the playback of the parallel Media Overlay `audio` is finished (in which case the last-played video frame **SHOULD** remain visible until the parent `par` container finally ends). This behavior is equivalent of the Media Overlay `audio` element implicitly carrying the behavior of the [\[SMIL\]](#) `endsync` attribute.

Furthermore, Reading Systems **SHOULD** expose User controls for the volume levels of each independent audio track (i.e., from the `audio` element of the Media Overlay, and from the embedded audio or video media within the EPUB Content Document), so that audio output can be adjusted to match listeners' requirements. Note that having overlapping audio tracks is typically an authoring-time concern: content producers usually add a layer of audio information over a video track for description purposes. It is **RECOMMENDED** that overlapping audio situations are carefully examined and dealt with at production stage, as Reading Systems are **NOT REQUIRED** to handle simultaneous volume levels in any particular way.

- When a `text` element becomes inactive in the Media Overlay, and when it points to embedded video or audio media, that referenced media **MUST** be reset to its initial "stopped" state, ready to be played from the zero-position within their content stream (possibly displaying the poster image specified using the HTML5 markup).

### › 4.3.3 Text-to-Speech

When a Media Overlay `text` element with no `audio` sibling element references text within the target EPUB Content Document, Reading Systems capable of [Text-to-Speech \(TTS\)](#) **SHOULD** render the referenced text using TTS.

As per Reading System conformance requirements, the speech-related information provided in the target EPUB Content Document **SHOULD** be used to play the audio stream as part of the Media Overlay rendering. See [Reading System Text-to-Speech Conformance Requirements \[Publications301\]](#) .

The Media Overlay `text` element's lifespan corresponds to the rendering time of the associated speech synthesis. The implicit duration of the `text` element (and by inference, of the parent `par`

element) is therefore determined by the execution of the Text-to-Speech engine, and cannot be known at authoring time (factors like speech rate, pauses and other prosody parameters influence the audio output).

## › 4.4 Skippability and Escapability

### › 4.4.1 Skippability

While reading, Users may want to turn on or off certain features of the content, such as sidebars, footnotes, page numbers, or other types of secondary content. This feature is called skippability. Reading Systems **SHOULD** use the semantic information provided by Media Overlay elements' **epub:type** attribute to determine when to offer Users the option of skippable features. In the following example, a Reading System **SHOULD** offer the User the option of turning on and off the page break/page number announcements, which are often cumbersome to listen to.

*The following example shows a Media Overlay Document with a pagebreak.*

```
<smil xmlns="http://www.w3.org/ns/SMIL"
      xmlns:epub="http://www.idpf.org/2007/ops"
      version="3.0">
  <body>
    <!-- a paragraph -->
    <par id="id1">
      <text src="chapter1.xhtml#para1"/>
      <audio src="chapter1_audio.mp3" clipBegin="0:23:22.000"
clipEnd="0:24:15.000"/>
    </par>

    <!-- a page number -->
    <par id="id2" epub:type="pagebreak">
      <text src="chapter1.xhtml#pgbreak1"/>
      <audio src="chapter1_audio.mp3" clipBegin="0:24:15.000"
clipEnd="0:24:18.123"/>
    </par>

    <!-- another paragraph -->
    <par id="id3">
      <text src="chapter1.xhtml#para2"/>
      <audio src="chapter1_audio.mp3" clipBegin="0:24:18.123"
clipEnd="0:25:28.530"/>
    </par>
  </body>
</smil>
```

*The following example shows an EPUB Content Document with a pagebreak.*

```
<html ... >
  ...
  <body>
    <p id="para1">This is the paragraph before the pagebreak ... </p>

    <br id="pgbreak1" epub:type="pagebreak" title="234"/>
```

```

    <p id="para2">This is the paragraph after the pagebreak ...</p>
  </body>
</html>

```

The following selection of terms from the [StructureVocab] for which Reading Systems **SHOULD** offer Users the option of skippability is provided as an informative reference:

- sidebar
- practice
- marginalia
- annotation
- help
- note
- footnote
- rearnote
- pagebreak

Media Overlays **MAY** use additional vocabularies by defining them in the **epub:prefix** attribute on the root **smil** element. Reading System support for skippability based on **epub:type** values **SHOULD NOT** be assumed.

#### › 4.4.2 Escapability

Escapable items are nested structures such as tables, lists, and sidebars that listeners may wish to skip over, continuing to read from the point immediately after the nested structure. The escapability feature differs from the skippability feature in that it does not enable or disable entire types of items, but provides an exit from them (e.g., a User can listen to some of the content before choosing to escape). Reading Systems **SHOULD** allow escaping of nested structures. Reading Systems **MUST** determine the start of nested structures by the value of the **epub:type** attribute (e.g., **glossary**) and **SHOULD** offer Users the option to skip playback of that structure and resume with whatever content comes after it.

*The following example shows the Media Overlay Document for an EPUB Content Document containing a paragraph, a glossary, and another paragraph. A Reading System that supported escapability would give the User the option to interrupt playback of the glossary and continue playing the document paragraphs.*

```

<smil xmlns="http://www.w3.org/ns/SMIL"
  xmlns:epub="http://www.idpf.org/2007/ops"
  version="3.0">
  <body>
    <!-- a paragraph, part of the regular document text -->
    <par id="id1">
      <text src="chapter1.xhtml#para1"/>
      <audio src="chapter1_audio.mp3" clipBegin="0:23:22.000"
clipEnd="0:24:15.000"/>
    </par>

```

```

        <!-- a glossary, which is a nested structure -->
        <seq id="id2" epub:textref="chapter1.xhtml#g0"
epub:type="glossary">
            <par id="id3" epub:type="glossterm">
                <text src="chapter1.xhtml#g1"/>
                <audio src="chapter1_audio.mp3" clipBegin="0:24:15.000"
clipEnd="0:24:18.123"/>
            </par>
            <par id="id4" epub:type="glossdef">
                <text src="chapter1.xhtml#g2"/>
                <audio src="chapter1_audio.mp3" clipBegin="0:24:18.123"
clipEnd="0:25:28.530"/>
            </par>
            <par id="id5" epub:type="glossterm">
                <text src="chapter1.xhtml#g3"/>
                <audio src="chapter1_audio.mp3" clipBegin="0:25:28.530"
clipEnd="0:25:45.515"/>
            </par>
            <par id="id6" epub:type="glossdef">
                <text src="chapter1.xhtml#g4"/>
                <audio src="chapter1_audio.mp3" clipBegin="0:25:45.515"
clipEnd="0:27:04.123"/>
            </par>
        </seq>

        <!-- another paragraph, part of the document text that comes
after the glossary -->
        <par id="id7">
            <text src="chapter1.xhtml#para2"/>
            <audio src="chapter1_audio.mp3" clipBegin="0:27:04.123"
clipEnd="0:27:59.000"/>
        </par>
    </body>
</smil>

```

## › Appendix A. Media Overlays Schema

The schema for Media Overlays is available at [../schema/media-overlay-30.nvdl](https://www.iso.org/standard/70431.html).

Validation using this schema requires a processor that supports [NVDL], [RelaxNG], [ISOSchematron] and [XSD-DATATYPES].

### NOTE

The NVDL schema layer can be substituted by a multi-pass validation using the embedded RELAX NG and ISO Schematron schemas alone.

## › Appendix B. Examples of Clock Values

## This appendix is informative

The following are examples of allowed clock values:

- **5:34:31.396** = 5 hours, 34 minutes, 31 seconds and 396 milliseconds
- **124:59:36** = 124 hours, 59 minutes and 36 seconds
- **0:05:01.2** = 5 minutes, 1 second and 200 milliseconds
- **0:00:04** = 4 seconds
- **09:58** = 9 minutes and 58 seconds
- **00:56.78** = 56 seconds and 780 milliseconds
- **76.2s** = 76.2 seconds = 76 seconds and 200 milliseconds
- **7.75h** = 7.75 hours = 7 hours and 45 minutes
- **13min** = 13 minutes
- **2345ms** = 2345 milliseconds
- **12.345** = 12 seconds and 345 milliseconds

## › Appendix C. Acknowledgements and Contributors

---

### This appendix is informative

EPUB has been developed by the International Digital Publishing Forum in a cooperative effort, bringing together publishers, vendors, software developers, and experts in the relevant standards.

The EPUB 3 specifications were prepared by the International Digital Publishing Forum's EPUB Maintenance Working Group, operating under a charter approved by the membership in May, 2010 under the leadership of:

Gylling, Markus (International Digital Publishing Forum (IDPF)) *Chair*  
 Conboy, Garth (Google Inc.) *Vice-chair*  
 Duga, Brady (Google Inc.) *Vice-chair, Subgroup Lead*  
 McCoy, Bill (International Digital Publishing Forum (IDPF)) *Secretary*  
 Kasdorf, Bill (Apex CoVantage) *Subgroup Lead*  
 MURATA, Makoto (JEPA EPUB Study Group) *Subgroup Lead*  
 Sorotokin, Peter (Adobe) *Subgroup Lead*

Active members of the working group included:

### › IDPF Members

Abrams, Willie (Ingram Digital)  
 Acton, Daniel (Google)  
 Allesi, Ana Maria (HarperCollins)

Amos, Dan (DNAML (DNL eBooks))  
Arany, Steve (John Wiley & Sons)  
Artin, Michael (Barnes & Noble)  
Badger, Brandon (Google)  
Ballard, Kevin (Apple Inc.)  
Beard, Elliot (HarperCollins)  
Belfanti, Paul (Pearson)  
Bell, Graham (EDItEUR)  
Bell, Jeff (Microsoft)  
Bide, Mark (EDItEUR)  
Bogaty, Nick (Adobe)  
Bowers, Micah (Bluefire Productions)  
Brantley, Peter (Internet Archive)  
Breglio, Melissa (Apple Inc.)  
Broome, Karen (Sony)  
Brugge, John (Benetech)  
Carbonell, Oliver (Sony)  
Chang, Phobos (Chinese Foundation for Digitization Technology)  
Chen, Mei-Li (Institute for Information Industry)  
Chen, Peter (ITRI)  
Choi, Soo (HarperCollins)  
Chow, King-Wai (ASTRI (Hong Kong Applied Science & Technology Research Institute))  
Clutter, Mat (Random House)  
Conboy, Garth (Google)  
Cramer, Dave (Hachette Book Group)  
Cronin, Margot (Bowker)  
Daly, Liza (Threepress)  
De Meulemeester, Eric (Jouve/Publishing Dimensions)  
DeMeglio, Marisa (DAISY Consortium)  
Deltour, Romain (DAISY Consortium)  
Dougherty, Casey (Apple Inc.)  
Drake, Jama (Impelsys)  
Duga, Brady (Google)  
Elliott, Ray (Crossway)  
Fahlgren, Keith (Threepress)  
Fain, Guy (Crossway)  
Freese, Eric (Aptara)  
Gardeur, Hadrien (Feedbooks)  
Gold, Eric (Digital Divide Data)  
Goodwin, Jonathan (Appfoundry)  
Gopinath, Anith (Impelsys)  
Gosling, Andreas (Penguin)  
Grazioli, Frank (John Wiley & Sons)  
Gunn, Dave (RNIB)  
Gylling, Markus (International Digital Publishing Forum (IDPF))  
Haas, Matt (Pearson)  
Hadfield, Tom (CourseSmart)  
Hagino, Masaaki (Voyager Japan)  
Hawkins, Kevin (University of Michigan Library)  
Hayashi, Junichi (Voyager Japan)  
Heiberger, Richard (HarperCollins)  
Hepp, Mike (Dartmouth Journal Services)  
Herren, Matthew (BlankPage)  
Hisashi, Saiga (Sharp)  
Hoda, Hisashi (Voyager Japan)  
Howard, William (Easypress)  
Hughes, Dan (Liguori Publications)  
Hulse, Leslie (HarperCollins)

Imsieke, Gerrit (le-tex)  
 Jain, Anupam (Innodata Isogen)  
 Jie, Fan (Gansu DUZHE Digital Sci&Tech)  
 Johnson, Rick (Ingram Digital)  
 Jung, Kanghee (Incube Technologies)  
 Kakar, Samir (Aptara)  
 Kanai, Takeshi (Sony)  
 Kasdorf, Bill (Apex CoVantage)  
 Kasher, Bob (BookMasters and Newgen Imaging)  
 Kato, Kazuyuki (East Co.)  
 Keating, Patrick (Bluefire Productions)  
 Kerscher, George (DAISY Consortium)  
 Kida, Yasuo (Apple Inc.)  
 Kim, Jean (Barnes & Noble)  
 Kim, HyunYoung (Incube Technologies)  
 Kim, Terry (INKA Entworks)  
 Kitagawa, Masahiro (Impress Holdings)  
 Koike, Toshiaki (Voyager Japan)  
 Kok, Dan (Crossway)  
 Kotrch, Steve (Simon & Schuster)  
 Larroque, Benoit (Feedbooks)  
 Levantovsky, Vladimir (Monotype Imaging)  
 Lu, Cho-Chin (Institute for Information Industry)  
 Lynch, Ryan (Apple Inc.)  
 MacFarlane, James (Easypress)  
 Makower, Dave (Apple Inc.)  
 Mandelbaum, David (Barnes & Noble)  
 Manis, Will (Metaplates)  
 McCloy-Kelley, Liisa (Random House)  
 McCoy, Bill (IDPF)  
 Menzies, Tracey (HarperCollins)  
 Mitchell, Chris (Random House)  
 Moore, Helen (HarperCollins)  
 Muller, Eric (Adobe)  
 Murata, Makoto (JEPA EPUB Study Group)  
 Mussinelli, Christina (Associazione Italiana Editori)  
 Nagai, Yoshinori (Sharp)  
 Novelli, Joe (Sony)  
 O'Connor, Theresa (Apple Inc.)  
 Ohmura, Yoshinori (Impress Holdings)  
 Olenick, Michael (Bowker)  
 Oshiyama, Taka (East Co.)  
 Pagano, Pat (Barnes & Noble)  
 Picco, Marty (AppFoundry)  
 Prabhu, John (HOV Services)  
 Pritchett, James (Learning Ally (formerly RFB&D))  
 Rao, Vishnu (Sharp Laboratories)  
 Rivlin, John (Google)  
 Rubino, Frank (Kaplan Publishing)  
 Ruffino, Daniel (Penguin)  
 Rui Hua, Wang (Gansu DUZHE Digital Sci&Tech)  
 Ruse, Tyler (LibreDigital)  
 Sanicola, Daniel (Penguin)  
 Schirmer, Lorenz (Monotype Imaging)  
 Severdia, Ron (Metrodigi, Inc.)  
 Shiohama, Daihei (Voyager Japan)  
 Shrivastava, Abhishek (CourseSmart)  
 Siegman, Tzviya (John Wiley & Sons)

## ISO/IEC 23736-5:2020(E)

Slavin, Wayne (Barnes & Noble)  
Slye, Christopher (Adobe)  
Smith, Michael (IDPF)  
Soiffer, Neil (Design Science)  
Sorotokin, Peter (Adobe)  
Stevenson, Tobias (eBookArchitects)  
Tahara, Kyoji (Toppan Printing)  
Takase, Hiroshi (East Co.)  
Tallent, Joshua (eBookArchitects)  
Tanabe, Shu (Toppan Printing)  
Thomas, Vinu (Impelsys)  
Tsumagari, Koichiro (Voyager Japan)  
Valentine, Chelsea (LibreDigital)  
Vangage, Peter (Harlequin)  
Vido, Ariel (Geografica Editora)  
Wait, John (Pearson)  
Walkley, George (Hachette Book Group)  
Watters, Kevin (Harlequin)  
Webster, Roger (Barnes & Noble)  
Weck, Daniel (DAISY Consortium)  
Wei, Selena (Chinese Foundation for Digitization Technology)  
White, Russell (Random House)  
Wiles, Alexis (Overdrive, Inc.)  
Witwer, Adam (O'Reilly)  
Wright, Ric (Adobe)  
Young, Liz (Crossway)  
Zu, Alex (ASTRI (Hong Kong Applied Science & Technology Research Institute))

### › Invited Experts/Observers

Bowes, Rick  
Cazenove, Rhys  
Collingridge, Peter  
Cook, Mike  
Etemad, Elika J. *W3C CSS WG Liason*  
Forster, Karen  
Freed, Geoff  
Fujisawa, Jun  
Garrish, Matt  
Gould, Bryan  
Görner, Martin  
Hsieh, Michael  
Ishii, Koji  
Johar, Kenny  
Karlsson, Mattias  
Kennedy, Dianne  
Kilborn, Peter  
Koppel, Josh  
Lee, Tommy  
Lu, Kenny  
Lubeck, Scott  
Masanori, Kusunoki  
McKinney, Steven  
Murakami, Shinyu  
Ning, Elliott  
Noring, Jon

Norton, Paul  
 Oishi, Yasuharu  
 Passey, Lee  
 Rosmaita, Gregory  
 Seaman, David  
 Shan, Walter  
 Smith(tm), Michael (W3C) *W3C Liason*  
 Sperberg, Roger  
 Walsh, Norman  
 Zergaoui, Mohamed

For more detailed acknowledgements and information about contributors to each version of EPUB, refer to [Acknowledgements and Contributors \[EPUB3Overview\]](#) .

## › References

---

### Normative References

- [ContentDocs301] [EPUB Content Documents 3.0.1](#) .
- [HTML5] [HTML5: A vocabulary and associated APIs for HTML and XHTML](#) .
- [ISOSchematron] [ISO/IEC 19757-3: Rule-based validation — Schematron](#) .
- [MediaOverlays301] [EPUB Media Overlays 3.0.1](#) .
- [NVDL] [ISO/IEC 19757-4: NVDL \(Namespace-based Validation Dispatching Language\)](#) .
- [OCF301] [Open Container Format 3.0.1](#) .
- [Publications301] [EPUB Publications 3.0.1](#) .
- [RFC2119] [Key words for use in RFCs to Indicate Requirement Levels \(RFC 2119\)](#) . March 1997.
- [RFC3987] [Internationalized Resource Identifiers \(IRIs\) \(RFC 3987\)](#) . M Duerst, et al. January 2005.
- [RelaxNG] [ISO/IEC 19757-2: Regular-grammar-based validation — RELAX NG. Second Edition](#) . 2008-12-15.
- [SMIL] [SMIL Version 3.0](#) . D. Bulterman, et al. 01 December 2008.
- [StructureVocab] [EPUB 3 Structural Semantics Vocabulary](#) .
- [XML] [Extensible Markup Language \(XML\) 1.0 \(Fifth Edition\)](#) . T. Bray, et al. 26 November 2008.
- [XMLNS] [Namespaces in XML \(Third Edition\)](#) . T. Bray, D. Hollander, A. Layman, R. Tobin. W3C. 8 December 2009.
- [XSD-DATATYPES] [XML Schema Part 2: Datatypes Second Edition](#) . Paul V. Biron et al. 28 October 2004.

## Informative References

[EPUB3Overview] [EPUB 3 Overview](#) . Garth Conboy, et al.

[XPTRSH] [XPointer Shorthand Notation](#) .



