# Nordic Guidelines for the Production of Accessible EPUB 3 TEST

## February 16, 2021

## Contents

1		oduction 2
	1.1	Situating the Nordic Guidelines in the World of Specifications
2	Fori	mat Requirements 3
	2.1	Required EPUB Standard
		2.1.1 Accessible Publishing Knowledge Base
	2.2	Container
		2.2.1 META-INF
	2.3	Publication Resources
	2.4	Package Document
		2.4.1 Metadata
		2.4.1.1 schema.org Accessibility Metadata
		2.4.1.2 Optional Increased Title and Creator Metadata Granularity 5
		2.4.1.3 Optional Additional Metadata
		2.4.2 Manifest
		2.4.3 Spine
	2.5	Content Documents
		2.5.1 XHTML
		2.5.1.1 XML Declaration and Encoding
		2.5.1.2 Document Type Declaration
		2.5.1.3 HTML Root Attributes
		2.5.1.4 Namespaces
		2.5.1.5 Language Definition
		2.5.2 Title
		2.5.3 Metadata
	2.6	Navigation Documents
		2.6.1 EPUB 3.2 Navigation Document
		2.6.1.1 The Table Of Contents
		2.6.1.2 Page List
		2.6.1.3 Landmarks
		2.6.2 NCX Navigation Document
	2.7	Images
		2.7.1 Image File Naming and Directory Structure
		2.7.2 Resizing of Images
		2.7.3 Image Editing
	2.8	CSS
	2.9	Fonts
	2.10	Javascript 11

	2.11	Validation	12
3		Structural Divisons and Semantics	12 13 13 14 14 14 15 16 16 16 17 17 18 18 18 18 19 19 20 20 20 21 21
4	<b>Spe</b> 6 4.1 4.2 4.3	Mathematical Content	21 21 21 22
5	App 5.1 5.2	External Resources	22 22 23 23 23 24 24

## 1 Introduction

This guidelines document has been developed as a joint effort between the (mostly) Nordic Agencies dedicated to providing accessible literature in different formats to children and adults with various reading impairments or special needs. The participating organisations are Celia, Dedicon, HBS, MTM, NLB, Nota, SBS, SPSM, and Statped.

Fundamental to the process of adapting text-based media is the role of beginning with well-structured digital

content. Previously this has been achieved through XML structures as defined by the ANSI/NISO Z39.86 specification for digital talking books (DTBook). The structures specified in these guidelines, however, are based on a profile of HTML5 requiring the use of XML serialization. This ensures that content can be reliably manipulated and rendered. Moreover, the EPUB 3.2 specification and the accompanying EPUB Accessibility 1.0 specification provide constructs that further ensure semantically meaningful structures and increase the accessibility of the content.

This document can be seen as the successor of the previous EPUB 3 guidelines used by the Nordic agencies (2015-1), and are based on the most recent EPUB 3 specification, EPUB 3.2.

The target audience of the document is mainly contracted EPUB 3 suppliers, but the guidelines will also be used by staff at the Nordic agencies, other vendors and interested parties, etc.

### 1.1 Situating the Nordic Guidelines in the World of Specifications

The Nordic Guidelines work as an application of higher-level standards for the purpose of accessible EPUB 3 production at the Ordering Agencies. As such, they build on these higher-level specification documents (EPUB 3.2, EPUB Accessibility 1.0, etc.) and provide more detailed instructions for their application in the specific context of production at the Ordering Agencies.

The application of the Nordic Guidelines can also be further specified on lower levels, e.g. instructions on a per-Ordering Agency and/or per-title basis.

The different levels of specification can be expressed as the following hierarchy:

- 1. High-level specifications, e.g. EPUB 3.2, EPUB Accessibility 1.0, HTML5, etc.
- 2. The Nordic Guidelines
- 3. General Ordering Agency-specific Guidelines (usually these build on the Nordic Guidelines, but might also contain some deviations from them)
- 4. Title-specific instructions, usually expressed in the form of Editing Instructions, containing further guidance on how to treat a specific title

Note that the forms of levels 3 and 4 will vary between the Ordering Agencies, and their standardisation is outside the scope of this document.

## 2 Format Requirements

### 2.1 Required EPUB Standard

Suppliers are required to refer to the specifications provided in version 3.2 of the EPUB standard. See https://www.w3.org/publishing/epub32/epub-spec.html.

#### 2.1.1 Accessible Publishing Knowledge Base

In addition to following the recommended version of the EPUB standard, suppliers are required to follow the recommendations in the Accessible Publishing Knowledge Base that is maintained by the Daisy Consortium: https://kb.daisy.org/publishing/docs/.

The recommendations in the Knowledge Base may differ from the specification. As long as no conflicts occur during validation, the recommendations in the Knowledge Base are to be followed.

### 2.2 Container

The EPUB container file must be given the production number provided with the order and correspond with the identifier stored in the dc:identifier element located in the Package metadata.

The container archive is required to have the .epub extension. The file extension must be lowercase.

Note that the mimetype file must be archived as the first file in the Container.

#### 2.2.1 **META-INF**

The container.xml file must identify no more than one media alternative, unless indicated otherwise by the Ordering Agency.

The container.xml file shall look like this, unless indicated otherwise:

No other files, optional or otherwise, are allowed in the META-INF directory unless specifically indicated by the Ordering Agency.

### 2.3 Publication Resources

All publication resources are required to be located in a directory called EPUB. Publication resources other than the Package Document (.opf) and XHTML Content Documents are to be located in dedicated resource sub-directories.

### 2.4 Package Document

The name of the Package Document file is required to be package.opf. Suppliers are required to use the file extension .opf for the Package Document.

The following xml declaration must be placed at the first line of the document:

```
<?xml version="1.0" encoding="utf-8" ?>
```

Suppliers are required to apply the following attributes and values to the <package> element:

xmlns="http://www.idpf.org/2007/opf"
xmlns:dc="http://purl.org/dc/elements/1.1/"
prefix="nordic: http://www.mtm.se/epub/"
version="3.0"
xml:lang="xx" (The publication language code)
unique-identifier="pub-identifier"

Required child elements to the <package> element are:

- <metadata>
- <manifest>
- <spine>

### 2.4.1 Metadata

The following metadata are required to be placed in the <metadata> element:

Note that the default value to use in <dc:publisher> is the shorthand for the Ordering Agency, e.g. "NLB", "MTM", etc. Optionally, the Ordering Agency can request the original source publisher to be expressed in <meta property="dc:publisher.original">.

If the source material does not have an ISBN, ISSN or any other systematic source identifier the content of the <dc:source> element will be a text string based on whatever available information about the source there is (publisher, year of publication etc.). This will be provided by the Ordering Agency via Editing Instructions.

#### 2.4.1.1 schema.org Accessibility Metadata

Also required are schema.org accessibility metadata, http://kb.daisy.org/publishing/docs/metadata/sc hema.org/index.html. Which metadata that are relevant depend on the type of content included in the package. Use the Accessibility Checker for EPUB tool to find out which metadata are relevant, https://inclusivepublishing.org/toolbox/accessibility-checker/getting-started/. It will typically be something like this, but not necessarily exactly the same:

```
<meta property="schema:accessibilitySummary">This publication conforms to the Nordic
Guidelines for the Production of Accessible EPUB 3, version 2020-1.</meta>
<meta property="schema:accessMode">textual</meta>
<meta property="schema:accessMode">visual</meta>
<meta property="schema:accessModeSufficient">textual visual</meta>
<meta property="schema:accessModeSufficient">textual </meta>
<meta property="schema:accessibilityFeature">structuralNavigation</meta>
<meta property="schema:accessibilityFeature">MathML</meta>
<meta property="schema:accessibilityFeature">alternativeText</meta>
<meta property="schema:accessibilityFeature">alternativeText</meta>
<meta property="schema:accessibilityHazard">none</meta>
<meta property="ally:certifiedBy">_[the ordering agency]_</meta>
link rel="dcterms:conformsTo" href="https://format.mtm.se/nordic_epub/2020-1"/>
```

As with <dc:publisher>, the default value of <meta property="a11y:certifiedBy"> is the shorthand for the Ordering Agency.

Please note that the version number forming the end of the dcterms:conformsTo URL must match the value of the statement in nordic:guidelines, i.e. (for this version of the guidelines) "2020-1".

Section 5.2.4 of this document contains language specific texts for <meta property="schema:accessibilitySummary">.

#### 2.4.1.2 Optional Increased Title and Creator Metadata Granularity

The Ordering Agency may request increased metadata granularity for the representation of title and creator details, making use of a linked meta element refining the character of the main metadata element.

The following is an example of a more granular way of expressing title metadata, separating main title, subtitle, and edition title statements from each other:

Similarly, creator/author information can be expressed with explicit refinement of the creator role of each author, as in the following example:

```
<dc:creator id="creator1"> _[coauthor of the publication]_ </dc:creator>
<meta refines="#creator1" property="role" scheme="marc:relators" id="role">aut</meta>
<dc:creator id="creator2"> _[coauthor of the publication]_ </dc:creator>
<meta refines="#creator1" property="role" scheme="marc:relators" id="role">aut</meta>
<dc:creator id="creator3"> _[illustrator of the publication]_ </dc:creator>
<meta refines="#creator1" property="role" scheme="marc:relators" id="role">ill</meta>
```

The codes to use as values for the creator role are taken from the MARC relators vocabulary, https://id.loc.gov/vocabulary/relators.html.

#### 2.4.1.3 Optional Additional Metadata

Additional metadata for inclusion can be supplied by the Ordering Agency on a per-title or per-Agency basis.

#### 2.4.2 Manifest

In the <manifest> all publication resources are declared. Each resource is declared using the <item> element with required attributes id, href and media-type. Some items must also have a properties attribute.

The id is a unique identifier used to address a particular publication resource. The only formal requirement here is its uniqueness, but the Ordering Agency will require certain naming conventions to be applied.

The href is simply a path to the publication resource. The path will be relative to where the package document is located.

The media-type attribute specifies what type of resource it is and what format. Here are some of the more common examples:

- application/xhtml+xml
- text/css
- image/jpeg
- application/x-dtbncx+xml
- application/javascript

Refer to https://www.w3.org/publishing/epub3/epub-spec.html#sec-core-media-types for additional information about various media types.

Some resources have additional properties that need to be identified in the corresponding item in order for a reading system to be able to implement these properties. This is done by using the properties attribute. Common examples are:

- nav for the xhtml navigation document
- cover-image for the cover image
- scripted for any content document where javascripts are being used

• mathml for any content document containing MathML

Refer to https://www.w3.org/publishing/epub3/epub-packages.html#app-item-properties-vocab for additional information about the properties attribute.

### 2.4.3 Spine

In the <spine> all content documents are listed in the correct reading order. This is done by using an <itemref> element for each content document and simply listing them in the desired order. The <itemref> element is associated with the corresponding <item> for the content document in the <manifest> by using the idref attribute and setting the value to the id attribute of the <item> element. The idref attribute is a required attribute and the id it refers to must be unique. Here is an example:

In the <manifest>:

```
<item id="item_7" href="007-chapter.xhtml" media-type="application/xhtml+xml"/>
In the <spine>:
<itemref idref="item_7"/>
```

The <itemref> has an optional attribute called linear. The linear attribute can have the values yes (default) or no. If the attribute is omitted it is set to yes. The linear attribute indicates whether the referenced item contains content that contributes to the primary reading order and has to be read sequentially (yes) or auxiliary content that enhances or augments the primary content and can be accessed out of sequence (no). Examples of auxiliary content include: notes, descriptions and answer keys.

linear="no" should generally be applied to all content documents representing content that precedes the title page, e.g. cover, half-title page, etc. This is to ensure a reader always lands on the title page when opening the book in their reading system.

If a fall-back ncx navigation document is included, this is required to be referenced by adding the toc attribute to the <spine> element and assign as value the id attribute of the <item> referring to the ncx file, like this:

```
<spine toc=" [id value of the manifest item that refers to the ncx file] ">
```

### 2.5 Content Documents

#### 2.5.1 XHTML

The XHTML content files specified by the EPUB 3.2 specification are based on HTML5. Suppliers are required to use the extension .xhtml.

### 2.5.1.1 XML Declaration and Encoding

The following xml declaration must be used:

```
<?xml version="1.0" encoding="utf-8"?>
```

### 2.5.1.2 Document Type Declaration

The following document type declaration must be included:

```
<!DOCTYPE html>
```

#### 2.5.1.3 HTML Root Attributes

Suppliers are required to include the following attributes on the html root element:

- xmlns XML namespace
- xmlns:epub EPUB namespace
- xml:lang XML Language definition
- lang HTML Language definition

#### 2.5.1.4 Namespaces

The following namespace values are required to be applied to the namespace attributes:

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

#### 2.5.1.5 Language Definition

Suppliers are required to identify primary languages for each content file instance using the xml:lang and lang attributes. The values for xml:lang and lang must be the same.

If not requested otherwise by the Ordering Agency, language should be coded using values from the IANA registry of valid language codes, https://www.iana.org/assignments/language-subtag-registry/language-subtag-registry.

Please note that the most granular language tag should be used, e.g. the macrolanguage value "no" for Norwegian could be used, but only in instances where the sublanguages Norwegian Nynorsk ("nn") or Norwegian Bokmål ("nb") are not known.

The accuracy of language coding is vital, and Suppliers are instructed to contact the Ordering Agency in case of any uncertainties.

#### 2.5.2 Title

The <title> element of every xhtml content document must match the dc:title metadata of the package file.

#### 2.5.3 Metadata

The content documents are required to contain two <meta> elements.

```
<meta name="dc:identifier" content="_[production UID]_"/>
<meta name="viewport" content="width=device-width"/>
```

The production UID must match the dc:identifier in package.opf.

#### 2.6 Navigation Documents

#### 2.6.1 EPUB 3.2 Navigation Document

The principle navigation document of the EPUB package is the xhtml file with the properties attribute set to nav in the <manifest> section of the package document. For matters of convenience mostly, this file is required to be named nav.xhtml.

Section 5.2.1 of this document contains language specific headings for the <nav> sections of the navigation document that are listed below.

#### 2.6.1.1 The Table Of Contents

The first required <nav> element in the file is for the main table of contents. The <nav> element is required to have a <h1> as the first child element. The <nav> element must have the aria-labelledby attribute set to the id of the <h1> element. The role and epub:type elements must be set as follows:

```
<nav role="doc-toc" aria-labelledby="n1" epub:type="toc">
<h1 id="n1">Contents</h1>
```

All headings in the main body of text must be included in the <nav role="doc-toc"...> element and the heading levels must be implied through nesting. Headings of sidebars, text boxes or other secondary content should not be included, unless specific instructions are given by the Ordering Agency.

The links must always reference the corresponding sectioning element for the heading in the content file, not the h[x] element directly. This is usually a <section> element, but can also be <aside> or any other sectioning element. Thus, in the example below, the link should point to the id "level3\_2" in that content file:

A section without any heading must be referenced using its aria-label value.

Note that this is not a representation of the table of contents in the source material. Only headings, or references using the aria-label value of the corresponding <section> element, should be included.

See the following link for further information (but disregard examples of unlinked headings or hidden branches):

http://kb.daisy.org/publishing/docs/navigation/toc.html

### 2.6.1.2 Page List

If the source material contains pagination, the next required <nav> element is:

```
<nav role="doc-pagelist" aria-labelledby="n2" epub:type="page-list">
<h1 id="n2">Pages</h1>
```

See the following link for further information:

http://kb.daisy.org/publishing/docs/navigation/pagelist.html

If no pagination occurs in the source material this may be omitted.

### 2.6.1.3 Landmarks

The final required <nav> element is:

```
<nav role="navigation" aria-labelledby="n3" epub:type="landmarks">
<h1 id="n3">Landmarks</h1>
```

The Landmarks section should contain references to any main sections of the front- and rearmatter parts of the source material, as well as a reference to the start of the main content of the book, which will be the start of the first content file with the word "bodymatter" in its epub:type attribute of the top-level <section> element. The list of references may include the following, when applicable:

- Table of Contents (the toc of the source material)
- · List of Images
- List of Tables
- Acknowledgments

- Preface
- Start of Content
- Answers
- Glossary
- Bibliography
- Index

Note that the links, the <a> elements, in the Landmarks section must have appropriate epub:type values. Examples can be found here:

http://kb.daisy.org/publishing/docs/navigation/landmarks.html

### 2.6.2 NCX Navigation Document

The EPUB package may include an ncx navigation document as fall-back for older reading systems that have not implemented functionality for the EPUB 3 navigation document. It is not required by the EPUB specification, but it may be requested by the Ordering Agency. No ncx file should be included unless specifically requested. If requested, the file is required to be named nav.ncx.

Refer to http://www.idpf.org/epub/20/spec/OPF\_2.0.1\_draft.htm#Section2.4 for information about how the ncx file should be formed. Required elements are:

- <head>
- <docTitle>
- <navMap>
- <pageList>

The <head> element must contain the following <meta> element:

```
<meta name="dtb:uid" content=" _[production UID provided by the ordering agency]_ "/>
```

All headings must be included in the <navMap> and the heading levels must be implied through nesting.

### 2.7 Images

Image content is required to be captured in the jpeg or png formats. The format extension is required to be .jpg or .png. Photos, drawings and illustrations must be captured in the jpeg format. For line graphics, charts and diagrams the preferred format is png.

Suppliers are required to maintain the highest possible image quality. The quality requirements are the following:

- The aspect ratio of the original image must always be maintained.
- Colours must be reproduced and rendered accurately.
- The image must be free of any compression artefacts or any distortion effects arising from scanning.
- Text in images, line graphics and small detail must be crisp and fully legible, or with no degradation in legibility in comparison with the original image.
- If the jpeg format is used, the quality setting is required to be set at around 90% (or a corresponding setting, if a different scale of measurement is used). A setting of 100% is usually unnecessary as it yields substantially larger files with little or no visible improvement of quality. If, however, there would be visible difference in quality between 90% and a higher setting, then the higher setting should be used.

#### 2.7.1 Image File Naming and Directory Structure

Image files are required to be stored in a folder named *images* placed at the same level relative to the package document. The images folder must not contain subfolders.

Image files should be named according to the following file naming convention:

[UID] - [XXX] . [png|jpg]

The *UID* must be identical to the value of the dc:identifier metadata element. *XXX* is a three-digit numeric string corresponding to the order of the image in the content, with leading zeros as needed. (If the number of images in a single book exceeds 999, a four-digit number should be used.)

Example:

X41001A-013.png

An exception to this is the cover image, which should be named simply cover.png or cover.jpg.

#### 2.7.2 Resizing of Images

The maximum image size is set to 800 pixels on the longest side of the image, unless an increase in the size of an image is required to achieve the legibility of text rich images, see point 4 above, or other small details.

In those circumstances where this requirement conflicts with requirements for legibility, the Supplier is required to contact the Ordering Agency.

### 2.7.3 Image Editing

Images are required to be captured exactly as they are in the scans or publisher files provided by the Ordering Agency. There is generally no need for editing images to enhance them above the quality of the originals (e.g. cloning to smooth edge of images spanning over spread, enhancing bad scans or old photos with bad quality already in print, etc.).

When whole sections or parts of text is placed on top of a background image, the text may need to be edited out of the background, if possible. However, no advanced restoration of lost image details are required. Note that if a publisher file is provided, images and text may have been placed in different layers and the image can be extracted without any text.

### 2.8 CSS

Suppliers are required to include the standard css file issued by the Ordering Agency. The css file is required to be stored in a folder named css and placed at the same level relative to the package document.

The clink> element is required to be applied to the relevant content documents.

### 2.9 Fonts

Fonts present in PDF source material must not be included in the EPUB, unless specifically indicated by the Ordering Agency.

Fonts, when present, must be stored in a folder named *fonts* and placed at the same level relative to the package document.

#### 2.10 Javascript

Javascript files requested by the Ordering Agency are required to be stored in a folder named *javascript* and placed on the same level relative to the package document.

The **script** element is required to be applied to the relevant content documents.

Suppliers must not include javascript files in the EPUB unless specifically instructed to do so by the Ordering Agency.

#### 2.11 Validation

Validation of EPUB files should be done using EPUBCheck for structural validity and conformance to the EPUB standard and the Ace tool for accessibility evaluation. The latest distributions of these tools should be used. The tools can be found here:

- https://github.com/w3c/epubcheck/releases
- https://inclusivepublishing.org/toolbox/accessibility-checker/getting-started/

The Ordering Agencies are currently developing a specific validation tool based on these guidelines. When this tool is available it will be the primary validation method.

## 3 General Requirements for Content Documents

### 3.1 Structural Divisons and Semantics

The EPUB Content File structure specified in these guidelines is generally made up of a multi-page HTML file set. Major divisions of the publication are to be captured in individual XHTML content files. The individual content files will typically correspond to *Part* or *Chapter* divisions of the book. Other major book components such as colophon, index or appendix, which can be found in front-matter and back-matter, will normally be stored in separate files as well.

The structural divisions of the publication are required to be semantically inflected by wrapping every structural part of the main text body in <section> elements, with proper nesting of subsections. Usually, a subsection ends whenever a new heading of the same level or higher occurs in the source material. However, there may be situations where a <section> element needs to be closed sooner. If so, this should be specified in the Editing Instructions given by the Ordering Agency. If the structure of the source material is unclear and the Editing Instructions do not provide answers, please contact the Ordering Agency.

The semantic role of each content document must be specified, when possible. This is done by adding the role and epub:type attributes to the top level <section> element of each content file. The available role attribute values are listed here:

https://www.w3.org/TR/dpub-aria-1.0/#roles

When the content file can be matched to one of the roles listed, the top level **<section>** element is required to have the **role** attribute with that matching value set. If no suitable role exists the **role** attribute may be omitted.

The epub:type attribute is also required for the top level <section> element of each content file and must first contain one of the following partition values:

- cover
- frontmatter
- bodymatter
- backmatter

The frontmatter, bodymatter and backmatter partition values must also be combined with a divisioning or sectioning value found here:

https://idpf.github.io/epub-vocabs/structure/

If no matching divisioning or sectioning value can be found, the epub:type will be set to only the appropriate partition value. Also, the content file containing the cover image of the printed book will have simply epub:type="cover".

If an EPUB file contains both parts and chapters, and each part and all its chapters is contained in a single content file, the above is also required for the second level **<section>** elements, which will contain the chapters.

Furthermore, all <section> elements are required to have a label. When the <section> has a heading, usually <h1>, <h2>, etc., the heading serves as a label but must be associated with the <section> element. This is done by using the aria-labelledby attribute and setting the id of the associated header as value:

If the section is untitled, the aria-label attribute is required to be used instead and have an appropriate label assigned to it. This will usually be addressed in the Editing Instructions, but below you will find a list of default values for generic sections often untitled in books, that should be used if there are no other instructions given by the Ordering Agency:

Section identification	Default aria-label value
@epub-type="cover"	Cover
<pre>@class="frontcover"</pre>	Front cover
<pre>@class="backcover"</pre>	Back cover
<pre>@class="leftflap"</pre>	Jacket left flap
<pre>@class="rightflap"</pre>	Jacket right flap
@role="doc-colophon"	Publisher information
<pre>@role="doc-dedication"</pre>	Dedication
@role="doc-epigraph"	Epigraph
@role="doc-toc"	Contents

Section 5.2.2 of this document contains language specific values.

See http://kb.daisy.org/publishing/docs/html/sections.html for more information.

### 3.2 File Naming Convention

The basic scheme for naming individual files is:

```
[UID] - [XXX] - [role] .xhtml
```

The *UID* must be identical to the value of the dc:identifier metadata element. *XXX* is a three-digit numeric string corresponding to the order in the <spine>, with leading zeros as needed. *role* corresponds to the ARIA role of the main section element in the content file, minus the "doc-" part. In the absence of an ARIA role for the top-level section of the content document, the value from the epub:type attribute should be used instead. In the case of multiple epub:type (e.g. frontmatter titlepage) values, the most specific (e.g. titlepage) should be used.

Example:

X41001A-010-chapter.xhtml

### 3.3 Special Content Requirements

Some content files have certain contents that are required to be included and marked up correctly.

#### 3.3.1 Cover

The following class attributes must applied to the child **section** elements where appropriate:

- <section class="frontcover">
- <section class="rearcover">
- <section class="leftflap">
- <section class="rightflap">

The front cover, when available, must be captured as a .jpg or .png image file and given the name cover.jpg/png. The properties="cover-image" attribute must be applied to the manifest item element corresponding to this image. The <img> element referencing the cover image must have aria-role set to "doc-cover".

Note that the linear="no" attribute must be applied to the itemref element in the package spine corresponding to this content file.

#### 3.3.2 Title Page

Content corresponding to the publication's full title must be included in an <n1> element.

```
<h1 epub:type="fulltitle" class="title" id="booktitle">
```

Use of the <span> element is only required when a subtitle is present. The following attribute usage must be applied to the appropriate <span> element:

<span epub:type="title"><span epub:type="subtitle">

#### 3.3.3 Parts

Some publications are divided into parts, each containing a number of chapters. By default, the part heading and any associated contents must be placed in a separate content file and the <section> element associated with the part heading must be closed at the end of that file. Each chapter of the part must then have its own content file. Note that chapter headings must be marked up as <h2>, even though they are the first heading of its content file. Semantic attributes, role and epub:type, must be given to the <section> elements of both part and chapters.

Ordering Agencies might give instructions to keep whole parts in single content files, if the size of the content is small enough to not inhibit the performance of reading systems. The part heading will be the <h1> and the chapter headings <h2> and the <section> element associated with the part heading will wrap all the contained chapters. Of course, semantic attributes, role and epub:type, must still be given to the <section> elements of both part and chapters. Note that this alternative option is only to be used if cleared by the Ordering Agency.

### 3.4 Mark-up Requirements

These guidelines will not give highly detailed descriptions of how to correctly handle general content. Common recommendations for making valid and accessible HTML content will apply. As a general rule, the simple solution is almost always the best solution. Using common elements like , <blockquote>, <aside>, , , <dl>, , <figure> etc. and the common structural elements like <section> and headings will almost always be sufficient. More specialised elements, or special attributes, may be needed occasionally, though. Some of the more common cases will be described below, and more obscure ones will be covered specifically in Editing Instructions.

For further information about the common HTML elements and their attributes, please refer to any online resource on HTML 5, for instance:

- The formal HTML(5) specification: https://html.spec.whatwg.org/multipage/
- https://www.w3schools.com/

### 3.4.1 Pagination

All page breaks occurring in the source copy are required to be indicated with one of the following elements unless stated otherwise by the Ordering Agency:

- Inline: <span epub:type="pagebreak" role="doc-pagebreak">
- Block: <div epub:type="pagebreak" role="doc-pagebreak">

Additionally required attributes:

- aria-label=""
- class=""
- id=""

The value for the aria-label attribute must be identical to the page number in the source copy. For empty pages occurring for example between chapters, this attribute must have a value corresponding to the number implicit for that page.

In those cases where pagination of a text cannot be effectively represented using the following rules, the Supplier is required to contact the Ordering Agency.

The class attribute must have either of the following values:

- page-front
- page-normal
- page-special

The attribute value page-normal is required for the pagination of the main contents of the book, basically the bodymatter, but can also be used for the frontmatter and backmatter if the pagination is continuous.

The attribute value page-front is required when the page numbering series of the frontmatter is not logically continued in the publication bodymatter, i.e. roman numerals.

The attribute value page-special is required for any parts of the book not numbered in a standard manner, for example appendices, suites of photography etc.

The attribute id is simply a unique identifier.

By default, the pagebreak elements are required to be empty, like in the following examples:

- Inline: <span epub:type="pagebreak" role="doc-pagebreak" class="page-normal" id="page-38" aria-label="38"></span>
- Block: <div epub:type="pagebreak" role="doc-pagebreak" class="page-normal" id="page-38" aria-label="38"></div>

However, there may be agency specific instructions to place the page number, as it is displayed in the source material, as content in the pagebreak elements. For example:

This is only to be done if specific instructions are given by the Ordering Agency.

There must be no unnumbered pages. If unnumbered pages in the source material are implicitly numbered, for instance the initial pages of a book, they must be numbered accordingly. That is, if the pagination starts with page 9 in the source material, the previous pages must be numbered 1-8. If there are other unnumbered

pages in the source material they must be assigned numbers and given the class attribute page-special. This will be specified by the Ordering Agency via Editing Instructions.

### 3.4.2 Headings

The <h1>-<h6> elements are used to reflect the heading structure present in the source copy. Note that <h[x]> tag must be contained within its respective sectioning element. Sectioning elements that may require headings are, for instance, <section>, <aside> and <nav>.

Headings that do not contribute to the hierarchical structure of the work and that are not desired to be included in the navigation document can be marked up using . Bridgehead markup may be specifically requested by the Ordering Agency via Editing Instructions and must never be used otherwise.

#### 3.4.3 Figures

All proper figures, illustrations, photographs, icons and other symbols must be captured as images and stored in the EPUB file, unless other instructions are given (see section 2.7). Purely decorative graphics that have no other purpose than layout can be ignored. If there are any doubts about whether to include certain graphics or not, the Supplier is required to contact the Ordering Agency.

Unless the image occurs inline in the source material, any image is required to be placed inside a <figure> element with a class attribute set to image. If the image has a caption, the caption is required to be marked up with the <figcaption> element and placed as the first child of the <figure> element. If no caption is present, the <img> element will be the first child.

Small symbols or other non-typographical content that might occur inline, are required to be represented with <img> elements alone, without the <figure> container. There may be cases where it is unclear whether to regard symbols or icons as inline or not, for instance small icons in connection with exercises. Specific instructions on how to handle these may be given in Editing Instructions. If doubt remains, the Supplier is required to handle the image as a block element and use <figure> or, alternatively, contact the Ordering Agency.

Images in tables or lists may be handled as inline images if there are no captions or similar.

### 3.4.3.1 Text Extraction from Images

When images contain text that is integral to the image itself, i.e. not a caption or similar, this text is required to be extracted as accessible text. This text must be placed in a placeholder within the <figure> element of the image. The placeholder should be placed after the <img> element, before the closing </figure> tag. Suppliers are required to use the <aside> element for this placeholder with the following attributes:

- class="fig-desc"
- id=""

where id is a unique identifier. Furthermore, the corresponding <img> element must then be given the following attribute:

• aria-describedby=""

with the same value as the id of the <aside> placeholder.

The extracted text is then placed inside the placeholder, marked up correctly and placed in a logical reading order, if there is any. Here is an example of how a figure with extracted text should be handled:

Note that this is the construction that is used by several Ordering Agencies for adding image descriptions in post-production.

#### 3.4.3.2 Alt-texts

Accessibility guidelines require all images to be supplied with a short, descriptive text as value of the alt attribute of the <img> elements. Suppliers are not required to provide these descriptive texts, but should instead use one of the following generic values:

- photo for photographs
- illustration for illustrations
- figure for schematic images, graphs, diagrams, scientific models etc.
- symbol for icons, signs, inline images etc.
- map for maps
- drawing for drawings
- comic for comic strips and panels
- logo for logos

Section 5.2.3 of this document contains language specific alt-text values.

If there are any doubts about which value to assign the alt attribute, Suppliers are required to use figure.

#### 3.4.3.3 Image Series

If the source material contains a series of images, images that are linked in some way, each individual image must be marked up as described above with each image wrapped in separate <figure> elements. The whole series must then be wrapped in a <figure> element where the class attribute is set to image-series.

If there is a caption for the image series as a whole, this must be marked up with the <figcaption> element and placed inside the outer <figure> element, before the individual <figure> elements containing the images. Any caption for individual images are placed together with that image as described above.

This is how the markup will look like:

#### 3.4.4 Lists

Lists are a number of connected items (single words, sentences or whole paragraphs) written consecutively. They can be numbered or unnumbered. Any such content is required to be marked up with either 
 (ordered list) or 
 (unordered list). Each item of any list must be marked up with .

A list item may either contain inline content or block elements, but not a mixture of both. As a rule of thumb, if all items in the list consist of single words or short phrases no further block elements are needed. If one or more of the list items consist of sentences or paragraphs, use one or more elements inside every list item of the list.

#### 3.4.4.1 Numbered Lists

The numbering of an ordered list must not be included as content in the <1i> elements of the list. The numbering will be rendered by the reading system. The default type for the numbering is numeric. This can be changed by using the type attribute. The default starting point is 1 (regardless of which type of numbering the ordered list uses), but can be changed using the start attribute.

#### 3.4.4.2 Unnumbered Lists

Unnumbered lists often have some sort of bullet markers for each list item. The default is a solid black circle. The type attribute has been used before, to change the type of bullet symbol, but this is not supported in HTML 5. Using CSS is the proposed method of controlling this. Suppliers are not required to modify the CSS file to match the bullet markers of the source material unless specifically instructed to do so.

Lists without any bullet markers are required to have the attribute:

• class="plain"

By default, should be used here, but Ordering Agencies may give specific instructions to use .

#### 3.4.4.3 Tables of Contents

Any table of contents in the source material is required to be marked up in the following way:

This is typically used for the main table of contents that most books have somewhere in the frontmatter part, but it could also be a table of content for a single part or chapter. The <span class="lic"> element, used to separate the headings from the page references, must not be used in any other context. Any line that does not have a page reference must be marked up using only , without any <span class="lic">.

Sometimes, mostly in educational books, the table of contents can be more complicated, including other type of content than simply headings and page references. In these cases, specific instructions will be given by the Ordering Agency of how to handle that. Contact the Ordering Agency if anything is unclear.

#### **3.4.5** Tables

All tables or table-like structures are required to be marked up as . If the table has a caption it is required to be marked up with <caption> and placed just after the starting tag of the element. It can sometimes be unclear what content should go into the <caption> element. Sometimes there is a title in the table itself, spanning the entire width. This must be removed from the table structure and placed in the <caption> element. Sometimes there could be a regular caption above the table and a source reference at

the bottom. These must both go into the <caption> element in individual paragraphs. Non-standard use of the <caption> element should be specified by the Ordering Agency in the Editing Instructions.

The element is required to be used for containing the main body of table data. It is recommended, although not formally required, to use <thead> for any column heading at the top of the table. For the sake of consistency, it is required to use <thead> if there is at least one row of column heading at the top of the table, unless specific instructions are given to omit it. The element <tfoot> can be used if there are, for instance, a row at the bottom of the table where columns are summed up. This is not required, but can be included in Editing Instructions.

Tables are required to have a consistent number of table cells per row. If colspan or rowspan are used, take extra care that the total number of cells is correct.

Never use tables solely for the purpose of mimicking the layout of the source material. The colspan and rowspan attributes may be used with or elements, if necessary, but if the purpose of the layout in the source material is unclear and no instructions are given, Suppliers are required to contact the Ordering Agency for clarification.

#### 3.4.6 Definition Lists

All paired lists of words, phrases, expressions etc. and corresponding definitions, translations etc. are required to be marked up as <dl>. Note that language attributes may be required, for example with glossaries.

#### 3.4.7 Notes and Note References

Footnotes and endnotes are required to be handled in accordance with the recommendations in the Accessible Publishing Knowledge Base:

http://kb.daisy.org/publishing/docs/html/notes.html

Footnotes are to be placed in an <aside> at the end of the relevant <section>. If there are more than one note reference in the section, notes must **not** be grouped in a single <aside>, but separated so that each note is contained in its own <aside>.

Endnotes or chapter notes must be placed in a <section> element at the end of the content file, before the closing of the top-level <section> element. The <section> element containing the notes must have the role attribute set to doc-endnotes. It must also have the aria-label attribute set to Endnotes. Note that the <aside> element must not be used in this case.

Backlinks are required for each note, done according to the examples in the Knowledge Base. Unless other instructions are given, the text of the backlink in should read (in English):

```
<a href="#ref" role="doc-backlink">Go to the note reference.</a>
```

Section 5.2.5 of this document contains language specific backlink texts.

Note references are not to be marked up with <sup>. This is handled by the default CSS provided by the Ordering Agency.

### 3.4.8 Sidebars, Text Boxes etc.

The **<aside>** element is required to be used for any material that is placed in the margin, breaks the flow of the main text or is in some other way to be considered optional or non-essential.

Any material that is in some way contained or has a clearly visible beginning and end, but is still an integral part of the main text or content is required to be marked up with <div>. This could be, for instance, examples, definitions, tips, boxes containing words or phrases in exercises etc.

In both cases the elements are required to have a class attribute. The default value of the class attribute is text-box, but other values can be given in Editing Instructions.

For certain publications there may be instructions given to add aria-label attributes to either of the elements, but unless such instructions are given no aria-label attributes are to be included.

If there are any doubts about which element to use and there are no further instructions available in the Editing Instructions, Suppliers are required to contact the Ordering Agency.

Headings in text boxes is by default required to be marked up using . These are structurally unimportant headings and will not be included in the navigation document. If, however, certain text boxes should be included in the navigation document, they must be marked up using a <h[x]> element of an appropriate level, following the level structure of the surrounding text. Only text boxes marked up with <aside> can have <h[x]> headings. The Ordering Agency will specify in the Editing Instructions if <h[x]> markup is required.

### 3.4.9 Computer Code

Suppliers are required to mark up code content with the <code> element. For block instances containing several lines of code, the <code> element must be contained in a element.

In blocks of computer code, spaces and empty lines must be preserved.

#### 3.4.10 Bolding and Italics

The only elements to be used are <strong> for bold and <em> for italics. Other type of formatting may be required in certain publications and, if such is the case, specific instructions will be given in the Editing Instructions.

Principles for how <strong> and <em> are used can be found here:

http://kb.daisy.org/publishing/docs/html/emphasis.html

Please take care that text marked up with em or strong is identical with the original book. Do not include space or punctuation which is not emphasised in the original. Words in em or strong at the end of sentences should not include end-of-sentence punctuation, like full stops, unless the whole sentence is emphasised. If a sentence ends with an emphasised word, and the next sentence begins with a new emphasised word, make sure the words are marked up using separate instances of em or strong, and that the end-of-sentence punctuation is not included.

#### 3.4.11 Poetry and Verse

Poetry, song lyrics or any content written in verse, where lines of text must be preserved just as they are in the source material, is required to be marked up with <div class="verse">.

Each group of lines of text must be contained in a separate class="linegroup"> element and each line
of text must be marked up with span class="line">. HTML line breaks, class="line\_indent">
or consecutive lines within a line group. Indented lines may be marked up using span class="line\_indent">
or cspan class="line\_longindent">.

Line numbers must only be included if specific instructions are given about it, even if they are present in the source material. If line numbers are to be included, they must be marked up with <span class="linenum">.

If the content written in verse has a title it may be handled as a normal heading, with <section> elements wrapping the content. However, if it does not make sense to use a heading, it may be marked up with class="verse-title"> and placed within the <div class="verse"> container. This option must not be used unless specific instructions are given by the Ordering Agency.

If there is an author name placed under the verse it may be marked up with class="verse-author">
and placed at the end of the <div class="verse"> container.

### **3.4.12** Quotes

#### 3.5 Thematic Breaks in the Text Flow

In circumstances where depth of structure is not amenable to structural markup using **<section>** elements, Suppliers are required to use the **<hr>>** element to provide distinguishable paragraph-level thematic breaks in one of the following two ways:

- <hr class="emptyline"/> indicates thematic breaks represented by a vertical space between paragraphs.
- <hr class="separator"/> indicates thematic breaks represented by a visual marker such as an asterisk, horizontal rule or any other type of graphical symbol. The visual marker must not be rendered as content.

## 4 Specific Requirements for Advanced Content

#### 4.1 Mathematical Content

All mathematical or scientific content, content that contains mathematical operators, special characters, subscripts, superscripts etc., must be marked up using MathML. These guidelines will not go into much detail about the MathML language. Please refer to the specification:

https://www.w3.org/TR/MathML3/

For more detail about the MathML structure, please refer to each individual Ordering Agency's requirements. Note that some Ordering Agencies may not require MathML at all, at present. ASCIIMath notation may be requested instead or in combination with MathML markup. See http://asciimath.org for information about the ASCIIMath notation.

### 4.2 Special Characters, Unicode and Phonetics

If not requested otherwise by the Ordering Agency, Unicode characters should be represented by the correct Unicode character rather than an entity reference using e.g. decimal or hexadecimal notation. An inverted question mark, used in for example Spanish, should be represented as ¿ rather than ¿, ¿, or similar.

Note that entity reference coding of special characters may be requested on a per-production or per-Ordering Agency basis.

Unicode character accuracy for special characters, e.g. phonetic characters, is vital for a correct representation of the source material. A visual likeness to the characters used in the printed source material is not enough, as the characters may be used by assistive technology to present text information to the user, or being used in various conversions to other types of end-user formats.

Here are some resources for commonly used special characters:

- Phonetics: https://www.phon.ucl.ac.uk/home/wells/ipa-unicode.htm
- The Greek alphabet: https://www.fileformat.info/info/unicode/block/greek\_and\_coptic/list.htm

Similarly to phonetics and other special characters, punctuation, such as quotation marks, dashes, etc., should be preserved as they are represented in the source material. This means that careful attention needs to be payed to ensure correct representation of e.g. hyphen minus (- (U+002D)) vs. en dash (- (U+2013)) vs. em dash (- (U+2014)), hyphen minus vs. mathematical minus sign (- (U+2212)), simple quotation marks ("" (U+0022)) vs. typographic quotation marks ("" (U+201D, U+201C)), etc. Any exceptions to this general rule will be noted in Editing Instructions.

### 4.3 Placeholders for User Input Areas

In educational material, especially for younger children, it is common that the user is supposed to answer questions or solve problems by writing directly in the printed book. Usually, this is indicated by printed lines where the user can write text or boxes that can be ticked etc. Suppliers are required to use the <span> element to provide placeholders for these input fields in one of the following three ways:

- <span class="answer">---</span> for a horizontal line or box where any amount of text or numbers are meant to be inserted. Multiple lines in succession must be represented by a single <span class="answer">---</span> unless other instructions are given by the Ordering Agency.
- <span class="answer\_1">-</span> for a single space where only one character is meant to be inserted, typically a missing letter in a word or similar. If there are two missing letters in a word, there must be two <span class="answer\_1">-</span> elements, without space between them.
- <span class="box">---</span> for check boxes.

## 5 Appendix

### 5.1 External Resources

- EPUB 3.2 Specification: https://www.w3.org/publishing/epub3/epub-overview.html
- EPUB Accessibility 1.0 Specification: https://www.w3.org/Submission/epub-a11y/
- DAISY Accessible Publishing Knowledge Base: https://kb.daisy.org/publishing/docs/
- HTML(5) Specification: https://html.spec.whatwg.org/multipage/
- Digital Publishing WAI-ARIA Module 1.0: https://www.w3.org/TR/dpub-aria-1.0/
- Ace by DAISY Accessibility Checker for EPUB 3: https://daisy.github.io/ace/
- MARC Code List for Relators Scheme: https://id.loc.gov/vocabulary/relators.html
- IANA registry of valid language codes:https://www.iana.org/assignments/language-subtag-registry/language-subtag-registry
- EPUB 2 .ncx Legacy Navigation Document Specification: http://idpf.org/epub/20/spec/OPF\_2.0.1 draft.htm#Section2.4.1
- MathML Specification: https://www.w3.org/TR/MathML/
- ASCIIMath Specification: http://asciimath.org
- International Phonetic Alphabet in Unicode Reference: https://www.phon.ucl.ac.uk/home/wells/ipa-unicode.htm
- Greek and Coptic Letters in Unicode Reference: https://www.fileformat.info/info/unicode/block/greek and coptic/list.htm

### 5.2 Language Dependent Fixed Text Values

The default text values given in this document are, for practical reasons, only given in English. Unless otherwise instructed by the Ordering Agency, these values, when applied to a specific publication, should follow the primary language of that publication.

For convenience, the values for all the main languages of the Ordering Agencies are listed in the tables below. The Ordering Agencies will provide Suppliers with canonical translations if the main language of a title is not one of those listed here.

### 5.2.1 nav.xhtml Headings

English (default)	Swedish	Norwegian (Bokmål)	Norwegian (Nynorsk)	Finnish	Dutch	Danish	Iceland	ic German
Pages Sidindelningste		nin <b>l</b> giste over	Liste over	Sivut	Paginering	Liste	Blaðsíð	urSeiten
		sider	sider			over sider		
Contents	Innehåll	Innhold	Innhald	Sisällys	Inhoud	Indhold	Efni	Inhalt
Landmarks	Navigatio	onNavigasjon	Navigasjon	Navigoii	nt <b>0</b> riëntatiep	u <b>N</b> texrigation	Leiðarn	ne <b>Oxi</b> ientierur

### 5.2.2 aria-label Values

English		Norwegian	Norwegian					
(default)	Swedish	(Bokmål)	(Nynorsk)	Finnish	Dutch	Danish	Icelandic	German
Cover	Omslag	Omslag	Omslag	Kansi	Omslag	Omslag	Kápa	Umschlag
Front	Framsida	Forside	Framside	Etukansi	Voorkant	Forside	Forsíða	Buchvorderseit
cover								
Back cover	Baksida	Bakside	Bakside	Takakans	iAchterkant	Bagside	Bakhlið	Buchrückseite
Jacket left	Vänsterflik	k Venstre	Venstre	Etulieve	Linker	Venstre	Vinstra	Vorderer
flap		innbrett	innbrett		flaptekst	inderflap	innábrot	Klappentext
Jacket	Högerflik	Høyre	Høgre	Takalieve	Rechter	Højre	Hægra	Hinterer
right flap		innbrett	innbrett		flaptekst	inderflap	innábrot	Klappentext
Publisher	Förlagsinfo	ou <b>utgtien</b> inf	o <b>Ungsjen</b> info	rı <b>halsjəis</b> uti	d <b>ditt</b> geversin	fd <b>koletie</b> n	Útgefand	i Verlagsangaber
information			- V					
Dedication	Dedikation	n Dedikasjor	n Dedikasjon	Omistus	Opdracht	Dedikation	Tileinkun	Widmung
Epigraph	Epigraf	Epigraf	Epigraf	Epigrafi	Epigraaf	Epigraf	Tilvitnun	Epigraf
Contents	Innehåll	Innhold	Innhald	Sisällys	Inhoud	Indhold	Efni	Inhalt
Start of	Innehållets	s Start av	Start av	Sisällön	Begin	Indholdets	Byrjun	Beginn
Content	början	innhold	innhald	alku	inhoud	begyndelse		des
	-							Inhalts

### 5.2.3 Image Alternative Text Values

English (default)	Swedish	Norwegian	Finnish	Dutch	Danish	Icelandic	German
Photo	Foto	Foto	Valokuva	Foto	Foto	Ljósmynd	Foto
Illustration	Illustration	Illustrasjon	Kuvitus	Illustratie	Illustration	Mynd	Illustration
Figure	Figur	Figur	Kuvio	Figuur	Figur	Myndefni	Figur
Symbol	Symbol	Symbol	Symboli	Symbool	Symbol	Tákn	Symbol
Map	Karta	Kart	Kartta	Kaart	Kort	Kort	Karte
Drawing	Teckning	Tegning	Piirros	Tekening	Tegning	Teikning	Zeichnung
Comic	Tecknad	Tegneserie	Sarjakuva	Stripverhaal	Tegneserie	Myndasag	aComic
	serie	-	-	_	-		
Logo	Logotyp	Logo	Logo	Logo	Logo	Lógó	Logo

## 5.2.4 schema.org Accessibility Metadata Values

English (default)	Swedish	Norwegian	Finnish	Dutch	Danish	Icelandic	German
This publication conforms to the Nordic Guidelines for the Production of Accessible EPUB 3, version 2020-1.	Den här publikationen är framställd i enlighet med the Nordic Guidelines for the Production of Accessible EPUB 3, version 2020-1.	Denne pub- likasjo- nen er i samsvar med Nordic Guide- lines for the Pro- duction of Acces- sible EPUB 3, version	Finnish  Tämä julkaisu noudat- taa ohjeis- tusta Nordic Guide- lines for the Pro- duction of Acces- sible EPUB 3, version	Dutch  Deze publicatie voldoet aan de Nordic Guide- lines voor productie van toeganke- lijke EPUB3, versie 2020-1.	Danish  Denne publikation er i ov- erensstem- melse med de Nordiske Retningslin- jer for Produktion af Tilgæn- gelighed i Epub 3, version 2020-1.	Icelandic  Pessi aðgengisjöfnun er unnin í samræmi við Norrænu viðmiðunar- reglurnar um aðgengi- legar rafbækur á EPUB3 sniði, útgáfa	Diese Veröffentlichung entspricht den Nordic Guidelines für die Herstellung von barriere- freiem EPUB 3, Version 2020-1.
		2020-1.	2020-1.			2020-1.	

### 5.2.5 Footnote and Endnote Backlink

English							
(default)	Swedish	Norwegian	Finnish	Dutch	Danish	Icelandic	German
Go to the	Gå tillbaka	Gå til	Siirry	Ga naar	Gå til	Aftur í	Gehe
note	till	notereferans	se <b>wi</b> itteeseen.	nootreferent	ienotereferer	nc <b>e</b> ilvísun.	zur
reference.	notreferensen.						Referenz.