



Accessible Publishing

Four common challenges and how to overcome them



Contents

INTRODUCTION	1
CHALLENGE 1: HTML MARKUP & ARIA	2
CHALLENGE 2: MATH, TABLES & MEDIA	5
CHALLENGE 3: IMAGE DESCRIPTIONS	7
CHALLENGE 4: ACCESSIBILITY METADATA	10
CONCLUSION	11
RESOURCES	12



INTRODUCTION:

Accessibility is easier than ever before, so why is it still so hard?

The time is ripe for publishers to produce born-accessible publications. New globally aligned standards and specifications provide a sturdy foundation for accessibility. Publishers are adept at producing EPUB 3, the recommended format for accessible content. Consumers, those requiring assistive technology, such as screen reading software, and the general population, want content that “just works” on whatever platform they choose.

However, the majority of publications produced today are still not accessible. So why is accessibility still so hard?

Ultimately, it is a workflow issue.

Workflows remain largely designed for print. Publications require expensive post-production “fixing” or remediation to make them accessible, versus building them to be accessible from the start.

Thankfully, adopting accessibility best practices is not as hard, or as expensive, as you think. And, if you are already using an XML-first workflow and producing EPUB 3, you are one step ahead of the game.

This guide explores four of the most common and important accessible publishing challenges and how to overcome them.

Accessibility Standards



OWP
Open Web Platform
EPUB 3 is based on this.



WCAG
Web Content Accessibility Guidelines
Provides the foundation for accessibility.



WAI-ARIA
Web Accessibility Initiative /
Accessible Rich Internet Applications
Provides semantics for assistive tech.



IDPF
EPUB Accessibility 1.0
Baseline spec. for accessible publications.



US Section 508,
EU Accessibility Requirements
Global Standards
Summarizes regulations for government entities.

MORE THAN
90%

of published materials are inaccessible to people with dyslexia or blindness. *(United Nations [UN], 2015).*



CHALLENGE 1:

HTML Markup and ARIA

Getting the HTML Markup Right

There is a common misconception that HTML 5 and EPUB 3 guarantee accessibility. They do not. What they do guarantee is **accessibility-ability**, or the ability to be accessible, but the HTML markup has to be right.

Here, we'll walk through a real-life example that illustrates HTML encoding that is very common, but does not provide screen reader software with an appropriate level of semantic tagging to assist users with vision impairment – and better options for correcting it.

Keep in mind, publishers don't necessarily have to do the detailed markup themselves. Many publishing service providers are capable of providing properly accessible HTML markup with marginal extra effort.



HTML and EPUB enable **accessibility-ability**:
the ability to be accessible.

What Not to Do: Overuse <div>

The code pictured to the right looks like HTML, correct? And it is. However, screen readers cannot interpret the <div> tag on every component other than simply reading the undifferentiated text. <div> undermines the HTML structural semantics screen readers depend on, and that enable users to appropriately interact with the content.

```
<div class="chaptitle">CHAPTER TITLE</div>
<div class="noindent1">First paragraph...</div>
<div class="indent">Second paragraph...</div>
<div class="head1">First-level subhead</div>
<div class="noindent1">First paragraph after
subhead...</div>
<div class="image">
  
</div>
```



Better: Use Standard HTML

Standard structural semantics inherent in HTML (sections, headings, paragraphs, image tags, and image descriptions) provide a logical reading order for screen readers to follow. By making this straightforward encoding alteration, publications will work fairly well for most screen readers.

It is also fairly easy to implement standard HTML as most publishing services providers already use it.

```
<section>
  <h1 class="chaptitle">CHAPTER TITLE</h1>
  <p>First paragraph...</p>
  <p>Second paragraph...</p>
  <section>
    <h2 class="subhead1">First-level
    subhead</h2>
    <p> First paragraph after subhead...</p>
    <figure>
      </img>
      <figcaption>Text of caption...</
      figcaption>
    </figure>
  </section>
</section>
```

Best: WAI-ARIA

To create a fully enriched and accessible file, publishers must supplement HTML with the detailed structural semantics in the WAI-ARIA specification (Web Accessibility Initiative-Accessible Rich Internet Applications).

```
<section role="doc-chapter" aria-
labelledby="hd01">
  <h1 class="chaptitle" id="hd01">CHAPTER TITLE
  </h1>
  <p>First paragraph...</p>
  <p>Second paragraph...</p>
  <section aria-labelledby="hd02">
    <h2 class="subhead1" id="hd02">First-level
    subhead</h2>
    <p> First paragraph after subhead...</p>
    <figure>
      </img>
      <figcaption>Text of caption...</figcaption>
    </figure>
  </section>
</section>
```

Next Steps

Open one of your EPUB files and view the markup. How does it compare to the previous examples? If you are already using proper HTML markup, you are one step closer to accessibility. Work with your publishing services partner or other knowledgeable accessibility entity to outline how best to augment your publications with ARIA.

HTML structural semantics:
<section> <aside> <h1>-<h6>



Navigation

HTML markup provides a logical reading order within a single document, but EPUB files commonly contain multiple documents. For screen readers to navigate between and within the different chapters or resources within an EPUB, there needs to be an **EPUB Navigation Document**. To make this interaction possible, accessible publications require the following three navigation elements.

<NAV> ELEMENTS

TOC <nav>

Labels chapters and sections.

Your EPUB 3 files probably already have this.

Landmarks <nav>

Points out important features like glossary or chronology.

Only necessary if you have these features to point out.

Page-list <nav>

Provides corresponding page breaks.

Important for classroom resources. Think cross-references, citations, indexes.

A Note on Page Breaks

While many publishers think page breaks are not important, users of screen readers would disagree. The strongest exemplar for inserting page break coding in EPUB is in education. Page breaks are not just a desire, they are a need. Why? Because content, particularly scholarly academic content, has cross-references, citations, and indexes that all point to page numbers.

Imagine a print-disabled student in the classroom. The professor says, “turn to page 52” or “today’s assignment is on pages 48 to 64”. How does the student find the correct page to turn to?

The Page-List <nav> makes finding those pages possible. It is essential for accessibility when there is a corresponding print version or pre-paginated version, whether or not it is printed on actual paper. However, it is missing in the large majority of EPUBs.

Adding page breaks by hand would be intense and laborious, but publishing services providers can automate this piece of the accessibility puzzle and remove the burden from the publisher.

Links

Another accessibility subtlety that is very simple to resolve, but many publishers are not aware of, is accessible links. The key is to make sure the content of the link tells the user what it is being linked to.

Linking the words “this” or “here” does not provide someone using a screen reader details of what the link will deliver. Here are a few examples of how to address accessible links:

Do

The baseline standard in [EPUB Accessibility 1.0](#).

See [this full explanation](#).

Don't

Find the EPUB Accessibility 1.0 standard [here](#).

See a full explanation [here](#).

The key to accessible links:
link descriptive text, not the word “here”.



CHALLENGE 2:

Math, Tables, and Media

Math and MathML

Publishers may be surprised to know properly formatted math is closer than expected. The proper accessible format for math and equations is MathML. However, almost no publishers include MathML in their EPUBs. EPUB Reading Systems, particularly in retail book environments, have an extremely difficult time interpreting MathML, it usually does not look right, and so it gets left out.

This is a problem for accessibility and a particular problem for education publishers who rely on MathML for classroom texts and materials.

The good news is there are millions of equations already converted to MathML even if they are not included in the final EPUB publication – yet. Where is all this MathML coming from and how can publishers access it?



1. Your Publishing Services Provider

Almost all publishing service providers use MathML as the master equation format in their workflows. At some stage of the editorial production workflow, math is in fact in MathML even if it is not delivered in the final product. You may have MathML even if you don't know you have it!



2. Microsoft Word and Word Equation Editor

In other good news, recent versions of Microsoft Word use an XML format and automatically generate equations in MathML. While it may not be perfect, Word files submitted by authors at the very start of your workflow contain some form of MathML. The Word Equation Editor also produces fairly good MathML. InDesign workflows often use MathType, which will also generate MathML.



3. MathML Cloud

Another helpful, little-known resource is MathML Cloud from Benetech and the DIAGRAM Center. This tool converts AsciiMath, MathML, and LaTeX into PNG and SVG image files complete with alt text – and it's free. Benetech and the Daisy Consortium are also working on a solution so that standard EPUB Reading Systems can better display MathML.

While it may require a little extra work, there is substantial baseline MathML out there and conversion processes to automate the coding. Even if MathML is not included in the EPUB distributed through regular retail channels, hold on to it for accessibility.



Tables

Tables can present challenges. A table may exceed the viewport dimensions of an EPUB reader or the tables may be incorrectly tagged causing interpretation issues.

The most important rule to improve the accessibility of tables is this: include headers on both columns and rows. Without both headers, screen readers will simply read the table from left-to-right and top-to-bottom, which may not provide an accurate interpretation to the user. Avoiding merged and blank

cells also helps improve accessibility.

Disabled Student Services (DSS) offices in colleges and universities all over the world spend an exorbitant amount of unnecessary time remediating tables by hand.

While it is difficult to make tables perfect, they can be improved. Talk to your publishing services provider about how tables are processed and how to augment them for accessibility.

Accessible tables need column and row headers:
`<th scope="row">` and `<th scope="col">`.

Media

Audio-visual content presents exciting opportunities and new challenges for publishers, particularly educational platforms. EPUBs can include audiovisual content, but depending on the type, size, and delivery mechanism of the publication, there are different routes publishers can take.

Audio

For audio, EPUB provides “media overlays” that synchronize text and audio to enable read aloud features. It requires an HTML audio tag and a transcript to function correctly.

Video

Videos require an HTML video tag, captioning, and an audio description that describes what is being shown on the screen. Captioning provides the text a deaf person requires, while the audio description tells a blind or visually-impaired person what is being shown.

Where publishing service providers do not typically create the actual media, they are wonderful resources for integrating audio/visual – sometimes providing transcripts or closed-captioning – and ensuring your publication works exactly as it should for the end user.

The Challenge: File Size

HTML coding for media is not complicated. The real challenge is size. Video in particular drives up file size and makes EPUBs harder to use. It may take so long to download the file that the user gives up before they even get to the first page!

For relatively small books – like children’s books – with less than 30 pages and minimal text on each page, size is generally not an issue.

For longer novels or text books that have narration or supplemental video, many publishers adopt an ebook hybrid scenario.

Media files are referenced in the EPUB with an external URL and are rendered inside the publication only when the device is connected to the internet. Educational publishers with large EPUBs find this model to work well and meet their end-users' appetite for interactive media.



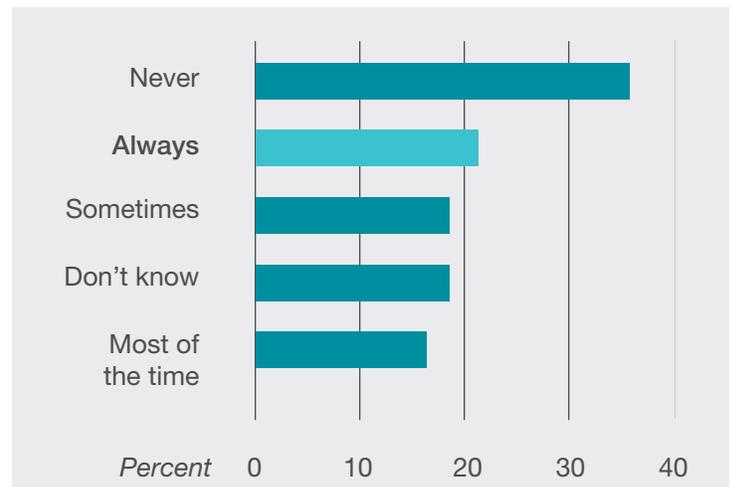
CHALLENGE 3:

Image Descriptions

Image descriptions are one of the most talked about and “sticky” pieces of producing accessible publications. In an anonymous survey of a small group of publishers, only 20% of respondents said they always provide meaningful alt text. That means the majority of the time, alt text is either bad or just not there.

Why is alt text so tricky? Well, not everyone understands the correct way to use, or not use, alt text for different kinds of images. Whether it is a photo, graph, diagram, flowchart, or map, it changes the complexity of the alt text and who is able to provide it. There is a lot of variability and subjectivity that make it difficult.

Do you provide meaningful alt text for images?

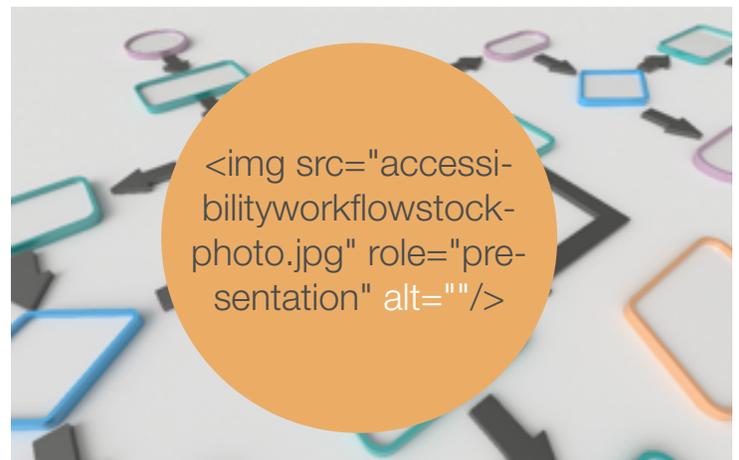


Do not use alt text when:

1. The image is merely decorative
2. The caption sufficiently describes it
3. Surrounding text sufficiently describes it

When Not to Use Alt Text

Before outlining the subtleties of good alt text, it is important to understand when not to use it. Alt text is unnecessary in situations where it will not provide value to the end user. For example, when the image is merely decorative.



Common Bad Practices to Avoid

For the majority of images that do require real alt text, here are a few bad practices publishers and their vendor partners should avoid.

Bad Practice	Why it Matters
Using “image” or leaving the alt text field blank alt=”” alt=”image”	Technically “legal” HTML, but not helpful to users. Screen readers interpret this as “image: image” which is redundant and does not explain what the image is meant to convey.
Using a generic identifier alt=”[filename]” alt=”[figure number]” alt=”bar chart”	Says what the image is by object name or as a high-level description of the image type, but does not explain what the image is meant to convey.
Repeating caption text alt=”[same as caption]”	Very frustrating common occurrence for a visually disabled user. The caption gets read to them twice!
Forgetting about pictures of text alt=”[transcript of text in graphic]”	Title pages like part and chapter are sometimes incorporated as images. Screen readers will not recognize this as text without a transcription included as alt text.

Attaining Meaningful Alt Text

First, what exactly does “meaningful” mean in terms of alt text? Alt text should describe the purpose of an image and why it matters; not just its appearance. What is the image there to convey? Alt text is read aloud to users and its main purpose is to help further his or her understanding of the publication.

Secondly, where does this meaningful alt text come from? The answer is: it depends.

Alt text: What is the image there to convey?



Scholarly or STM publications

The best person to provide an initial image description here is the author. The author knows why the image is included or what it is there to show the reader, which is typically also another scholar or scientist.

Trade Books

Trade books are completely different. Trade book authors often do not specify the images. In this case, an editor may be the best person to provide the image description.

Educational Texts

In education, texts have a pedagogical purpose so the person overseeing the aspects of a larger course program may be the best provider of alt text.

Service Providers

Some publishing services partners are able to provide original alt text that works well; or provide excellent copyediting at the very least.

Ultimately, it is a workflow issue.

Multiple hands may touch an image description before it is considered final.

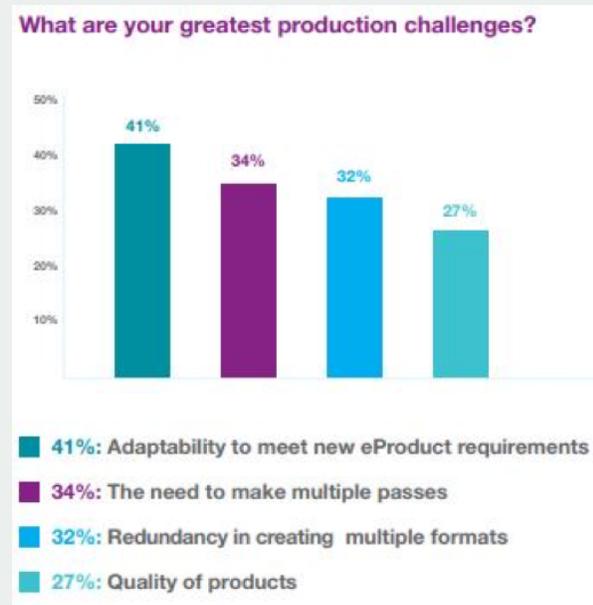
Typically, the author or a subject-matter expert is the best resource to provide the initial content, which can then be passed on to an editor or copyeditor. Ideally, someone who really understands the needs of a print-disabled person can weigh-in, but this is not always a reality.

Depending on the type of publication and intended use, authors, editors, or service providers may all be involved – or not.

At what point in your workflow is the original alt text provided, who provides it, and who ensures it is appropriate and helpful? Outline these questions as the first step to attaining meaningful alt text for accessibility.

SHORT ALT TEXT:

A bar graph shows percentage of unique difficulties that constitute the greatest production challenges.



LONG ALT TEXT :

Data from the graph are summarized as follows:

- 41 percent: Adaptability to meet new eProduct requirements
- 34 percent: The need to make multiple passes
- 32 percent: Redundancy in creating multiple formats
- 27 percent: Quality of products



CHALLENGE 4:

Accessibility Metadata

Accessibility metadata is, thankfully, not difficult. It is simply documentation of how accessible a publication is. It goes into the EPUB file and is aligned with schema.org so the same metadata can also be used online on publisher or retailer websites. Publishing service providers can easily add this metadata, particularly because they are already familiar with the publication and all of its accessibility features.

Accessibility metadata documents two things:

1. The Nature of Accessibility

accessMode	How a person may perceive the information: textual, visual, auditory, or tactile.
accessibilityFeature	Does it have accessibility features like text-to-speech, text-to-braille, speech-to-text?
accessibilityHazard	Can it cause harm? Flashing can cause seizures.
accessibilitySummary	A human-readable explanation of the publications accessibility.
accessModeSufficient	For a figure to be fully comprehensible, the minimum level of content would be text + alt text.

2. Categories of Compliance

Discovery-enabled	“Nature of” metadata, even if it does not meet EPUB Accessibility 1.0.
Accessible	Meets EPUB Accessibility 1.0 and is WCAG-compliant.
Optimized	May not meet all of EPUB Accessibility 1.0, but is optimized for a specific access mode like braille or screen readers.



Conclusion

The day when all publications are created born-accessible is not far off and a good thing for publishers and consumers alike. Just as curb cuts, ramps, and automatic doors benefit both the able-bodied and disabled, more accessible content benefits all readers, whether disabled or not.

And the opportunities accessibility provides are enormous. The U.S. Department of Labor estimates that the 20% of Americans with disabilities have a collective \$175 billion in discretionary spending – and that is just the U.S. In the UK, the estimated

purchasing power of people with disabilities is £80 billion.

The tools, technology, and expertise exist for publishers to start producing born-accessible publications today. XML-first workflows and EPUB 3 provide the foundation. Publishers need only take the next few steps: get the HTML markup right; improve math, tables, and media; provide meaningful image alt text; and don't forget the accessibility metadata.

Ready to move towards more accessible publications?
Contact the experts at Apex CoVantage.

This guide is adapted from content presented in "What you need to know to get accessibility right" webinar hosted by Apex CoVantage and Bill Kasdorf & Associates.



About Apex CoVantage

Apex is a leading supplier of data conversion, editorial, and content-enhancement services to owners of large content collections worldwide. By balancing leading technology and broad industry expertise, Apex delivers exceptional content and media solutions to its customers. Apex has been managing large-scale content projects for libraries, publishers, and media companies around the world for 30 years.

+1.703.709.3000
content@apexcovantage.com
@apex_content
linkedin.com/company/apex-covantage
facebook.com/ApexCoVantage

apexcovantage.com

Apex CoVantage
200 President's Plaza
198 Van Buren Street
Herndon, VA USA 20170-5338



INDEX:

Accessibility Resources

Standards and Recommendations

HTML	w3.org/TR/html/
WCAG 2.1	w3.org/TR/WCAG21/
WAI-ARIA 1.1	w3.org/TR/wai-aria-1.1/
DPUB-ARIA 1.0	w3.org/TR/2017/REC-dpub-aria-1.0-20171214/
ARIA in HTML	w3.org/TR/html-aria/
EPUB Accessibility 1.0	w3.org/Submission/2017/SUBM-epub-a11y-20170125/

Resources

DAISY Knowledge Base	kb.daisy.org/publishing/
BISG Guide	bisg.site-ym.com/store/ViewProduct.aspx?id=6972996
Inclusive Publishing	inclusivepublishing.org/
Benetech DIAGRAM Center	diagramcenter.org/
MathML Cloud	mathmlcloud.org/
Image description resources	diagramcenter.org/making-images-accessible.html
Metadata	www.w3.org/Submission/2017/SUBM-epub-a11y-20170125/#sec-disc-package

Testing and Certification

DAISY Ace Checking Tool	inclusivepublishing.org/toolbox/accessibility-checker/
Benetech Certified Accessible	benetech.org/our-work/born-accessible/certification/

