

03 Accessibility Standards

- Digital Content Standards
- Digital Document Standards
- Metadata Standards that include accessibility tags
- Markup languages that allow provision of accessibility features and tools

Digital Content Standards

Web Content Accessibility Guidelines (WCAG)

Web Content Accessibility Guidelines (WCAG) standards are the most commonly used standards that are mandated in many countries' legal requirements. They are all based on 4 design principles:

1. Perceivable - you have to be able to perceive all available content with senses the end user possesses
2. Operable - you have to be able to use all available content with different interfaces the end user can use
3. Understandable - you have to be able to understand all available content and how to access it with the ability the end user has
4. Robust - you have to be able to use the content in an interoperable and compatible way with third party technologies the end user can use

WCAG has different versions that appear in various legislations across the world.

- The UK requires 2.2, which was released in October 2023
- The EU requires a standard similar to 2.1 (called EN 301 549 Annex A). WCAG 2.1 was released in June 2018
- The US requires 2.0, which was released in December 2008. WCAG 2.0 is identical to ISO/IEC 40500:2012

Generally each version includes everything from the previous one, and adds extra, making them all backwards compatible. Between 2.0 and 2.1 there were 17 additions, and between 2.1 and 2.2 there were 9 additions, and 1 removal (this is Parsing, which has become obsolete). Therefore you could interpret this as the newer version therefore being stricter, higher standards.

The WCAG guidelines are split into 3 levels that increase in strictness, A, AA and AAA, with A being basic or minimum accessibility with 25 success criteria, AA strong with an additional 13 criteria (38 total) and AAA outstanding with an additional 23 criteria (61 total). AA is the default level that is captured in legislative requirements, and reaching WCAG A will not make a digital resource legally compliant. Some aspects of AAA are not applicable in many situations.

The WCAG standards are split into 13 guidelines, that are further split into more detailed success criteria - the number of success criteria depends on the WCAG level.

Count	WCAG Guideline Number	WCAG Guideline Title
1	1.1	Text Alternatives
2	1.2	Time Based Media
3	1.3	Adaptable
4	1.4	Distinguishable
5	2.1	Keyboard Accessible
6	2.2	Enough Time
7	2.3	Seizures and Physical Reactions
8	2.4	Navigable
9	2.5	Input Modalities
10	3.1	Readable
11	3.2	Predictable
12	3.3	Input Assistance
13	4.1	Compatible

More information on the success criteria is available here:

[Web Content Accessibility Guidelines - Quick Reference](#)

[Web Content Accessibility Guidelines in Plain English](#)

The full details of each version of WCAG are available here:

[Web Content Accessibility Guidelines \(WCAG\) 2.2 AA](#)

[Web Content Accessibility Guidelines \(WCAG\) 2.1 AA](#)

[Web Content Accessibility Guidelines \(WCAG\) 2.0 AA](#)

EN 301 549 Annex A

EN 301 549 Annex A is the standard required to be compliant with EU legislation, and it is roughly similar to WCAG 2.1 AA. It has a broader scope than WCAG (which just covers websites), and includes all ICT products and services in the public sector, including specific requirements around

web sites and documents that are both part of the website (HTML or embedded) or downloadable from them (called Non-web Documents).

It includes the same 4 design principles of Perceivable, Operable, Understandable and Robust, and the requirements are mapped to WCAG and directly reference them. This includes the requirements around Non-web Documents, which are described separately to the requirements for Websites.

WAI-ARIA

Web Accessibility Initiative - Accessible Rich Internet Applications (WAI-ARIA) provides an ontology of roles, states, and properties that define accessible user interface elements and can be used to improve the accessibility and interoperability of web content and applications. Accessibility of web content requires semantic information about widgets, structures, and behaviours, in order to allow assistive technologies to convey appropriate information to persons with disabilities. These semantics are designed to allow an author to properly convey user interface behaviours and structural information to assistive technologies in document-level markup.

WAI-ARIA provides Web authors with the following:

- Roles to describe the type of widget presented, such as “menu”, “treeitem”, “slider”, and “progressbar”
- Roles to describe the structure of the Web page, such as headings, and regions
- Properties to describe the state widgets are in, such as “checked” for a check box, or “readonly” for most form controls
- Properties to define live regions of a page that are likely to get updates (such as stock quotes)
- A way to provide keyboard navigation for the Web objects and events, such as those mentioned above

CrossRef recommend tagging DOIs with an ARIA label, more information here: [Accessibility for Crossref DOI Links](#)

Digital Document Standards

Accessible EPUBs

EPUB Accessibility 1.1 addresses two key needs in the EPUB® ecosystem:

- evaluation and certification of accessible EPUB Publications;
- discovery of the accessible qualities of EPUB Publications.

This specification sets formal requirements to meet to certify content as accessible, and provides Authors a clear set of guidelines to evaluate their content against, and allow certification of quality. It is designed to be applicable to EPUB Publications that conform to any version or profile, including future versions of the standard.

An EPUB Publication must meet the following criteria to be accessible per this specification:

- It must meet the requirements for Discoverable EPUB Publications, including providing metadata on accessibility.
- It must meet the requirements for [WCAG 2.0] conformance defined in WCAG Conformance Requirements. This specifies WCAG 2.0 A as mandatory, and AA as recommended.
- It must meet the requirements for EPUB Publications defined in EPUB Requirements. This section covers providing navigation to static page break locations and synchronised text and audio playback.

It must include accessibility conformance metadata as defined in Conformance Reporting. This includes stating which WCAG level (A, AA, AAA) it conforms to, and who provided this certification.

PDF/UA - ISO 14289-1:2014 and ISO 32000-1:2008

ISO 14289-1 is also known as PDF/UA (Portable Document Format, Universal Accessibility), and is aimed at everyone involved in creating PDF. It is based on the PDF standard ISO 32000-1 (also known as Adobe PDF 1.7) and directly references that. It sets minimum requirements that make sure documents are compliant with devices that support people with disabilities. PDF/UA files require the information on their pages to be tagged, and it also allows users to create structure

trees out of tags so that assistive programmes know in which order to read information.

The Matterhorn Protocol is a guide on compliance with PDF/UA: <https://pdfa.org/resource/the-matterhorn-protocol/>

DAISY - Digital Accessible Information System

Digital Accessible Information System (DAISY), also known as ANSI/NISO Z39.86-2005 (R2012), is a digital talking book standard which offers a flexible and navigable reading experience for people who are blind or print disabled, offering a significantly enhanced reading experience—one that is much closer to that of the sighted reader using a print book. A Digital Talking Book (DTB) is a collection of electronic files arranged to present information to the target population via alternative media, namely, human or synthetic speech, refreshable Braille, or visual display, e.g., large print. DAISY multimedia can be a book, magazine, newspaper, journal, computerised text, or a synchronised presentation of text and audio. It provides up to six embedded "navigation levels" for content, including embedded objects such as images, graphics, and MathML. In the DAISY standard, navigation is enabled within a sequential and hierarchical structure consisting of (marked-up) text synchronised with audio.

Metadata Standards that include accessibility tags

There are two ways that metadata accompanies a publication. In the first are digital publication formats that directly embed accessibility metadata (EPUB and PDF). In the second are external metadata record formats (ONIX and MARC) that accompany a digital publication as it moves through the supply chain. In some cases, a digital publication may include both internal and external metadata (e.g., an EPUB could have accessibility metadata in its package document and also be provided to a vendor with an ONIX record).

Source: <https://w3c.github.io/publ-a11y/a11y-meta-display-guide/2.0/draft/guidelines/>

EPUB

EPUB publications **MUST** include the following accessibility metadata:

- **accessMode** — a human sensory perceptual system or cognitive faculty necessary to process or perceive the content (e.g., textual, visual, auditory, tactile).
- **accessibilityFeature** — features and adaptations that contribute to the overall accessibility of the content (e.g., alternative text, extended descriptions, captions).
- **accessibilityHazard** — any potential hazards that the content presents (e.g., flashing, motion simulation, sound).

EPUB publications **SHOULD** include the following [**schema-org**] accessibility metadata:

- **accessibilitySummary** — a human-readable summary of the accessibility that complements, but does not duplicate, the other discoverability metadata. The summary also describes any known deficiencies (e.g., lack of extended descriptions, specific hazards).
- **accessModeSufficient** — a set of one or more access modes sufficient to consume the content without significant loss of information. An EPUB publication can have more than one set of sufficient access modes for its consumption depending on the types of content it includes (i.e., unlike **access modes**, this property takes into account any alternatives for content that is not broadly accessible, such as the inclusion of transcripts for audio).

content).

EPUB creators MAY include additional [[schema-org](#)] accessibility metadata not specified in this section.

There is some companion guidance on [Fixed Layout EPUBs](#) and some guidance on techniques for extracting information from [EPUB Accessibility Metadata](#)

PDF

The PDF/UA standard (see below) defines how to describe accessibility metadata. There is also some companion guidance on creating a [Well Tagged PDF](#)

MARC

[MACHINE-Readable Cataloging \(MARC\)](#) standards are a set of digital formats for the description of items catalogued by libraries, such as books. MARC 21 was designed to redefine the original MARC record format for the 21st century and to make it more accessible to the international community. MARC 21 has formats for the following five types of data: Bibliographic Format, Authority Format, Holdings Format, Community Format, and Classification Data Format.

Within the Bibliographic Format, there are [specific fields to include accessibility metadata](#).

ONIX

[ONIX](#) is an XML-based standard for rich book metadata, providing a consistent way for publishers, retailers and their supply chain partners to communicate a wide range of information about their products. An ONIX record is a separate XML file that is sometimes packaged together with an ebook, and sometimes left separate, but either way, it is distributed alongside an ebook. It contains all kinds of metadata about a book, like title, author, edition, page count, etc. and a set of accessibility metadata.

Most [ONIX Accessibility metadata](#) is carried in the data element. This uses ONIX codelist 196 to specify particular accessibility options that are provided by the product. and Codelist 196 in ONIX provides for granular description of particular accessibility features of the e-book, and can also specify the e-book's conformance with particular accessibility standards and provide links to

further detail. ONIX does not describe reading system functionality. As well as , and and the relevant codes from List 81 are important to highlight content types in the e-book (text, images, audio and the like) that require mitigations for potential inaccessibility.

There some guidance on techniques for extracting information from ONIX Accessibility Metadata for display.

Crosswalk between ONIX and MARC: <https://w3c.github.io/publ-a11y/drafts/a11y-crosswalk-MARC/>

BIBFRAME

Bibliographic Framework (BIBFRAME) was designed to replace the MARC standards, and to use linked data principles to make bibliographic data more useful both within and outside the library community. Bibframe includes the property Contentaccessibility

Schema.org

Schema.org is an initiative launched in 2011 by operators of the world's largest search engines at the time to create and support a common set of schemas for structured data markup on web pages. It includes the CreativeWork type 'Book' and includes several standard accessibility tags, including: AccessibilityFeature, AccessibilitySummary and AccessibilityHazard, plus others.

There are no accessibility sections associated with the following metadata standards

- Dublin Core
- BibTex
- DataCite
- CrossRef
- KBART
- OPDS

Markup languages that allow provision of accessibility features and tools

HTML Living Standard

Hyper Text Markup Language is the ubiquitous web page Markup standard, with [HTML Living Standard](#) being the current version beyond HTML5 that is now constantly updated, although HTML5 is used interchangeably with this. It was originally developed as a language for semantically describing scientific documents. It can be used on web pages, documents and applications.

XML

[Extensible Markup Language](#) is a meta-syntax that can be extended with new custom tags and used across different platforms and devices. In web applications, XML is used to store or transport data, while HTML is used to format and display the same data.

CSS

[Cascading Style Sheets](#) is a style sheet language used for specifying the presentation and styling of a document written in a markup language. It is designed to enable the separation of content and presentation, including layout, colours, and fonts, which can improve content accessibility. Separation of formatting and content also makes it feasible to present the same markup page in different styles for different rendering methods, such as on-screen, in print, by voice (via speech-based browser or screen reader), and on Braille-based tactile devices.

JavaScript

[JavaScript](#) is the scripting language used to programme the behaviour of web pages, while HTML defines the content and CSS the layout/styling. It is a step up from CSS, allowing complex web page functions such as dynamically updating content, control of multimedia, and animated images.

Document Type Definition

A Document Type Definition defines the structure and the legal elements and attributes of an XML document. An application can use a DTD to verify that XML data is valid.

PreTeXt

PreTeXt is a markup language that captures the structure of textbooks and research papers in the mathematical sciences. PreTeXt documents serve as a single source which can be easily converted to multiple other formats, current and future. The best of DocBook, LaTeX, and HTML. Before June 2017, PreTeXt was called “MathBook XML.”

MathML ISO/IEC 40314:2016

ISO/IEC 40314:2016 also known as MathML, MathML is a markup language for describing mathematical notation and capturing both its structure and content. The goal of MathML is to enable mathematics to be served, received, and processed on the World Wide Web, just as HTML has enabled this functionality for text.

LaTeX

LaTeX is a typesetting system which includes features designed for the production of technical and scientific documentation. It is widely used in academia for the communication and publication of scientific documents and technical note-taking in many fields, owing partially to its support for complex mathematical notation. LaTeX is available as free software.

The Difference Between MathML and LaTeX

LaTeX is an input format. It is how we mathematicians write our articles, books, webpages, and anything else where mathematics is involved. (And often anything where mathematics isn't involved.) It is not designed to be read as-is. It is intended to be processed into a suitable output format and then read.

MathML is an output format. It is not designed to be written directly, but it is designed to be read. Of course, one needs a suitable renderer: a browser for the sighted and something like MathPlayer for those who want their mathematics read aloud, but then the same is true of any output format.

It is possible, though not always straightforward, to convert LaTeX to MathML. The main difficulty is that most websites don't bother with this route. They convert the LaTeX mathematics to a graphic which is then displayed, with the original LaTeX as the alt text. Because of how it is produced, the LaTeX is usually very simple (no complicated macros), and so it may be possible to get by with reading the alt text.

So if you want to read mathematics, look for MathML. If you want to write mathematics, learn LaTeX (or another TeX variant).

Source: <https://www.access2science.com/latex/StaceyLatexNote.html>

MusicML

Music Markup Language (MML) is an attempt to mark music objects and events with an XML-based language. Marking such objects should enable managing music documents for various purposes, ranging from music theory and notation to practical performance. This project is not complete and a work in progress. The first draft of a possible DTD is available and a few examples are provided of music pieces marked with MML that result in well-formed as well as valid documents. The approach is modular. Many modules are still incomplete and need more research and attention.

TTML

Timed Text Markup Language provides a standard markup language for synchronising text with media, for example for captions and subtitles. It is widely supported, unifies the increasingly divergent set of existing caption formats, and offers more control over subtitles than simpler formats.

SVG

Scalable Vector Graphics is a language for describing two-dimensional graphics in markup on a web page. This can be advantageous for inclusive design because vector graphics can be easily resized, and scaled up or down to different resolutions without loss of quality. SVG can also be augmented with additional semantics that make them compatible with assistive technologies such as screen readers.

VoiceXML

Voice Extensible Markup Language is a markup language for structuring interactive voice response applications and specifying interactive media and voice dialogs between humans and computers. It is used for developing audio and voice response applications. In order to make these applications accessible to users who are deaf or hard of hearing, the language provides a mechanism for including text alternatives to audio content.

DocBook

DocBook is a markup language for publishing computing and other technical complex scientific documents including books. It allows you to convert one source format into multiple target formats.

DTBook

DTBook or DAISY XML is a markup language used in DAISY Digital Talking Books.