

W3C EPUB Annotation

PMWG 2025, Annotation TF

Scope of the project

- Create an interchange format for notes, bookmarks and reading location.
- Define an Annotation object, enabling the annotation of text, audio, video and images.
- Define a set of Annotation objects relative to the same publication.
- Define how to embed a set of Annotation objects in an EPUB, but also a separate file.
- Recommending selectors perfectly adapted to EPUB.
- Define a Web protocol for interacting with an Annotation Server. Using xAPI?
 - Different from the Annotation Collection?

References

- **W3C Annotation Data Model** (2017, [link](#))
 - Annotation is optimized for Web Resources.
 - Annotation Collection & Page are optimized for paginated requests.
- **W3C Selectors and States** (2017, note, [link](#))
- **Readium Annotations** (2024, [link](#)): A profile of W3C Annotations + Annotation Set (file exchange) + EPUB integration + best practices for reading systems.
- **Hypothes.is** ([link](#)): Web annotations, custom format, uses some W3C Annotation Selectors.
- **Kindle, Kobo, Apple Books, Google Play Books, Moon+Reader** can sync and/or export notes.
- **Readwise.io** ([link](#)): can export "Highlights" to apps or create CSV/markdown. Can import from and to many apps.
- **Zotero** supports EPUB and creates citations.

Other useful specifications?

IDPF Open Annotations (2015)

Web Annotation Protocol (2017)

Web Annotation Vocabulary (2017)

Web Annotation Extensions for Web Publications (2020)

Virtual Locators (GDoc)

Text Fragments (badly specified; WICG); see <https://web.dev/articles/text-fragments?hl=en>

The Accessibility tree, if it helps.

More links

- [DPUB TF on Identifiers](#)

Testimony

*I am writing you because I have recently downloaded your **Thorium Reader** for epub reading (public library ebooks). It is a great app. I have a question and a suggestion for future improvements.*

When I tried to read .annotation file pdf/google chrome/notepad/other apps...I cannot get a Text format, only the code format. How can I get all my annotations easily readable on pdf/doc/txt...?

*Is it possible in the near future to integrate **Zotero**? It would be great to extract all highlights of every epub on zotero software So I can gon on with my academic research using **zotero+thorium on epubs**.*

Thank you very much in advance for your time and your effort.

Felix H. (01/10/2025)

Interested parties

- **Radium Foundation.** 100+ mobile applications, Web applications soon.
- **EDRLab** (Thorium Reader) on desktop, mobile and the Web.
- **Zotero** for exchanging annotations/citations with EPUB reading systems.
- **AllVit**, a Norwegian high-education content provider.
- **Colibrio**, on desktop, mobile and the Web.
- **FBReader**, on desktop, mobile and the Web.
- **Readwise**, a social reading service, as a new export / import format.

Where we are

November 2025

- **Done**

- Use Cases
- Requirements

- **Active:** Body & Target, properties, values; AnnotationSet

- **Later**

- Selectors
- JSON Schema
- New API Protocol?
- Multiple implementations, interchange demos, presentations.

Use Cases & Requirements

- See <https://w3c.github.io/epub-specs/wg-notes/annotations-ucr/>
- Expecting more feedback from
 - Reading system developers.
 - Social reading platforms.

Specification has started

- <https://w3c.github.io/epub-specs/epub34/annotations/index.html>
- First draft edited recently.
 - Description of the model, as a subset of W3C Annotation Data Model
 - Details of the Annotation object
 - Details of the AnnotationSet object
 - Embedding Annotations in EPUB (location in the package, file name)
 - Best Practices for Reading Systems (may become a satellite document)
 - JSON Schema