

Accessible EPUB : Creating an universal accessible format for EPUBs and an IDE to create them

Sachin Rajgopal¹, Thorsten Schwarz¹, Rainer Stiefelhagen¹

Karlsruhe Institute of Technology

Abstract. Currently the electronic accessible document formats, like PDF and word documents, try too similar to their printed form and thus do not utilize the full ability of computers. This paper introduces a universal standard for accessible EPUB 3 formats, which will allow normally sighted, visually impaired and blind to share the same EPUB 3 document. Furthermore a IDE will be presented to simplify the creation process of accessible EPUBs.

1 Introduction

Over the last several years, accessibility is becoming a word with ever increasing importance. Several nations have passed regulatory acts which guarantee equal treatment between all people and that documents have to be accessible for everyone. [1]

Currently, the dominant format for electronic documents are PDFs (Portable Document Format, extension .pdf) and word documents(extension .doc, .docx, .odt). While PDF/UA (PDF/Universal Accessibility) has done much in terms of accessibility, there are still some shortcomings. First of all both formats have a predefined page size. While this is useful for printed documents, a computer screen can rarely display all contents of the document to the detriment of visually impaired people.[2] Therefore an electronic document format with no set document size, containing semantic and structure information and a set reading order would be better suited to meet the demands of accessibility.[2]

1.1 EPUB

EPUB stands for electronic publication and is a format primarily used for books in an electronic format (E-book). The EPUB format was created by the International Digital Publishing Forum (IDPF) and the current version is 3.1, which is a minor update to EPUB 3.[3] EPUB uses XML based formats like XHTML, and thus also uses the accessibility standards and guidelines already established in many nations like the Web Content Accessibility Guidelines (WCAG). [4] This was done as reading systems can have different screen sizes and the EPUB content can therefore be reflowable. Font type and size can also be changed. Visually

impaired people could therefore adjust the document to their preferences. The EPUB 3 specification also contains guidelines for accessibility so these features are built in and not an afterthought.[5]

The EPUB working group has also made some important changes from EPUB 2 to EPUB 3 to make it more accessible. Equations can now be displayed in MathML, there is better navigation and more support for Cascading Style Sheets (CSS). However, not all of these changes are supported yet by many EPUB reading programs and devices.[6] DAISY (Digital Accessible Information System), the audio substitute for print media for the blind, has now been integrated into EPUB 3.[7]

1.2 EPUB Creation Process

An EPUB file is a ZIP file but renamed.[8] Once the file extension is changed and the file contents are extracted, the individual files, like XHTML and image files, can be opened. So creating an EPUB can be described briefly in three steps: create the content documents like XHTML and SVG, then create the package document called package.opf and finally zip up the file with meta-data. The package document has five sections which describe how the EPUB document is structured.[5]

1.3 Existing EPUB editors

Adobe InDesign can be used to create EPUBs and is suitable for publishers, but it does not have inbuilt MathML support, and it is a commercial program which might not be affordable for teachers.

A word document can also be converted to an EPUB file, but there may be some adjustments which have to be made later with an EPUB editor, which usually involves coding. Some EPUB editors are purely WYSIWYG (What You See Is What You Get), but important features like mathematical equations and semantic information tend to be missing. EPUB editors where the program code has to be edited is limited to people with coding experience. Sigil is a WYSIWYG, open source EPUB editor with many features, but important features like text alternatives for images, can only be added with coding. Lastly, many of the editors are not accessible by visually impaired or blind people.

2 Specifications of Accessible EPUB

2.1 EPUB switching mechanism

From the introduction it is apparent that while EPUBs are a promising format, the lack of easily usable editors means that not everyone can create accessible EPUBS. Therefore the main goal of this article is to describe a program called Accessible EPUB, which allows educators to create an EPUB which can be used by blind, visually impaired and regularly sighted people. By changing the CSS

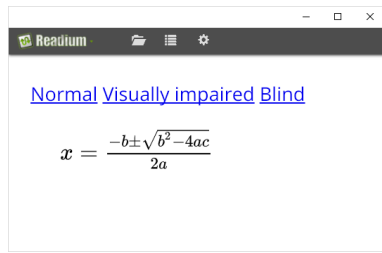


Fig. 1. Quadratic formula shown in 'Normal' mode

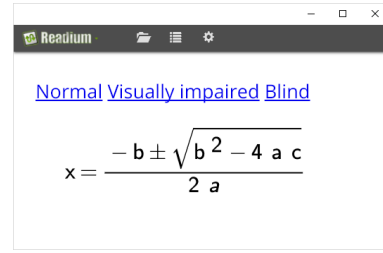


Fig. 2. Quadratic formula shown in 'Visual impairment' mode

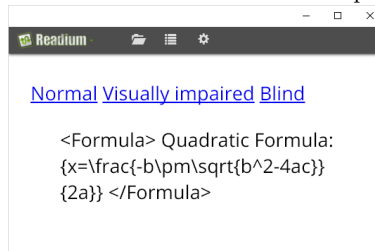


Fig. 3. Quadratic formula shown as LaTeX code in 'Blind' mode

with a click for each target group, the EPUB will be change its content to match the target group's requirements.

For example, the text alternatives for mathematical equations and images should be displayed automatically if the blind version is selected.[5] The switching mechanism can done with either JavaScript or CSS. While using JavaScript is easier to do, not all reading devices support JavaScript nor does the EPUB specification demand that it is supported. [5] The CSS version is much more limited as the whole content has to inserted into a single XHTML file, and CSS selectors are not supported widely yet, but CSS support is in the EPUB specification.

2.2 Accessible EPUB

Accessible EPUB will be a WYSIWYG program written in C# (.NET) and will be for creating accessible EPUB documents with the switching mechanism. The document creation process should be as simple as possible so that people with limited experience with computers can still use the program. A primary target group is teachers with students who have a need for accessible documents.

Some of the main features of Accessible EPUB will be:

- Creating an EPUB with the switching mechanism without coding it in
- An equation editor which accepts LaTeX as input and inserts the MathML equivalent, SVG image and the alternative text of the equation

- Ability to insert images with title, caption and alternative text
- Ability to insert tables
- Mark up sections using the HTML styles (h1, h2, p, etc.)

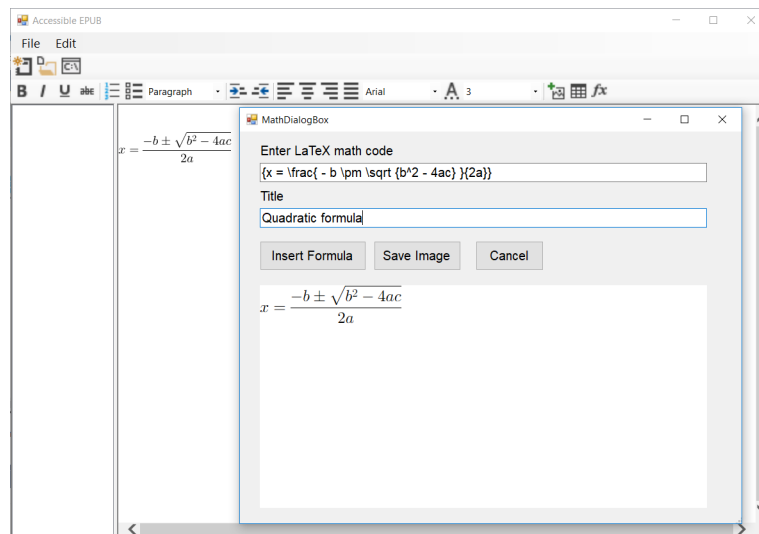


Fig. 4. Accessible EPUB equation editor

3 Outlook

The first iteration of Accessible EPUB is scheduled to be completed by the end of March. In the final paper the project will be presented in its complete form and the results of a evaluation with student tutors and teachers will also be presented here.

References

1. WebAIM (Web Accessibility in Mind): World laws: Introduction to laws throughout the world <https://webaim.org/articles/laws/world/>, visited on 30.01.2018.
2. IDPF: Understanding epub 3 <http://epubzone.org/epub-3-overview/understanding-epub-3>, visited on 29.01.2018.
3. IDPF: Epub 3 specification <http://www.idpf.org/epub/301/spec/epub-overview.html>, visited on 29.01.2018.
4. W3C: Web content accessibility guidelines (wcag) 2.0 <https://www.w3.org/TR/WCAG20/>, visited on 29.01.2018.
5. Garrish, M., Gylling, M.: EPUB 3 Best Practices. O'Reilly (2013)

6. IDPF: Epub 3 changes from epub 2.0.1 <http://www.idpf.org/epub/30/spec/epub30-changes.html>, visited on 29.01.2018.
7. DAISY Consortium: Baseline for accessible epub3 www.daisy.org/baseline, visited on 29.01.2018.
8. Garrish, M.: What is EPUB 3? O'Reilly (2011)