

# Accessible EPUB

Sachin Rajgopal, Thorsten Schwarz

## 1 Introduction

Over the last several years, accessibility is becoming a word with ever increasing importance. Over the last several decades several nations have passed regulatory acts which guarantee equal treatment between all people and. For example, in 1994 Germany passed the Accessibility Ordinance (Behindertengleichstellungsgesetz) which meant that no one can be disadvantaged due to their disability.

As a result of the regulatory acts in many nations it was required to have documents being able to be read by everyone. Everyone is supposed to have access to documents.

Making accessible documents in paper form is difficult, because each group needs to have its document printed separately. Blind people need to have documents printed in braille, which requires special equipment. Furthermore it is very heavy. Visually impaired people have document requirement which person to person, so it is very difficult to prepare printed documents for them.

Making documents accessible in electronic form is much easier, as the document can be adjusted to each group's requirements easily. The output is also adjustable so if a blind person prefers using a screen reader to a refreshable braille display they are free to do so with little to no extra effort. Visually impaired people can change font, size and colors with only a few clicks.

Currently, the dominant format for documents are PDFs (Portable Document Format, extension .pdf) and word documents(extension .doc, .docx, .odt). Unfortunately, they are not always accessible. First of all both formats have a predefined page size. While this is useful when printing documents, a computer screen can rarely display all contents of the document. The font size is also set in stone. While zooming can increase the apparent size of the font, the document width may not fit the screen. Furthermore semantic information normally is missing from the documents. For example, a PDF does not necessarily have the document language defined. A PDF might also not have a predefined reading order which means that the header and footer might be output by the screen reader every page.

Conversely, this means that 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.

## 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. EPUB uses XML based formats like XHTML, and thus also uses the accessibility standards and guidelines already established in many nations. The Web Content Accessibility Guidelines (WCAG) can be followed to create accessible EPUB documents. This was done as reading systems can have different screen sizes and the EPUB content must therefore be reflowable. Font type, size and color can also be changed. Visually impaired people could therefore adjust the document to their preferences. The EPUB3 specification also contains guidelines for accessibility so these features are built in and not an afterthought.

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 Cascading Style Sheets (CSS). However, not all of these changes are supported yet by many EPUB reading programs and devices. DAISY (Digital Accessible Information System), the audio substitute for print media for the blind, has now been integrated into EPUB 3.

### 1.1.1 EPUB Creation Process

An EPUB file is actually just a ZIP file but renamed. Once the file is renamed and 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:

1. Create the content documents like XHTML and SVG
2. Create the package document called package.opf
3. Zip up the file with meta-data

The package document has five sections which describe how the EPUB document is structured.

There are many EPUB editors and most of them do most of the work. A word document can also be converted to an EPUB file, but there may some adjustments which have to be made later with an EPUB editor. 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.

Figure 1: TODO: Add image showing switching mechanism

## 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 to describe a program called Accessible EPUB, which allows people, predominantly teachers, to create an EPUB which can be used by blind, visually impaired and regularly sighted people by changing the CSS on the fly for each target group.

For example, the text alternatives for mathematical equations and images should be displayed automatically if the blind version is selected. This special switching EPUB format can be 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. The CSS version is much more limited as the whole content has to be 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 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 and as straight forward as possible so that people with limited experience with computers can still use the program.