Siteimprove

HOW TO CREATE ACCESSIBLE PDFs

PDF

SITEIMPROVE
ACCESSIBILITY

Id

- **Overview of Guidelines**
- **Document Authoring Tools**
- **Creating the PDF**
- **Checking for Accessibility**

How to create accessible PDFs

In this whitepaper I will cover how to create accessible documents in authoring tools such as MS Word and Adobe InDesign, as well as how to convert them into PDF. I will also cover how to check a PDF for accessibility once it has been created.

To start I will present a brief overview of the main things to consider when making an accessible PDFs.

OVERVIEW

The accessibility level of a document can be scaled from completely inaccessible to greatly accessible.

The basis of an accessible document has two major dependencies:

1) How it was created in the original format

2) How it was turned into a PDF

The way the different page elements and text are defined is of major importance to the accessibility of a document. Assistive technologies interpret the tags of a document and renders content to the user accordingly. A well tagged document enhances accessibility whereas a poorly tagged document introduces major accessibility problems.

Always remember to have your document templates checked for accessibility.

The following can be considered to be of major importance:

Language

In order for screen readers to be able to correctly read the document language there should be an overall definition of the language the document is written in. Furthermore if lines or blocks of text within the document change language, then the text should be tagged separately.

Title

As a minimum documents should include basic information such as a document title. Providing the name of the author, as well as a description and some keywords is also a good idea.

Tagging text

It is important that all text in a document is tagged, be it a paragraph text, a heading, a list, or similar. Not only do you make this evident visually but also clear to all users by providing the correct tagging, as this is what assistive technologies use.

Images

An image can have different purposes depending on how it is used in the document. Many images have a purely decorative purpose and this purpose needs to be conveyed. This is done by giving it a definition of "artifact", described further on. Other images may have some sort of function or convey important information, and therefore these need an alternative text stating this.

Tables

When data tables are used it is important to tag the structure of these. As a minimum defining which are the column / row headings.

Reading order

The sequence in which a screen reader will render the page content depends on how the document was created. Therefore it is important to

ensure that a document has a sensible reading order.

Bookmarks

For many users the easiest and most accessible way to have a table of content is to have bookmarks provided from the headings in the document. Stating these gives the user the option to have a panel open with the document containing the bookmarks.

Exporting the right way

There are many different ways to export to / convert to / save as / PDF. Depending on the way this is done the document can vary from poorly accessible to having great accessibility. So make sure you have the right processes. This is described in detail further on.

Check the document

Even with the best intentions we can sometimes forget a step in the process and therefore it is important to check the PDF for accessibility to make sure there are no issues.

Security Settings

Finally make sure that the document is not locked in a way that makes assistive technologies unable to extract content and render it to the user.

Further resources

If you would like to know more about PDF accessibility issues there are many sources available. One standard that explores all the things that are relevant to creating accessible PDFs is ISO 14289-1 (PDF/UA-1)[1]. The official international guidelines for creating accessible web content,

WCAG 2.0[2], also include many of these. A way of testing for conformance is included in the Matterhorn Protocol[3].

How to work with authoring tools

Authoring tools such as MS Word or Adobe InDesign make it possible to achieve a high standard of accessibility within documents.

MS WORD

For instance in MS Word it is mostly a matter of using the tools as intended. The layout of the different features within the interface depends largely on which version of MS Office you use. I will however mostly cover these features in general.

Language

In MS Word you can define the language you are publishing in by default within the options menu. Additionally you can also change the language of individual sentences if the natural language for part of a text changes. Most often you can see the language of your text at the bottom of the screen.

Document information

Document information such as title, author and keywords can also be added. For instance in MS Word 2010 this is available in 'Document proper-

---

1 http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?ics1=35&ics2=240&ics3=30&csnumber=54564

2 http://www.w3.org/WAI/WCAG20/quickref/Overview.php

3 http://www.pdfa.org/2013/08/the-matterhorn-protocol-1-0/

ties' through the file menu.

Tagging text

Text that is not actively tagged as anything will automatically be tagged as paragraph text. If the text is a heading it is important to not only choose a large font or write in bold, but you must also choose a level of heading (heading 1, heading 2, heading 3...). These headings can be styled anyway you want with font type and color through modifying the heading style. Headings are most often available directly through the top menu.

Lists

If part of the text is a list it is also important to use the intended feature for this, and not just create dots or something similar that looks like a list. MS Word supports bullet lists, numbered lists and multilevel lists. Lists are most often available directly through the top menu.

Tables

When using tables make sure to define table headings. This can be done in MS Word, if these are given in the top row as column headings. In table properties for rows the setting 'Repeat as header row at the top of each page' has a dual function. Besides repeating headings across pages it also tags these table headings as <th> which is code recognizable for screen readers and other user agents.

Images

For images it is possible to provide them with an alternative text. In 'Format Picture' there is an 'Alt text' tab available. Just make sure that in versions where you have the option of writing a title as well as a description, that you use description. Unfortunately it is not possible to define images as 'artifact' (decorative) in MS Word. This has to be done with a remediation tool after converting to PDF.

Columns

If you wish to give content in columns you must use the built in feature for this, most often available

directly through the top menu. Do not simply tab content in a position that looks like columns. If you do this, the reading order will become very poor for screen readers.

Reading order

A documents reading order can vary depending on software version and conversion type. Most often if the document originates in MS Word it will have a logical structure. In some cases images are prioritized last, so if an image conveys information (by having an alternative text) and the position of this is important it needs to be moved to the right position in the reading sequence. This is also possible through remediation tools..

Bookmarks

You can create bookmarks in two ways:

1) From the headings in the document - this option is available by ticking a box once you save the file as a PDF through 'Options'

2) You can also create bookmarks from headings in the document after turning it into a PDF with a remediation tool

Conversion

When you are done creating the document and are ready to turn it into a PDF there are several options available. Depending on the method you use you will create either a tagged PDF or a non-tagged PDF (the latter of which is poorly accessible). If, for instance you choose only to 'print to PDF' then you are saying it only needs to go on a piece of paper and that tagging is of no importance. Therefore tagging is not included, and the file size is smaller. (If you wish to reduce the file size there are other ways of doing this that does not remove accessibility features). Using a method of 'saving as' PDF or 'converting to' PDF will keep the tagging in the document thereby providing accessibility. In later versions of MS Word this is readily available from the file menu. In older versions different plug ins can be downloaded for this purpose.
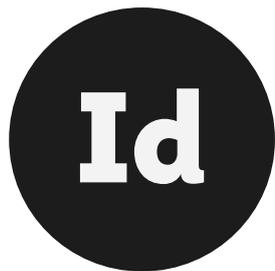
ADOBE INDESIGN

In previous versions of Adobe InDesign it was difficult to work with accessibility. It provided limited features that did not work well. This has changed significantly and especially versions 5.5 and later provide a lot of accessibility features – most of which are well integrated and intuitive to use. This said ensuring accessibility in InDesign still requires a bit of extra work.

The following will mainly focus on accessibility features available in InDesign 5.5 or later.

Language

The language of the document can be set by default. And the language for individual paragraphs can also be set through the 'Character' panel.

Document information

Providing document information such as title, author and keywords can be added. This is done via 'File Info' in the 'Description' tab.

Tagging text

When it comes to tagging text as a paragraph, heading etc. it is possible to map the styles you define with a tag. When you choose one of your styles you can choose 'export tagging' and for PDF define which tag should be mapped to the style. For instance a style you've named 'large header' could be mapped to the H1-tag.

Lists

When creating lists it is important to do so in the right way unless you want to touch up on accessibility afterwards. Using the 'Bulleted and Numbered Lists' feature will automatically create tagged lists (applying bullets or applying numbers). If you insert special characters such as bullet characters on the other hand, this only becomes a visual dot and not a tagged list item. So when using the special character, you need to touch up the tagging in a remediation tool afterwards.

Tables

Tables can be tagged in an accessible way in InDesign if the table is a simple data table. You then have the option of adding table headings for columns / rows, using the 'Convert Rows to Header' feature.

Images

Several features are available to ensure image accessibility. You can define a decorative image as 'artifact', and also give an alternative text for an image. If the image is purely decorative or it does not convey any important information, the image should be defined as 'artifact'. If the image does convey important information or links it should be given an alternative text. You define these through the 'Object' menu. When selecting the image and choosing these options the tabs for stating these becomes available.

Reading order

The reading sequence of the document can be ensured in several ways. In InDesign it can be done for instance by dragging the content into the 'Articles' panel in the desired reading sequence.

For images these can also be placed in a desired location of the reading order. The frame of an image has a small rectangle and by clicking and dragging this into a desired position and dropping, the rectangle turns into an anchor.

Finally remember in the articles panel to choose 'Use for Reading Order in Tagged PDF'.

It is worth mentioning that reading order is one of the few areas that I myself prefer to use a remediation tool to ensure. Some tools provide the ability to simply drag and drop the content into a desired order through a panel.

Bookmarks

An easy way to insert bookmarks is to highlight the headings one by one adding them in the bookmarks panel.

Conversion

After finalising the document, the way in which you convert it into pdf is important. In the 'Export' menu, choose 'Create Tagged PDF' and 'Use Structure for Tab Order'. Tags are then saved within the document and provide accessibility within the pdf document.

Checking the document

A number of tools are available to check PDF documents for accessibility and also to remediate documents. The following provides a general approach to checking the document without stating specific tools.

When opening your converted PDF document you need to ensure a number of things.

Make sure that 'Content Copying for Accessibility' is set to allowed. This information is available through the 'Security Settings' in the 'Document Properties'.

Most remediation tools provide the ability to check the reading order of a document. This is an important check and one that needs human evaluation. Several tools gives a good view of the reading order, Reading order helps ensure that assistive technologies can render the content in a meaningful way. Likewise the tab order should be ensured.

It is also a good idea to check the document title and language settings.

Tagging is something that should also be checked, especially headings, lists, tables and alternatives for images. Depending on the tool you are using, this can be done in a fully automated or semi-automated way. If you created your document in InDesign, remember to check if your role mapping is done correctly; mapping styles to tags. This information can be available through a menu in a tags panel in some remediation tools.

A final check to do is to make sure that the document opens in the desired way (on the first page in the right size).

You can also compress the document to reduce its file size.

Any of these areas that are not defined in the right way should be fixed in the original document if possible. If this is not possible a remediation tool can be used to fix problems.
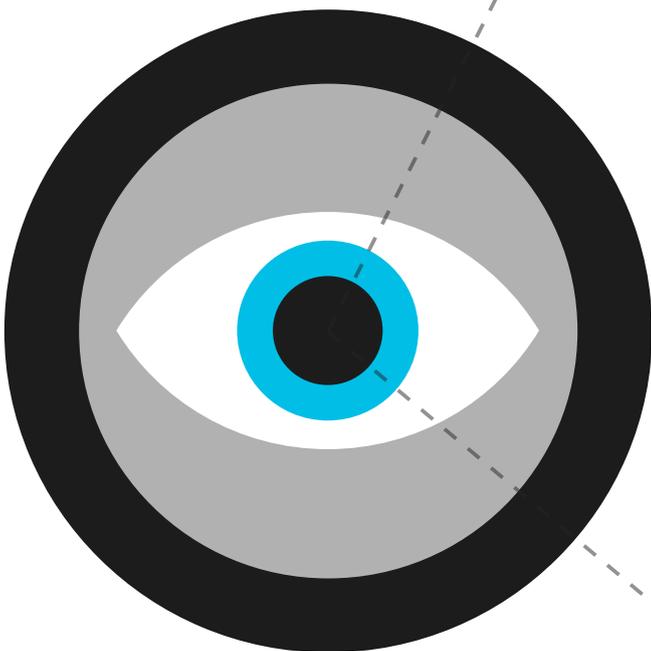
Other things to consider

It is also important to ensure that the color of the background and the color of text are in sufficient contrast to each other. The 'Web Content Accessibility Guidelines' gives guidance on this in term of recommendations for text sizes and compliance level[4].

Also avoid using references to content and information solely based on a location. Some users will receive the content in one long sequence so

for instance a 'box on the right' does not exist to them. Make sure to supplement this by referring to a heading also for instance.

Make sure that the documents can be zoomed to enlarge text without it becoming difficult to read. One example is when text becomes very pixelated.

Images of text should be avoided as they don't work for well for several user groups, such as those with reading difficulties. Text in an image can be identified by the fact that the text cannot be highlighted.

---

4 http://www.w3.org/WAI/WCAG20/quick-ref/#qr-visual-audio-contrast-contrast

CONTRAST vs contrast

Additional help

For both MS Word and InDesign the software providers give detailed guides on how to work with ensuring accessibility. There are also software providers that offer plug ins for creating accessible content.

When it comes to checking the PDF for accessibility several options are available. Some things can be checked in an automated way, whilst

others need to be reviewed manually. Siteimprove Accessibility is an example of an automated tool that checks for PDF accessibility. There are also several providers of remediation tools.

This whitepaper does not intend to recommend any specific tools and should instead be used as a guide when working with ensuring the accessibility of documents.