

Apache™ FOP: Accessibility

Table of contents

1 Overview.....	2
2 Enabling accessibility.....	2
3 Changes to your XSL-FO input files.....	2
4 Customized Tagging.....	2
5 Testing.....	3
6 Limitations.....	3
7 Related Links.....	3

1 Overview

This page describes the [accessibility](#) features of Apache™ FOP. [Section 508](#) defines accessibility in the context of electronic documents for the USA but other countries have similar requirements.

Accessibility features are available only for the PDF output format and there are some implementation limitations. Also, certain actions must be undertaken by the content creator to ensure that FOP can create a truly accessible document.

2 Enabling accessibility

There are 3 ways to enable accessibility:

1. **Command line:** The command line option `-a` turns on accessibility: `fop -a -fo mydocument.fo -pdf mydocument.pdf`
2. **Embedding:** `userAgent.setAccessibility(true);`
3. **Optional setting in fop.xconf file:** `<fop version="1.0"> <accessibility>true</accessibility> ... </fop>`

When accessibility is enabled, additional information relating to the logical structure of the document is added to the PDF. That information allows the PDF viewer (or a text-to-speech application) to retrieve the natural reading order of the document.

Note:

The processing of the logical structure is memory-hungry. You may need to adjust the Java heap size in order to process larger files.

3 Changes to your XSL-FO input files

Apache FOP cannot automatically generate accessible PDFs. Some of the work can only be performed by the content provider. Following are some changes that may be necessary to your XSL-FO content in order to generate really accessible documents:

- Table cells must have a table row as their parent.
- Images must have an alternate text: use the `fox:alt-text` extension attribute (in the [fox namespace](#)) on `fo:external-graphic` and `fo:instream-foreign-object` to specify a short text describing the image.
- Ensure that the order of `fo:block-container` elements in a page corresponds to the reading order.
- Specify the natural language of the document using the `language` and `country` properties (or via the `xml:lang` shorthand property).

4 Customized Tagging

The [PDF Reference](#) defines a set of standard Structure Types to tag content. For example, ‘P’ is used for identifying paragraphs, ‘H1’ to ‘H6’ for headers, ‘L’ for lists, ‘Div’ for block-level groups of elements, etc. This standard set is aimed at improving interoperability between applications producing or consuming PDF.

FOP provides a default mapping of Formatting Objects to elements from that standard set. For example, `fo:page-sequence` is mapped to ‘Part’, `fo:block` is mapped to ‘P’, `fo:list-block` to ‘L’, etc. You may want to customize that mapping to improve the accuracy of the tagging or deal with particular FO constructs. For example, you may want to make use of the ‘H1’ to ‘H6’ tags to make the hierarchical structure of the document appear in the PDF. This is achieved by using the `role` XSL-FO property:

```
...
<fo:block role="H1" font-weight="bold">I. A Level 1 Heading</fo:block>
<fo:block>This is the first paragraph of the first section...</fo:block>
...
```

If a non-standard structure type is specified, FOP will issue a warning and fall back to the default tag associated to the Formatting Object.

5 Testing

Accessible PDFs can be tested, for example, using Adobe Acrobat Professional. Its Accessibility Check feature creates a report indicating any deficiencies with a PDF document. Alternatively, you can just let a screen reader read the document aloud.

6 Limitations

Accessibility support in Apache FOP is relatively new, so there are certain limitations. Please help us identify and close any gaps.

- The natural language can currently only be specified at the page-sequence level. The document language is derived from the language of the first page-sequence. It is currently not possible to override the language inside the content below the page-sequence level.
- It's currently not possible to specify the expanded form of an abbreviation or acronym.
- SVG graphics (or images in general) are treated as a single figure. Text contained in SVGs is not accessible. It's only possible to work with `fo:x-alt-text`.
- The side regions (`region-before`, `region-after` etc.) are currently not specially identified. Screen readers may read their content at page changes.

7 Related Links

Many resources providing guidance about creating accessible documents can be found on the web. Here are a few links, along with additional resources around the topic:

- [US Government - Website on Section 508](#)
- [Wikipedia on Accessibility in general](#)
- [Wikipedia on Accessibility in PDF](#)
- [PDF Reference 1.4](#) (look up chapters 9.7 "Tagged PDF" and 9.8 "Accessibility Support")
- [PDF/A support in Apache FOP](#)
- [Developer-oriented details on the accessibility features \(on the Wiki\)](#)