

Apache FOP Output Formats

\$Revision: 392909 \$

by Keiron Liddle, Art Welch

Table of contents

1 General Information.....	2
1.1 Fonts.....	2
1.2 Output to a Printer or Other Device.....	2
2 PDF.....	2
2.1 Fonts.....	3
2.2 Post-processing.....	3
2.3 Watermarks.....	4
3 PostScript.....	4
3.1 Limitations.....	5
4 RTF.....	5
5 XML (Area Tree XML).....	5
6 Java2D/AWT.....	5
7 Print.....	5
8 Bitmap (TIFF/PNG).....	6
9 TXT.....	6
10 Output Formats in the Sandbox.....	6
10.1 PCL.....	7
10.2 MIF.....	7
10.3 SVG.....	7
11 Wish list.....	7

FOP supports multiple output formats by using a different renderer for each format. The renderers do not all have the same set of capabilities, sometimes because of the output format itself, sometimes because some renderers get more development attention than others.

1. General Information

1.1. Fonts

Most FOP renderers use a FOP-specific system for font registration. However, the Java2D/AWT and print renderers use the Java AWT package, which gets its font information from the operating system registration. This can result in several differences, including actually using different fonts, and having different font metrics for the same font. The net effect is that the layout of a given FO document can be quite different between renderers that do not use the same font information.

1.2. Output to a Printer or Other Device

The most obvious way to print your document is to use the FOP [print renderer](#), which uses the Java2D API (AWT). However, you can also send output from the Postscript renderer directly to a Postscript device, or output from the PCL renderer directly to a PCL device.

Here are Windows command-line examples for Postscript and PCL:

```
fop ... -ps \\computername\printer  
fop ... -pcl \\computername\printer
```

Here is some Java code to accomplish the task in UNIX:

```
proc = Runtime.getRuntime().exec("lp -d" + print_queue + " -o -dp -");  
out = proc.getOutputStream();
```

Set the output MIME type to "application/x-pcl" (MimeTypeConstants.MIME_PCL) and it happily sends the PCL to the UNIX printer queue.

2. PDF

PDF is the best supported output format. It is also the most accurate with text and layout. This creates a PDF document that is streamed out as each page is rendered. This means that the internal page index information is stored near the end of the document. The PDF version supported is 1.4. PDF versions are forwards/backwards compatible.

Note that FOP does not currently support "tagged PDF", PDF/X or PDF/A-1a. [Support for](#)

[PDF/A-1b](#) has recently been added, however.

2.1. Fonts

PDF has a set of fonts that are always available to all PDF viewers; to quote from the PDF Specification: *"PDF prescribes a set of 14 standard fonts that can be used without prior definition. These include four faces each of three Latin text typefaces (Courier, Helvetica, and Times), as well as two symbolic fonts (Symbol and ITC Zapf Dingbats). These fonts, or suitable substitute fonts with the same metrics, are guaranteed to be available in all PDF viewer applications."*

2.2. Post-processing

FOP does not currently support several desirable PDF features: XMP metadata and watermarks. One workaround is to use Adobe Acrobat (the full version, not the Reader) to process the file manually or with scripting that it supports.

Another popular post-processing tool is [iText](#), which has tools for adding security features, document properties, watermarks, and many other features to PDF files.

Warning:

Caveat: iText may swallow PDF bookmarks. But [Jens Stavnstrup tells us](#) that this doesn't happen if you use iText's PDFStamper.

Here is some sample code that uses iText to encrypt a FOP-generated PDF. (Note that FOP now supports [PDF encryption](#). However the principles for using iText for other PDF features are similar.)

```
public static void main(String args[]) {
    try {
        ByteArrayOutputStream fopout = new ByteArrayOutputStream();
        FileOutputStream outfile = new FileOutputStream(args[2]);
        FopFactory fopFactory = FopFactory.newInstance();
        Fop fop = fopFactory.newFop(MimeConstants.MIME_PDF, fopout);

        Transformer transformer = TransformerFactory.newInstance().newTransformer(
            new StreamSource(new File(args[1])));
        transformer.transform(new StreamSource(new File(args[0])),
            new SAXResult(fop.getDefaultHandler()));
        PdfReader reader = new PdfReader(fopout.toByteArray());
        int n = reader.getNumberOfPages();
        Document document = new Document(reader.getPageSizeWithRotation(1));
        PdfWriter writer = PdfWriter.getInstance(document, outfile);
        writer.setEncryption(PdfWriter.STRENGTH40BITS, "pdf", null,
            PdfWriter.AllowCopy);
        document.open();
    }
}
```

```

PdfContentByte cb = writer.getDirectContent();
PdfImportedPage page;
int rotation;
int i = 0;
while (i < n) {
    i++;
    document.setPageSize(reader.getPageSizeWithRotation(i));
    document.newPage();
    page = writer.getImportedPage(reader, i);
    rotation = reader.getPageRotation(i);
    if (rotation == 90 || rotation == 270) {
        cb.addTemplate(page, 0, -1f, 1f, 0, 0,
            reader.getPageSizeWithRotation(i).height());
    } else {
        cb.addTemplate(page, 1f, 0, 0, 1f, 0, 0);
    }
    System.out.println("Processed page " + i);
}
document.close();
} catch( Exception e) {
    e.printStackTrace();
}
}

```

Check the iText tutorial and documentation for setting access flags, password, encryption strength and other parameters.

2.3. Watermarks

In addition to the [PDF Post-processing](#) options, consider the following workarounds:

- Use a background image for the body region.
- (submitted by Trevor Campbell) Place an image in a region that overlaps the flowing text. For example, make region-before large enough to contain your image. Then include a block (if necessary, use an absolutely positioned block-container) containing the watermark image in the static-content for the region-before. Note that the image will be drawn on top of the normal content.

3. PostScript

The PostScript renderer has been brought up to a similar quality as the PDF renderer, but may still be missing certain features. It provides good support for most text and layout. Images and SVG are not fully supported, yet. Currently, the PostScript renderer generates PostScript Level 3 with most DSC comments. Actually, the only Level 3 feature used is FlateDecode, everything else is Level 2.

3.1. Limitations

- Images and SVG may not be displayed correctly. SVG support is far from being complete. No image transparency is available.
- Only Type 1 fonts are supported.
- Multibyte characters are not supported.
- PPD support is still missing.

4. RTF

JFOR, an open source XSL-FO to RTF converter has been integrated into Apache FOP. This will create an RTF (rich text format) document that will attempt to contain as much information from the fo document as possible. The RTF output follows Microsoft's RTF specifications and produces best results on Microsoft Word.

5. XML (Area Tree XML)

This is primarily for testing and verification. The XML created is simply a representation of the internal area tree put into XML. We use that to verify the functionality of FOP's layout engine.

The other use case of the Area Tree XML is as FOP's "intermediate format". More information on that can be found on the page dedicated to the [Intermediate Format](#).

6. Java2D/AWT

Warning:

The Java2D/AWT, Print and Bitmap renderers may not yet fully work as expected. There are some known bugs and missing features.

The AWT viewer shows a window with the pages displayed inside a Java graphic. It displays one page at a time. The fonts used for the formatting and viewing depend on the fonts available to your JRE.

7. Print

It is possible to directly print the document from the command line. This is done with the same

code that renders to the Java2D/AWT renderer.

8. Bitmap (TIFF/PNG)

It is possible to directly create bitmap images from the individual pages generated by the layout engine. This is done with the same code that renders to the Java2D/AWT renderer.

Currently, two output formats are supported: PNG and TIFF. TIFF produces one file with multiple pages, while PNG output produces one file per page. The quality of the bitmap depends on the resolution setting on the FOUserAgent.

9. TXT

The text renderer produces plain ASCII text output that attempts to match the output of the PDFRenderer as closely as possible. This was originally developed to accommodate an archive system that could only accept plain text files, and is primarily useful for getting a quick-and-dirty view of the document text. The renderer is very limited, so do not be surprised if it gives unsatisfactory results.

The Text renderer works with a fixed size page buffer. The size of this buffer is controlled with the textCPI and textLPI public variables. The textCPI is the effective horizontal characters per inch to use. The textLPI is the vertical lines per inch to use. From these values and the page width and height the size of the buffer is calculated. The formatting objects to be rendered are then mapped to this grid. Graphic elements (lines, borders, etc) are assigned a lower priority than text, so text will overwrite any graphic element representations.

Because FOP lays the text onto a grid during layout, there are frequently extra or missing spaces between characters and lines, which is generally unsatisfactory. Users have reported that the optimal settings to avoid such spacing problems are:

- font-family="Courier"
- font-size="7.3pt"
- line-height="10.5pt"

10. Output Formats in the Sandbox

Due to the state of certain renderers we moved some of them to a "sandbox" area until they are ready for more serious use. The renderers and FOEventHandlers in the sandbox can be found under src/sandbox and are compiled into build/fop-sandbox.jar during the main build. The output

formats in the sandbox are marked as such below.

10.1. PCL

Warning:

The PCL Renderer is in the sandbox and not yet functional in FOP Trunk!!! Please help us resurrect this feature.

This format is for the Hewlett-Packard PCL printers. It should produce output as close to identical as possible to the printed output of the PDFRenderer within the limitations of the renderer, and output device.

10.2. MIF

Warning:

The MIF handler is in the sandbox and not yet functional in FOP Trunk!!! Please help us resurrect this feature.

This format is the Maker Interchange Format which is used by Adobe Framemaker.

10.3. SVG

Warning:

The SVG renderer is in the sandbox and may not work as expected in FOP Trunk!!! Please help us improve this feature.

This format creates an SVG document that has links between the pages. This is primarily for slides and creating svg images of pages. Large documents will create SVG files that are far too large for an SVG viewer to handle. Since FO documents usually have text the SVG document will have a large number of text elements. The font information for the text is obtained from the JVM in the same way as for the AWT viewer. If the SVG is viewed on a system where the fonts are different, such as another platform, then the page may look wrong.

11. Wish list

Apache FOP is easily extensible and allows you to add new output formats to enhance FOP's functionality. There's a number of output formats which are on our wish list. We're looking for volunteers to help us implement them.

- [AFP \(Advanced Function Presentation\)](#): The renderer has already been built and is in the

- process of being brought into the Apache FOP codebase.
- [ODF \(Open Document Format\)](#): The standardized successor to OpenOffice's file format.