YaHP Converter (Yet another Html to Pdf converter)

YaHP is a java library which permits to convert an html document into a pdf document.

YaHP is licensed under the LGPL (GNU).

YaHP use a pluggable renderer system. It has currently 3 renderer, each of them carrying advantages and disadvantages.

- The swing renderer, most stable, good rendering, can set header/footer, choose page size.
- The OpenOffice.org renderer, average rendering, cannot set header footer nor page size.(worst renderer)
- The firefox renderer, best rendering, can set header/footer, choose page size, must patch and compile a version of firefox, tested only on linux. Download pre-compiled version of the patched firefox (compiled on ubuntu dapper)mozila-print.tar.gz**!! UPDATED !!**

To use the openoffice.org renderer, you must first launch openoffice in listen mode like this:

oowriter "-accept=socket,port=8100;urp;"

(with the quotes), then set the correct properties (see CSimpleConversion.java).

To use the firefox renderer, you must first download the firefox source : firefox-1.5.0.1-source.tar.bz2 at http://www.mozilla.org/ , then patch it with the files "layout.printing.nsPrintEngine.diff" and "layout.printing.nsPrintEngineH.diff", then compile it by issuing :

"./configure --prefix=/yourchoosenpath/mozilla-print/ --disable-logging --disable-tests --disable-oji --disable-view-source --disable-accessibility --disable-composer --disable-ldap --enable-canvas --disable-gnomeui --enable-application=browser --with-user-appdir=.print.mozilla --enable-system-cairo", then "make install"

Copy the sh script "fireprint" in /yourchoosenpath/mozilla-print/, modify the path inside the script, then set the correct properties (see CSimpleConversion.java). -- script refer to the path /yourchoosenpath/mozilla-print/fireprint.

You also need ps2pdf installed and in the path.

Sample pdf 1

Sample pdf 2

Sample pdf 3

Sample pdf 4 (This page)

Sample pdf 5 (javadoc)

Sample pdf 6 (Google in chinese)

Sample pdf 7 (OpenOffice.org renderer)

Sample pdf 8 (OpenOffice.org renderer)

Sample pdf 9 (OpenOffice.org renderer)

Sample pdf 10 (This page) (OpenOffice.org renderer)

Sample pdf 11 (Google in chinese) (OpenOffice.org renderer)

Sample pdf 12 (Firefox renderer)

Sample pdf 13 (Firefox renderer)

Sample pdf 14 (Firefox renderer)

Sample pdf 15 (Firefox renderer)

Sample pdf 16 (Google in chinese) (Firefox renderer)

Sample pdf 17 (Firefox renderer)

Sample pdf 18 (This page) (Firefox renderer)

Javadoc

!! Download v0.20 (tar.gz) - 07/04/2006 !!

Changes:

- Support header/footer in utf-8 with the firefox renderer.
- Support concurrent rendering with the firefox renderer.
- Better rendering of comboboxes and buttons with the swing renderer, they are painted as vector instead of bitmap.
- Sample application updated.
- Javadoc updated.
- All samples files updated.

Download v0.19 (tar.gz) - 05/04/2006

Changes:

- Can set header/footer and page size with the firefox renderer.
- Recompiled iText to work on 1.4 JVM and corrected a LinkageError on 1.4 JVM.
- Sample application updated to use all the new properties.
- Javadoc updated.
- Firefox samples files updated.

Download v0.18 (tar.gz) - 02/04/2006

Changes:

- Added a new renderer which use firefox as html renderer.
- Sample application updated to use all the new properties.
- Javadoc updated.
- Samples files updated.

Download v0.17 (tar.gz) - 30/03/2006

Changes:

- Renderers are now pluggable.
- Added a new renderer which use OpenOffice.org writer as pdf generator.
- Sample application updated to use all the new properties.
- Refactoring and cleanup of code.
- Does not copy yahpxxx.jar in the temp directory anymore.
- Javadoc updated.
- Samples files updated.

Download v0.16 (tar.gz) - 24/03/2006

Changes:

- Added handling of pdf encryption.
- Added several properties in IHtmlToPdfTransformer interface, see javadoc.
- Cleanup of code.
- Updated the javadoc.
- Updated to the new ShaniXmlParser 1.3.6.
- All samples files updated.

Download v0.15 (tar.gz) - 23/03/2006

Changes:

- Correct rendering of page containing chinese characters.
- Better rendering of button/checkbox components.
- Updated to the new ShaniXmlParser 1.3.6-pre.
- Samples files updated.

Download v0.14 (tar.gz) - 20/03/2006

Changes:

- Intelligent cutting of pages.
- Better rendering of page footer.

Download v0.13 (tar.gz) - 17/03/2006

Changes:

- Detect if rendering in the event thread and avoid calling SwingUtilities invokeAndWait in this case.
- Ensure synchronized rendering of image inside the document.
- Updated to the new ShaniXmlParser 1.3.5.

Download v0.12 (tar.gz) - 16/11/2005

Changes:

- Use SwingUtilities.InvokeAndWait to synchronize with the swing paint thread.
- Better rendering of page header. (page footer still rendered as image)
- Samples files updated.

Download v0.11 (tar.gz) - 07/11/2005

Changes:

- Add property "FAST_TRANSFORM" default to true, which permits to have faster transformation, but will produce black background on transparent gif under kpdf (only so far).
- Circumvent a NullPointerException in JDK ParagraphView class under jdk 1.4.2

Download v0.10 (tar.gz) - 02/11/2005

Changes:

- Set the org.apache.commons.logging.Log System property to force the use log4j instead of setting wrongly with a LogFactoryImplementation.

Download v0.9 (tar.gz) - 28/10/2005

Changes:

- Correct rendering of comboboxes.
- Correct rendering of images with transparent zone (no more black background)
- Set the org.apache.commons.logging.Log System property to force the use of the default logger inside the Yahp context.
- Samples files updated.
- Javadoc updated

Download v0.8 (tar.gz) - 26/10/2005

Changes:

- Corrected a memory leak in the classloader due to commons logging.
- Destroy the classloader on finalization.
- Added the META-INF/services/org.apache.commons.logging.LogFactory file to force the use of the default logger inside the Yahp context.

Download v0.7 (tar.gz) - 23/10/2005

Changes:

- Remove not selected option tag from DOM.
- Updated xml parser.

Download v0.6 (tar.gz) - 24/09/2005

Changes:

- Better rendering quality.
- Render directly in the pdfgraphics2d and dot not use an offscreen buffer which had bad rendering quality.
- Fonts are now vectorized and not as bitmap.
- Samples updated.
- Javadoc updated.

Download v0.5 (tar.gz) - 23/09/2005 (modified, first 0.5 has still buttons display problem, consider the first 0.5 as nuked ;)

Changes:

- Force HTMLEditorKit on the JTextPane used for rendering. (prevent source display)
- Remove the doctype node if any before giving the source to the JTextPane
- Now render correclty the Text field, buttons, combobox, ... (before was blank)

Download v0.4 (tar.gz) - 22/09/2005

Changes:

- Remove the use of TimeoutException in the CMutex clas because this exception only exists in JDK 1.5
- Set the thread context classloader to prevent Duplicate Class.

Download v0.3 (tar.gz) - 21/09/2005

Changes:

- Document/javadoc
- Set antialiasing on the graphics2d object.
- handle '../' in css and image links.

[Download v0.2 (tar.gz)](#)

Changes:

- Use a specialised classloader to load inner jar.
- Compile FOP for jdk1.4 instead of 1.5.

[Download v0.1 (tar.gz)](#)

Contact : quentin.anciaux@advalvas.be