



Getting started with

ScroogeXHTML for the Java™ platform

Version 9.0

LIMITED WARRANTY

No warranty of any sort, expressed or implied, is provided in connection with the library, including, but not limited to, implied warranties of merchantability or fitness for a particular purpose. Any cost, loss or damage of any sort incurred owing to the malfunction or misuse of the library or the inaccuracy of the documentation or connected with the library in any other way whatsoever is solely the responsibility of the person who incurred the cost, loss or damage. Furthermore, any illegal use of the library is solely the responsibility of the person committing the illegal act.

By using this program you accept these responsibilities, and give up any right to seek any damages against the authors in connection with this program.

Specifications subject to change without notice.

Trademarks

Habari is a trademark or registered trademark of Michael Justin in Germany and/or other countries. Android is a trademark of Google Inc. Use of this trademark is subject to Google Permissions. The Android robot is reproduced or modified from work created and shared by Google and used according to terms described in the Creative Commons 3.0 Attribution License. Embarcadero, the Embarcadero Technologies logos and all other Embarcadero Technologies product or service names are trademarks, service marks, and/or registered trademarks of Embarcadero Technologies, Inc. and are protected by the laws of the United States and other countries. Mac and OS X are trademarks of Apple Inc., registered in the U.S. and other countries. Oracle, WebLogic and Java are registered trademarks of Oracle and/or its affiliates. Other brands and their products are trademarks of their respective holders.

Copyright (c) Michael Justin

Contents

Installation.....	5
Requirements.....	5
Android requirements.....	5
Installation.....	5
Directory structure.....	6
Maven and Gradle coordinates.....	6
Important changes in release 9.0.....	7
New features.....	7
Improved support for header and footer text.....	7
Breaking changes.....	7
Moved classes.....	7
Path based methods.....	7
Removed deprecated methods.....	7
Tutorial 1: simple conversion.....	8
What it does.....	8
Java code.....	8
Result HTML.....	9
Source code.....	9
Tutorial 2: additional HTML code.....	10
What it does.....	10
Java code.....	11
Result HTML.....	12
Source code.....	12
Embedding HTML.....	13
Usage of AddOuterHTML.....	13
Character encoding and document type.....	13
Hypertext support.....	14
Overview.....	14
Hyperlink field detection.....	14
Table conversion.....	15
Overview.....	15
Vertical Alignment in Cells.....	15

Size optimization	17
Default Font Properties	17
Example.....	17
Picture extraction	18
PictureAdapter Interface	18
MemoryPictureAdapter.....	18
MemoryPictureAdapterBase64.....	19
Language support	20
How to set the lang Attribute	20
Why you should use the lang attribute on the <html> element.....	20
Cascading Style Sheets	21
Suggested CSS code fragments	21
Post Processing	22
Overview: manipulation of the result DOM tree	22
Technical background.....	22
PostProcessListeners property.....	22
Performance.....	22
Example.....	23
No post processing by default.....	23
The PostProcessListener interface	23
How to: add elements to the HTML head section	24
How to: fix missing http:// in hyperlinks	24
Advanced conversion options	27
Overview	27
SUPPORT_LIST_TABLE.....	27
SUPPORT_STAR_PN.....	28
SUPPORT_MULTILEVEL.....	28
CONVERT_HEADERS_AND_FOOTERS.....	28
CONVERT_PARAGRAPH_BORDERS.....	29
Experimental feature notice	29
Frequently Asked Questions	30
Conversion	30
Why are empty paragraphs not shown in the result page?.....	30
Index	31

Installation

Requirements

ScroogeXHTML for the Java™ platform requires

- Java SE 8
- SLF4J (logging framework)
- JDK 8 for development

Android requirements

- because of the Java 8 requirement, this library only supports Android 7.0 or later

Installation

The library installer is an executable JAR¹ file created with izPack² and works on Microsoft Windows™, Linux™, Solaris™ and Mac OS X™.

A Java Run-time Environment is required to execute it. To launch the installer, double-click it.

The installer will guide you through the installation steps.

1 <https://docs.oracle.com/javase/7/docs/technotes/guides/jar/jarGuide.html>

2 <http://izpack.org/>

Directory structure

```

<inst>

\-- addons                               Add-ons
   \-- ...
\-- apidocs                               JavaDoc documentation
   \-- ...
\-- docs                                  This document
   \-- ScroogeXHTMLGettingStarted.pdf
\-- src                                   Library source code ①
   \-- main                               Test source code ①
   \-- test
\-- pictures                              Picture support code
   \-- ...
\-- Uninstaller
   \-- uninstaller.jar
.installationinformation
ScroogeXHTML-9.0.0.jar                    Precompiled library
ScroogeXHTML-9.0.0-javadoc.jar           Compressed JavaDoc
ScroogeXHTML-9.0.0-sources.jar           Compressed source code ①

① only included in the source edition

```

Maven and Gradle coordinates

Maven

```

<dependency>
  <groupId>com.scroogexhtml</groupId>
  <artifactId>ScroogeXHTML</artifactId>
  <version>9.0.0</version>
</dependency>

```

Gradle

```
implementation 'com.scroogexhtml:ScroogeXHTML:9.0.0'
```

Important changes in release 9.0

New features

Improved support for header and footer text

If the `CONVERT_HEADERS_AND_FOOTERS` option is set, headers and footers will be placed at the beginning and the end of the output document instead of placing them at the top. Important: the converter will emit multiple header and footer tags in the result document if there is more than one RTF header or footer definition.

Breaking changes

Moved classes

`MemoryPictureAdapter` and `MemoryPictureAdapterBase64` moved to new `ScroogeXHTML-Pictures` artifact. Optional `DefaultFontStatistics` class moved to `ScroogeXHTML-Addons` artifact. Optional post processing classes moved to `ScroogeXHTML-Addons` artifact

Path based methods

All file based conversion methods now use `java.nio.Path` instead of `java.io.File`.

Removed deprecated methods

All deprecated methods have been removed

Tutorial 1: simple conversion

What it does

This example converts a hard-coded RTF document to a HTML5 document named 'tutorial-1.html' in the current directory.

It sets the `AddOuterHTML` property which causes generation of surrounding HTML head and body code. The converted HTML is inserted within the body of the document.

Java code

Code example

```
public class Tutorial1 {  
    public static final void main(String[] args) throws IOException {  
        String rtf = "{\\rtf1 {\\b bold \\i Bold Italic \\i0 Bold  
again} \\par}";  
  
        // create a converter instance  
        ScroogeXHTML scrooge = new ScroogeXHTML();  
  
        // configure conversion options  
        scrooge.setAddOuterHTML(true);  
  
        // convert RTF and write HTML to file  
        scrooge.convert(rtf, Paths.get("tutorial-1.html"));  
    }  
}
```


Result HTML

HTML

```
<!DOCTYPE html>
<html>
  <head>
    <META http-equiv="Content-Type" content="text/html; charset=UTF-8">
    <title>Untitled document</title>
    <meta content="ScroogeXHTML for the Java(tm) platform 9.0.0"
name="generator">
  </head>
  <body>
    <p>
      <span style="font-weight:bold;">bold </span><span style="font-
weight:bold;font-style:italic;">Bold Italic </span><span style="font-
weight:bold;">Bold again</span>
    </p>
  </body>
</html>
```

Source code

The full source code is included in `examples/src/main/java/com/scroogexhtml/tutorials`

Tutorial 2: additional HTML code

What it does

This example converts a hard-coded RTF document to a HTML5 document named 'tutorial-2.html' in the current directory.

It shows how a HTML fragment generated by the converter can be embedded in other HTML code and then saved to a file.

Java code

Code example

```
/*
 * ScroogeXHTML for the Java(tm) platform
 * Copyright (c) 1998-2020 Michael Justin
 *
 */
package com.scroogexhtml.tutorials;

import com.scroogexhtml.ScroogeXHTML;
import java.io.IOException;
import java.nio.charset.StandardCharsets;
import java.nio.file.Files;
import java.nio.file.Paths;

/**
 * Tutorial2.
 *
 */
public class Tutorial2 {

    public static void main(String[] args) {

        String rtf = "{\\rtf1 Hello {\\b World} from ScroogeXHTML \\par}";

        // create a converter instance
        ScroogeXHTML scrooge = new ScroogeXHTML();

        // convert RTF and store HTML in String variable
        String converted = scrooge.convert(rtf);

        // wrap with required HTML5 elements
        String html = "<!DOCTYPE html>\n"
            + "<html>\n"
            + "  <head>\n"
            + "    <title>\n"
            + "      Untitled document\n"
            + "    </title>\n"
            + "    <meta http-equiv=\"content-type\"
content=\"text/html; charset=UTF-8\">\n"
            + "  </head>\n"
            + "  <body>\n"
            + "    <p>additional paragraph before</p>\n"
            + converted
            + "    <p>additional paragraph after</p>\n"
            + "  </body>\n"
            + "</html>";

        try {
            writeHtmlFile(html);
        } catch (IOException ex) {
```

```
        ex.printStackTrace(System.err);
    }
}

private static void writeHtmlFile(String html) throws IOException {
    Files.write(Paths.get("tutorial-2.html"),
        html.getBytes(StandardCharsets.UTF_8));
}
}
```

Result HTML

HTML

```
<!DOCTYPE html>
<html>
  <head>
    <title>
      Untitled document
    </title>
    <meta http-equiv="content-type" content="text/html; charset=UTF-8">
  </head>
  <body>
    <p>additional paragraph before</p>
    <p>Hello <span style="font-weight:bold;">World</span> from ScroogeXHTML
  </p>
    <p>additional paragraph after</p>
  </body>
</html>
```

Source code

The full source code is included in the `com.scroogexhtml.tutorial` package of the library unit tests.

Embedding HTML

Usage of AddOuterHTML

If you convert RTF using the methods

- `void convert(String rtf)`
- `void convert(String rtf, Charset charset)`
- `String convert(final ByteArrayInputStream rtf)`

the converter by default returns only the content of the document **body element**, without enclosing it in `<html>...<body>...</body></html>` tags. This fragment can be used in a larger document.

The property `AddOuterHTML` controls whether the enclosing HTML will be generated by the converter. Use `setAddOuterHTML(true)` to switch it on.

For conversions to files, the `AddOuterHTML` property must always be set to `true`. If the property is `false`, the converter will throw a `UnsupportedOperationException`.

Character encoding and document type

Choosing the correct charset³ and document type (HTML5 or XHTML) for the result document is also important.

Note

Always specify the result document charset when you save the HTML to a file or write it to a HTTP response to avoid encoding problems on the receiver side

³ <https://docs.oracle.com/javase/7/docs/api/java/nio/charset/Charset.html>

Hypertext support

Overview

Hyperlink field detection

Many RTF documents use specific hidden fields to store the Hyperlink target and the corresponding display text.

To enable hyperlink conversion of these RTF hyperlink fields, in addition to

```
setConvertHyperlinks(true)
```

also use

```
setConvertFields(true) .
```

If a hidden field does not specify a hyperlink, the converter will only insert the display text (the 'result value' of the hidden field) in the output document.

Table conversion

Overview

ScroogeXHTML for the Java™ platform supports conversion of **simple** RTF tables to HTML.

The library does not convert tables by default. By default, tables in the RTF input document will be converted to text paragraphs.

To enable conversion of simple tables, set the `ConvertTables` property to true.

Code example

```
scrooge.setConvertTables(true);
```

Note

As some RTF document writers create highly complex RTF code for tables, conversion results may not be perfect.

Vertical Alignment in Cells

In HTML, the default value for vertical alignment of table cell content is "middle".⁴

ScroogeXHTML for the Java™ platform assumes that RTF table cell content should be aligned with the top of the cell by default. To keep the output document size small, the conversion adds style elements only for cells which are aligned vertically or with the bottom of the cell.

In order to apply the vertical alignment as the default for table cells, the CSS code `td {vertical-align: top}` must be included.

⁴ See <https://stackoverflow.com/questions/33487148/>

Code example

```
scrooge.setStyleSheetInclude("td {vertical-align: top;}");
```


Size optimization

Default Font Properties

Document size can be optimized with the usage of CSS for frequently used font properties which can be set using the `DefaultFontSize`, `DefaultFontName` and `DefaultFontColor` properties.

Setting the `IncludeDefaultFontStyle` property to true then has these effects:

- if `AddOuterHTML` is true, the HTML head section will contain a CSS definition for the default font style
- the converter will create font style attributes only for text parts which differ from the values of the `DefaultFontSize`, `DefaultFontName` and `DefaultFontColor` properties

Example

If most text in the document uses "Arial, 14 pt, black", set the `DefaultFontSize`, `DefaultFontName` and `DefaultFontColor` properties to these values, and set `IncludeDefaultFontStyle` to true.

If the document is converted with `AddOuterHTML` set to true, the HTML head section will contain the following CSS definition:

Code example

```
<style>
  <!--generated styles-->
  body {font-family:Arial,sans-serif;font-size:14pt;color:#000000}
</style>
```

Picture extraction

PictureAdapter Interface

Picture extraction is activated by assigning a `PictureAdapter` implementation and `setConvertPictures(true)`.

Basic implementations of the `PictureAdapter` interface are provided in the `pictures` folder.

MemoryPictureAdapter

`MemoryPictureAdapter` keeps all extracted picture data in memory and returns hyperlinks to HTTP picture resources, which are then inserted in the result document.

This implementation is useful for web server environments where the server returns the image data back to the client. In the most simple implementation, the server keeps the image data in memory for the duration of a client session, and returns the image data dynamically when the browser requests the image resource URLs. Of course this requires HTTP session management and sufficient memory.

Example for a link element:

Code example

```

```

The image URLs will be numbered automatically.

The class allows to set a base path with `setBase(String base)`, for example `adapter.setBase("/images/")`, so that the result URL will be `"/images/image1.png"`.

MemoryPictureAdapterBase64

`MemoryPictureAdapterBase64` extends `MemoryPictureAdapter` but returns Image Data URIs for pictures which do not exceed a given maximum size. For larger images, it will return the external image URL as defined by its super class.

By default, the size threshold is set to 32 kB. The threshold can be set with the `maxSize` constructor argument.

Data URIs are fully supported by most major browsers, and partially supported in Internet Explorer and Microsoft Edge.

Code example

```
scrooge = new ScroogeXHTML();  
scrooge.setConvertPictures(true);  
PictureAdapter adapter = new MemoryPictureAdapterBase64();  
scrooge.setPictureAdapter(adapter);  
  
// run the conversion  
...
```

Example for a Data URI link:

Code example

```
 element

By using `lang`, you get the benefits of hyphen support in your (modern) browser that you otherwise would not get (assuming you use `hyphens: auto` in your CSS).

Add `hyphens: auto` in your CSS:

### Code example

```
scrooge.setStyleSheetInclude("p {hyphens: auto;}");
```

## References

<http://blog.adrianroselli.com/2015/01/on-use-of-lang-attribute.html>

<https://www.w3.org/International/questions/qa-lang-why>

## Cascading Style Sheets

### Suggested CSS code fragments

---

Suggestions for your custom CSS in documents which use tables:

#### Code example

```
scrooge.setStyleSheetInclude("body, p {\n"
 + " margin: 0px;\n"
 + "}\n"
 + "table {\n"
 + " border-collapse: collapse;\n"
 + "}\n"
 + "td {\n"
 + " vertical-align: top;\n"
 + "}\n");
```

| Code                                              | Reference                                                                               |
|---------------------------------------------------|-----------------------------------------------------------------------------------------|
| <pre>body, p {   margin: 0px; }</pre>             | Reduce space between lines                                                              |
| <pre>table {   border-collapse: collapse; }</pre> | TABLE collapse default is "separate", with collapse, adjacent cells share their borders |
| <pre>td {   vertical-align: top; }</pre>          | <a href="#">Vertical Alignment in Cells</a> (HTML default is "middle")                  |

# Post Processing

## Overview: manipulation of the result DOM tree

---

### Technical background

The converter internally uses an XML DOM tree to create the HTML document structure. Before converting the DOM to the result HTML5 or XHTML string, the converter calls a sequence of post processing handlers, which apply optimizations and custom modifications on the DOM tree. Post processing handlers must implement the `PostProcessListener` interface.

### PostProcessListeners property

The converter stores the event handlers in its `PostProcessListeners` property which is a list of `PostProcessListener` implementations.

### Performance

Post processing may cause a significant increase of the conversion time.

## Example

### Code example

```
// create a converter instance
ScroogeXHTML scrooge = new ScroogeXHTML();

// create and add a post processor
PostProcessListener listener = new PostProcessListenerExample();

scrooge.getPostProcessListeners().add(listener);

// run the conversion
...
```

## No post processing by default

The converter creates and adds no post process handlers by default (on creation).

## The PostProcessListener interface

---

The converter stores a list of post progress listeners in its `PostProcessListener` property. The listeners implement the `PostProcessListener` interface, which has one method, `postProcess`.

The converter passes an instance of `PostProcessEventObject` to the `postProcess` method. This event object carries references to the converter and the `org.w3c.dom.Document` instance with the result DOM.

## How to: add elements to the HTML head section

---

This example shows how a post process listener can be used to add a meta date element to the HTML head.

### Code example

```
public class Main {

 public static void main(String... args) {
 ScroogeXHTML scrooge = new ScroogeXHTML();
 scrooge.setAddOuterHTML(true);

 scrooge.addPostProcessListener(new PostProcessListener() {
 @Override
 public void postProcess(PostProcessEventObject e) {
 Document doc = e.getDocument();
 Element html = doc.getDocumentElement();
 Node head = html.getFirstChild();

 // add meta date
 Element metaDate = doc.createElement("meta");
 metaDate.setAttribute("name", "date");
 metaDate.setAttribute("content", new Date().toString());
 head.appendChild(metaDate);
 }
 });

 String html = scrooge.convert("{\\rtf1 Hello world}");
 System.out.println(html);
 }
}
```

## How to: fix missing http:// in hyperlinks

---

This example converts a RTF documents which contains a simple (blue and underlined formatted) hyperlink.



## ScroogeXHTML for the Java™ platform 9.0

The hyperlink text does not begin with a valid protocol name such as `http://` and this causes a non functional hyperlink in the result HTML:

### HTML

```
<p>
 example.com
</p>
```

To fix this, we want to apply post processing code which modifies all `<a>` elements so that they begin with `http://`

The result should be:

### HTML

```
<p>
 example.com
</p>
```

Our solution will use the XPath expression `//a[not(contains(@href, '://'))]` to find all `<a>` elements in the document whose `href` attribute do not contain the character sequence `://"`.

For all found elements, our code then inserts `"http://"` in the value of the `href` attribute.

### Notes

- this is pure demonstration code
- there is no guarantee that the result `href` value will be a valid internet address

## Source code

The following source code example shows the `PostProcessListener` implementation and its `postProcess` method.

The full source code is included in the examples folder (`Tutorial3.java` in package `com.scroogexhtml.tutorials`).

## Code example

```

// create a converter instance
ScroogeXHTML scrooge = new ScroogeXHTML();

// we want simple HTML output for this example
scrooge.setConvertFontSize(false);
scrooge.setConvertFontName(false);

// enable hyperlink conversion
scrooge.setConvertHyperlinks(true);

// add post process listener
scrooge.getPostProcessListeners().add(new PostProcessListener() {
 @Override
 public void postProcess(PostProcessEventObject e) {
 try {
 XPathFactory xpathFactory =
XPathFactory.newInstance();
 // XPath to find hyperlink nodes.
 XPathExpression xpathExp =
xpathFactory.newXPath().compile(
 "//*[not(contains(@href, '://'))]");
 NodeList links = (NodeList)
xpathExp.evaluate(e.getDocument(), XPathConstants.NODESET);
 for (int i = 0; i < links.getLength(); i++) {
 Element a = (Element) links.item(i);
 String href = a.getAttribute("href");
 a.setAttribute("href", "http://" + href);
 }
 } catch (XPathExpressionException ex) {
Logger.getLogger(Tutorial3.class.getName()).log(Level.SEVERE, null, ex);
 }
 }
});

// convert RTF to HTML
scrooge.setAddOuterHTML(true);
scrooge.convert(rtf, Paths.get("tutorial-3.html"));

```

## Advanced conversion options

### Overview

This table lists available conversion options, their valid values, defaults, and which area of the conversion they affect.

Key	Values	Area
SUPPORT_LIST_TABLE <sup>5</sup>	yes   <b>no</b>	Lists
SUPPORT_STAR_PN <sup>6</sup>	yes   <b>no</b>	Lists
SUPPORT_MULTILEVEL <sup>7</sup>	yes   <b>no</b>	Lists
CONVERT_HEADERS_AND_FOOTERS <sup>8</sup>	yes   <b>no</b>	Text
CONVERT_PARAGRAPH_BORDERS <sup>9</sup>	yes   <b>no</b>	Paragraphs

### SUPPORT\_LIST\_TABLE

This conversion option enables support for the RTF list table (Word 97); the value must be “yes” or “no”. The list table is a section in the RTF header which stores styles for numbered and unnumbered lists. The document then refers to a specific style by its id.

Note

This option is experimental and unsupported because some RTF writers generate malformed list tables. Use with caution.

<sup>5</sup> Experimental, introduced in version 5.3 (UseListTable property)

<sup>6</sup> Experimental, introduced in 7.0

<sup>7</sup> Experimental, introduced in 7.0

<sup>8</sup> Since 6.3.0

<sup>9</sup> Since 7.0 it is possible to disable paragraph border box conversion (introduced in version 6.5)

## SUPPORT\_STAR\_PN

This conversion option enables support for list formatting based on `\*\pn` RTF tokens (Word 6.0/95 RTF); the value must be "yes" or "no".

Note	This option is <u>experimental and unsupported</u> . Use with caution.
------	------------------------------------------------------------------------

## SUPPORT\_MULTILEVEL

This conversion option enables support for multilevel numbered and unnumbered lists in conjunction with `SUPPORT_LIST_TABLE` (Word 97); the value must be "yes" or "no".

Note	This option is <u>experimental and unsupported</u> because some RTF writers generate malformed list tables. Use with caution.
------	-------------------------------------------------------------------------------------------------------------------------------

### Code example

```
// create a converter instance
ScroogeXHTML scrooge = new ScroogeXHTML();

// enable experimental list table support
scrooge.setOutputProperty(ConversionKeys.SUPPORT_LIST_TABLE, "yes");

// enable experimental multi-level support
scrooge.setOutputProperty(ConversionKeys.SUPPORT_MULTILEVEL, "yes");

// run the conversion
```

## CONVERT\_HEADERS\_AND\_FOOTERS

This conversion option enables support for header and footer text conversion. By default header and footer will not appear in the output document; the value must be "yes" or "no".

Note	This option is <u>experimental and unsupported</u> . The converter will emit multiple header and footer tags in the result document if there is more than one RTF header or footer definition.
------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## CONVERT\_PARAGRAPH\_BORDERS

This conversion option enables support for paragraphs which are formatted with a simple border box; the value must be "yes" or "no".

### Code example

```
// enable paragraph border conversion
scrooge.setOutputProperty(ConversionKeys.CONVERT_PARAGRAPH_BORDERS,
"yes");
```

If consecutive paragraphs use the same border style, they can be merged with a post processor.<sup>10</sup>

### Code example

```
// merge identical paragraph borders
scrooge.addPostProcessListener(new MergeBorderDivNodes());
```

This post processor removes all identical style (div) nodes between paragraphs and replaces them with one parent div node.

## Experimental feature notice

---

- Advanced conversion options are [experimental and unsupported features](#)
- If you use them, be aware that not all RTF writers generate correct or conforming code
- Experimental features may be removed or significantly changed in future releases

# Frequently Asked Questions

## Conversion

---

### Why are empty paragraphs not shown in the result page?

HTML browsers do not show empty/white space only `<p>` elements. Example:

#### RTF view

```
Line 1
Line 2

Line 3
```

will look different in the HTML browser

#### Browser view

```
Line 1
Line 2
Line 3
```

You can set the `ConvertEmptyParagraphs` property to true. The result HTML then will contain `<br>` or `<br />` instead of empty `<p>` elements, and look as expected.

# Index

## Reference

AddOuterHTML.....	8, 13, 17	Hyperlink.....	14
Border-collapse.....	21	Hyperlinks.....	24
ByteArrayInputStream.....	13	Hypertext.....	14
ConvertEmptyParagraphs.....	30	Hyphens:auto.....	20
ConvertFields.....	14	IncludeDefaultFontStyle.....	17
ConvertHyperlinks.....	14	Installation.....	5
ConvertPictures.....	18	Language support.....	20
ConvertTables.....	15	Margin.....	21
CSS.....	21	MemoryPictureAdapter.....	18
Data URI.....	19	MemoryPictureAdapterBase64.....	19
Default Font Properties.....	17	PictureAdapter.....	18
DefaultFontColor.....	17	Post Processing.....	22
DefaultFontName.....	17	PostProcess.....	23
DefaultFontSize.....	17	PostProcessEventObject.....	23
DefaultLanguage.....	20	PostProcessListener.....	22p.
DOM tree.....	22	Size optimization.....	17
Empty paragraphs.....	30	SLF4J.....	5
Experimental.....	29	Table.....	15
Fragment.....	10	Tutorial.....	8
Head.....	24	UnsupportedOperationException.....	13
Hidden fields.....	14	Vertical-align.....	21
HTML5.....	10, 13, 22	XHTML.....	13, 22
Http.....	25	XPath.....	25