



链滴

使用 itext 将 html 页面转 pdf(完美支持中文)

作者: [washmore](#)

原文链接: <https://ld246.com/article/1528278305572>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

最近有用到使用itext将html页面转pdf的操作,做个笔记

技术预研

受限于某些外部原因(该环境无法使用java执行sh脚本),wkhtmltopdf被排除在外,最终选择了itext全家, [itextpdf](#)作为pdf生成核心, [xmlworker](#)用于html转pdf, [itext-asian](#)用于支持中文显示,视情况而定还可引入[nekohtml](#)对非标准html页面进行标签补全.

引入maven依赖

```
<dependency>
  <groupId>com.itextpdf</groupId>
  <artifactId>itextpdf</artifactId>
  <version>5.5.11</version>
</dependency>
<dependency>
  <groupId>com.itextpdf.tool</groupId>
  <artifactId>xmlworker</artifactId>
  <version>5.5.11</version>
</dependency>
<dependency>
  <groupId>com.itextpdf</groupId>
  <artifactId>itext-asian</artifactId>
  <version>5.2.0</version>
</dependency>
```

核心功能代码

```
import com.itextpdf.text.Document;
import com.itextpdf.text.Font;
import com.itextpdf.text.PageSize;
import com.itextpdf.text.pdf.BaseFont;
import com.itextpdf.text.pdf.PdfWriter;
import com.itextpdf.tool.xml.XMLWorkerFontProvider;
import com.itextpdf.tool.xml.XMLWorkerHelper;

import java.io.ByteArrayInputStream;
import java.io.InputStream;
import java.io.OutputStream;
import java.nio.charset.Charset;

public class PDFUtil {
  public static void writeStringToOutputStreamAsPDF(String html, OutputStream os) {
    writeToOutputStreamAsPDF(new ByteArrayInputStream(html.getBytes()), os);
  }

  public static void writeToOutputStreamAsPDF(InputStream html, OutputStream os) {
    try {
      Document document = new Document(PageSize.A4);
      PdfWriter pdfWriter = PdfWriter.getInstance(document, os);
      document.open();
      XMLWorkerHelper worker = XMLWorkerHelper.getInstance();
```

```

        worker.parseXHtml(pdfWriter, document, html, Charset.forName("UTF-8"), new AsianFontProvider());
        document.close();
    } catch (Exception e) {
    }
}
}
/**
 * 用于中文显示的Provider
 */
class AsianFontProvider extends XMLWorkerFontProvider {
    @Override
    public Font getFont(final String fontname, String encoding, float size, final int style) {
        try {
            BaseFont bfChinese = BaseFont.createFont("STSongStd-Light", "UniGB-UCS2-H", BaseFont.NOT_EMBEDDED);
            return new Font(bfChinese, size, style);
        } catch (Exception e) {
        }
        return super.getFont(fontname, encoding, size, style);
    }
}

```

其中writeToOutputStreamAsPDF方法接受输入流形式的html文件,然后将pdf写入到输出流中,而writeStringToOutputStreamAsPDF方法则接受文本形式的html文件,方便做一些模板替换的预处理工作,然也可以结合vm,fm等模板引擎进行模板文件数据渲染,请自行了解,OutputStream输出流可作为文件出流写入到生成PDF文件,也可以作为web环境下的response提供PDF文档显示/下载功能