

[脚本之家社区](#) [服务器常用软件](#) [源码商城](#) [脚本云](#)

- [手机版](#)
- [投稿中心](#)
- [关注微信](#)

扫一扫

[快捷导航](#)

[软件下载](#)

[android](#) [MAC](#) [驱动下载](#) [字体下载](#) [DLL](#)

[源码下载](#)

[PHP](#) [ASP.NET](#) [ASP](#) [JSP](#)

[软件编程](#)

[C#](#) [JAVA](#) [C语言](#) [Delphi](#) [Android](#)

[网络编程](#)

[PHP](#) [ASP.NET](#) [ASP](#) [JavaScript](#)

[在线工具](#)

[CSS格式化](#) [JS格式化](#) [Html转化为Js](#)

[数据库](#)

[MYSQL](#) [MSSQL](#) [oracle](#) [DB2](#) [MARIADB](#)

[CMS](#)

[PHPCMS](#) [DEDECMS](#) [帝国CMS](#) [WordPress](#)

[常用工具](#)

[PHP开发工具](#) [python](#) [Photoshop](#) [必备软件](#)

- [网站首页](#)
- [网页制作](#)
- [网络编程](#)
- [脚本专栏](#)
- [脚本下载](#)
- [数据库](#)
- [服务器](#)
- [电子书籍](#)
- [操作系统](#)
- [网站运营](#)
- [平面设计](#)
- 其它 [媒体动画](#) [电脑基础](#) [硬件教程](#) [网络安全](#)
- [vbs](#)
- [DOS/BAT](#)
- [hta](#)
- [htc](#)
- [python](#)
- [perl](#)
- [游戏相关](#)
- [VBA](#)

- [远程脚本](#)
- [ColdFusion](#)
- [ruby专题](#)
- [autoit](#)
- [seraphzone](#)
- [PowerShell](#)
- [linux shell](#)
- [Lua](#)
- [Golang](#)
- [Erlang](#)
- [其它](#)

-
-
-
-
-
-

您的位置: [首页](#) → [脚本专栏](#) → [python](#) → Python将HTML转成PDF

Python实现将HTML转成PDF的方法分析

更新时间: 2019年05月04日 12:38:50 作者: Tacey Wong
这篇文章主要介绍了Python实现将HTML转成PDF的方法,结合实例形式分析了Python基于pdfkit模块实现HTML转换成PDF文件的相关操作技巧与注意事项,需要的朋友可以参考下

本文实例讲述了Python实现将HTML转成PDF的方法。分享给大家供大家参考，具体如下：

主要使用的是wkhtmltopdf的Python封装——pdfkit

安装

1. Install python-pdfkit:

```
?  
l$ pip install pdfkit
```

2. Install wkhtmltopdf:

- Debian/Ubuntu:

```
?  
l$ sudo apt-get install wkhtmltopdf
```

- Redhat/CentOS

```
?  
lsudo yum intsall wkhtmltopdf
```

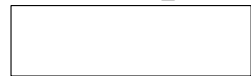
- MacOS

```
?  
lbrew install Caskroom/cask/wkhtmltopdf
```

使用

一个简单的例子:

```
1import pdfkit
2pdfkit.from_url('http://google.com', 'out.pdf')
3pdfkit.from_file('test.html', 'out.pdf')
4pdfkit.from_string('Hello!', 'out.pdf')
```



你也可以传递一个url或者文件名列表:

```
1pdfkit.from_url(['google.com', 'yandex.ru', 'engadget.com'], 'out.pdf')
2pdfkit.from_file(['file1.html', 'file2.html'], 'out.pdf')
```



也可以传递一个打开的文件:

```
1with open('file.html') as f:
2    pdfkit.from_file(f, 'out.pdf')
```



如果你想对生成的PDF作进一步处理, 你可以将其读取到一个变量中:

```
1# 设置输出文件为False, 将结果赋给一个变量
2pdf = pdfkit.from_url('http://google.com', False)
```



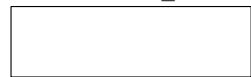
你可以制定所有的 wkhtmltopdf 选项 <<http://wkhtmltopdf.org/usage/wkhtmltopdf.txt>>. 你可以移除选项名字前面的 '--'. 如果选项没有值, 使用None, False or * 作为字典值:

```
1 options = {
2     'page-size': 'Letter',
3     'margin-top': '0.75in',
4     'margin-right': '0.75in',
5     'margin-bottom': '0.75in',
6     'margin-left': '0.75in',
7     'encoding': "UTF-8",
8     'no-outline': None
9 }
10 pdfkit.from_url('http://google.com', 'out.pdf', options=options)
```



默认情况下, PDFKit 将会显示所有的 wkhtmltopdf 输出. 如果你不想看到这些信息, 你需要传递一个 quiet 选项:

```
1 options = {
2     'quiet': ''
3 }
4 pdfkit.from_url('google.com', 'out.pdf', options=options)
```



由于wkhtmltopdf的命令语法, TOC 和 Cover 选项必须分开指定:

```

2
1toc = {
2  'xsl-style-sheet': 'toc.xsl'
3}
4cover = 'cover.html'
5pdfkit.from_file('file.html', options=options, toc=toc, cover=cover)

```



当你转换文件、或字符串的时候，你可以通过css选项指定扩展的 CSS 文件。

```

2
1# 单个 CSS 文件
2css = 'example.css'
3pdfkit.from_file('file.html', options=options, css=css)
4# Multiple CSS files
5css = ['example.css', 'example2.css']
6pdfkit.from_file('file.html', options=options, css=css)

```



你也可以通过你的HTML中的meta tags传递任意选项：

```

2
1 body = ""
2   <html>
3     <head>
4       <meta name="pdfkit-page-size" content="Legal"/>
5       <meta name="pdfkit-orientation" content="Landscape"/>
6     </head>
7     Hello World!
8   </html>
9   ""
10pdfkit.from_string(body, 'out.pdf') #with --page-size=Legal and --orientation=Landscape

```



配置

每个API调用都有一个可选的参数。这应该是pdfkit.configuration() API 调用的一个实例. 采用configuration选项作为初始化参数。可用的选项有：

- **wkhtmltopdf**——wkhtmltopdf二进制文件所在的位置。默认情况下pdfkit 会尝试使用which (在类UNIX系统中) 或 where (在Windows系统中)来判断。
- **meta_tag_prefix** -- pdfkit的前缀指定 meta tags（元标签） - 默认情况是pdfkit-

示例：针对wkhtmltopdf不在系统路径中（不在\$PATH里面）：

```

2
1config = pdfkit.configuration(wkhtmltopdf='/opt/bin/wkhtmltopdf')
2pdfkit.from_string(html_string, output_file, configuration=config)

```



问题

- IOError: 'No wkhtmltopdf executable found':

确保 wkhtmltopdf在你的系统路径中（\$PATH），会通过 configuration进行了配置 (详情看上文描述)。在Windows系统中使用 where wkhtmltopdf命令 或在 linux系统中使用 which wkhtmltopdf会返回 wkhtmltopdf二进制可执行文件所在的确切位置。

- IOError: 'Command Failed'

如果出现这个错误意味着 PDFKit不能处理一个输入。你可以尝试直接在错误信息后面直接运行一个命令来查看是什么导致了这个错误（某些版本的 wkhtmltopdf会因为段错误导致处理失败）

- 正常生成，但是出现中文乱码

确保两项：

- 1)、你的系统中有中文字体
- 2)、在html中加入<meta charset="UTF-8">

下面是我随便写的一个HTML表格：

```
2
1 <html>
2 <head><meta charset="UTF-8"></head>
3 <body>
4 <table width="400" border="1">
5   <tr>
6     <th align="left">Item...</th>
7     <th align="right">1</th>
8   </tr>
9   <tr>
10    <td align="left">衣服</td>
11    <td align="right">$241.10</td>
12  </tr>
13  <tr>
14    <td align="left">化妆品</td>
15    <td align="right">$30.00</td>
16  </tr>
17  <tr>
18    <td align="left">食物</td>
19    <td align="right">$730.40</td>
20  </tr>
21  <tr>
22    <th align="left">tOTAL</th>
23    <th align="right">$1001.50</th>
24  </tr>
25</table>
26</body>
27</html>
```



下面是生成的PDF截图

□

另：https://pdfcrowd.com/#convert_by_input

更多Python相关内容感兴趣的读者可查看本站专题：《[Python文件与目录操作技巧汇总](#)》、《[Python编码操作技巧总结](#)》、《[Python数据结构与算法教程](#)》、《[Python函数使用技巧总结](#)》、《[Python字符串操作技巧汇总](#)》及《[Python入门与进阶经典教程](#)》

希望本文所述对大家Python程序设计有所帮助。

您可能感兴趣的文章：

- [Python3转换html到pd的不同解决方案](#)
- [Python实现抓取HTML网页并以PDF文件形式保存的方法](#)
- [python将html转成PDF的实现代码\(包含中文\)](#)
- [Python使用到第三方库PyMuPDF图片与pdf相互转换](#)
- [Python合并同一个文件夹下所有PDF文件的方法](#)
- [Python使用reportlab模块生成PDF格式的文档](#)
- [Python多图片合并PDF的方法](#)
- [python使用pdminer解析pdf文件的方法示例](#)
- [Python生成pdf文件的方法](#)
- [Python实现将DOC文档转换为PDF的方法](#)

如您对本文有所疑义或者有任何需求，请[点击访问脚本社区](#)，百万网友为您解答！

原文链接: <http://www.cnblogs.com/taceywong/p/5643978.html>

□

微信公众号搜索“脚本之家”，选择关注

程序猿的那些事、送书等活动等着你

- [Python](#)
- [HTML](#)
- [PDF](#)

相关文章

-
-

[Python中关键字nonlocal和global的声明与解析](#)

这篇文章主要给大家介绍了关于Python中关键字nonlocal和global的声明与解析的相关资料，文中介绍的非常详细，相信对大家具有一定的参考价值，需要的朋友们下面来一起看看吧。

- 2017-03-03

[Python3 jupyter notebook 服务器搭建过程](#)

这篇文章主要介绍了Python3 jupyter notebook 服务器搭建过程，非常不错，具有一定的参考借鉴价值，需要的朋友参考下吧

- 2018-11-11

[通过字符串导入 Python 模块的方法详解](#)

这篇文章主要介绍了通过字符串导入 Python 模块的方法详解，本文通过实例结合，给大家介绍的非常详细，具有一定的参考借鉴价值,需要的朋友可以参考下

- 2019-10-10

[Python tkinter label 更新方法](#)

今天小编就为大家分享一篇Python tkinter label 更新方法，具有很好的参考价值，希望对大家有所帮助。一起跟随小编过来看看吧

- 2018-10-10

[python执行使用shell命令方法分享](#)

这篇文章主要介绍了python执行使用shell命令方法分享，具有一定参考价值，需要的朋友可以了解下。

- 2017-11-11

[使用 Python 在京东上抢口罩的思路详解](#)

受疫情影响全民真的在抢口罩，而且还是抢不到的那种。这篇文章主要介绍了用 Python 在京东上抢口罩的实例代码，代码简单易懂，非常不错，具有一定的参考借鉴价值,需要的朋友可以参考下

- 2020-02-02

[python使用paramiko实现远程拷贝文件的方法](#)

这篇文章主要介绍了python使用paramiko实现远程拷贝文件的方法,分析了paramiko库的安装以及远程下载文件的实现技巧,需要的朋友可以参考下

- 2016-04-04

[python方法生成txt标签文件的实例代码](#)

今天小编就为大家分享一篇python方法生成txt标签文件的实例代码，具有很好的参考价值，希望对大家有所帮助。一起跟随小编过来看看吧

- 2018-05-05

[Python3.5基础之函数的定义与使用实例详解【参数、作用域、递归、](#)

这篇文章主要介绍了Python3.5基础之函数的定义与使用,结合实例形式详细分析了Python3.5函数的定义、参数、作用域、

递归、重载、内置函数等基本概念与相关使用技巧,需要的朋友可以参考下

2019-04-04

[pycharm下打开、执行并调试scrapy爬虫程序的方法](#)

本篇文章主要介绍了pycharm下打开、执行并调试scrapy爬虫程序的方法,小编觉得挺不错的,现在分享给大家,也给大家做个参考。一起跟随小编过来看看吧

2017-11-11

最新评论

大家感兴趣的内容

- [1Python入门教程 超详细1小时学会](#)
- [2Pycharm 2020最新永久激活码\(附](#)
- [3Python 列表\(List\)操作方法详解](#)
- [4Python 元组\(Tuple\)操作详解](#)
- [5Python 字典\(Dictionary\)操作详解](#)
- [6pycharm使用心得\(一\) 安装和首](#)
- [7python strip\(\)函数 介绍](#)
- [8python 中文乱码问题深入分析](#)
- [9python中使用xlrd、xlwt操作exce](#)
- [10python逐行读取文件内容的三种方](#)

最近更新的内容

- [Python模块包中__init__.py文件功能分析](#)
- [python实现汽车管理系统](#)
- [Python实现常见的回文字符串算法](#)
- [PyQt 图解Qt Designer工具的使用方法](#)
- [EM算法的python实现的方法步骤](#)
- [Python实现的文本对比报告生成工具示例](#)
- [python各类经纬度转换的实例代码](#)
- [对python 多个分隔符split 的实例详解](#)
- [django 自定义filter 判断if var in list](#)
- [从零学Python之入门\(四\) 运算](#)

在线工具

[代码格式化等](#)

高防主机

[600G 防护](#)

枫信科技

[IDC服务商](#)

常用在线小工具

- [CSS代码工具](#)
- [JavaScript代码格式化工具](#)
- [在线XML格式化/压缩工具](#)
- [php代码在线格式化美化工具](#)
- [sql代码在线格式化美化工具](#)
- [在线HTML转义/反转义工具](#)
- [在线JSON代码检验/检验/美化/格式化](#)
- [JavaScript正则在线测试工具](#)
- [在线生成二维码工具\(加强版\)](#)
- [更多在线工具](#)

微信 [投稿](#) [脚本任务](#) [在线工具](#)

关注微信公众号

扫一扫



[关于我们](#) - [广告合作](#) - [联系我们](#) - [免责声明](#) - [网站地图](#) - [投诉建议](#) - [在线投稿](#)

©Copyright 2006-2020 JB51.Net Inc All Rights Reserved. 脚本之家 版权所有