

---

# beautifulsoup Documentation

发布 *4.4.0*

**delong**

**9月 15, 2016**







**Beautiful Soup** 是一个可以从HTML或XML文件中提取数据的Python库.它能够通过你喜欢的转换器实现惯用的文档导航,查找,修改文档的方式.**Beautiful Soup**会帮你节省数小时甚至数天的工作时间.

这篇文档介绍了BeautifulSoup4中所有主要特性,并且有小例子.让我来向你展示它适合做什么,如何工作,怎样使用,如何达到你想要的效果,和处理异常情况.

文档中出现的例子在Python2.7和Python3.2中的执行结果相同

你可能在寻找 **Beautiful Soup3** 的文档,**Beautiful Soup 3** 目前已经停止开发,我们推荐在现在的项目中使用**Beautiful Soup 4**, [移植到BS4](#)

这篇帮助文档已经被翻译成了其它语言: