



链滴

# Jsoup 入门

作者: [feng](#)

原文链接: <https://ld246.com/article/1497518580091>

来源网站: [链滴](#)

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

# Jsoup入门

## 简介

jsoup 是一款Java 的HTML解析器，可直接解析某个URL地址、HTML文本内容。它提供了一套非常力的API，可通过DOM，CSS以及类似于jQuery的操作方法来取出和操作数据。主要作用是 用HttpClient获取到网页后 具体的网页提取需要的信息的时候，就用到Jsoup，Jsoup可以使用强大的类似jQuery.css选择器，来获取需要的数据；

```
HttpEntity entity=response.getEntity(); // 获取返回实体
String webContent=EntityUtils.toString(entity, "utf-8");
Document doc=Jsoup.parse(webContent);
Elements elements=doc.getElementsByTag("title");
Element element=elements.get(0);
String title=element.text();
System.out.println("网页标题是: "+title)
```

通过一个字符串获取一个document对象，通过dom解析获取内容

再举例一些常用的方法

getElementById(String id) 根据id来查询DOM

getElementsByTag(String tagName) 根据tag名称来查询DOM

getElementsByClass(String className) 根据样式名称来查询DOM

getElementsByAttribute(String key) 根据属性名来查询DOM

getElementsByAttributeValue(String key,String value) 根据属性名和属性值来查询DOM

还有一种封装过的方法 Jsoup也可以使用css jquery选择器方式获取元素节点

```
Document doc= Jsoup.connect("http://www.evafjs.cn").get();
```

```
Element e = doc.select("选择器");
```

我感觉Jsoup和js操作dom树差不多，有js的基础就好了

针对**Json数据**的解析，这里我建议一款快速解析工具**fastjson**（阿里开源的一个工具）

github:[Jsoup](#)