

version 1.9.2.

Open source

# jsoup

[find](#) and extract data, using DOM traversal or CSS selectors

scrape and [parse](#) HTML from a URL, file, or string

[manipulate](#) the HTML elements, attributes, and text

[clean](#) user-submitted content against a safe white-list, to prevent XSS attacks

[output](#) tidy HTML

Java 1.5



Maven

```
<dependency>
<groupId>org.jsoup</groupId>
<artifactId>jsoup</artifactId>
<version>1.9.2</version>
</dependency>
```

Example

```
Document doc = Jsoup.connect("http://en.wikipedia.org/").get();
Elements newsHeadlines = doc.select("#mp-itn b a");
```

/jsoup

<https://jsoup.org/>

/jsoup/www.wenzlaff.de

<http://www.wenzlaff.de>

