

1-Selenium安装

1-什么是selenium

2-PhantomJs介绍

3-Chromedriver介绍

2-Selenium入门

加载页面:

定位和操作:

查看请求信息:

退出

3-selenium更多方法

selenium定位页面元素

获取元素的数据与操作

1-Selenium安装

```
1 pip3 install Selenium
```

1-什么是selenium

如果想在python爬虫操作动态的数据,就必须学习selenium自动化测试工具, 而使用selenium的前提是需要无头浏览器(PlantomJs无界面浏览器)

Selenium是一个Web的自动化测试工具, 最初是为网站自动化测试而开发的, Selenium 可以直接运行在浏览器上, 它支持所有主流的浏览器, 可以接收指令, 让浏览器自动加载页面, 获取需要的数据, 甚至页面截屏

2-PhantomJs介绍

PhantomJS 是一个基于Webkit的“无界面”(headless)浏览器, 它会把网站加载到内存并执行页面上的 JavaScript

- 下载地址: <http://phantomjs.org/download.html>
- 或者使用命令, 步骤如下:

```
1 wget https://bitbucket.org/ariya/phantomjs/downloads/phantomjs-2.1.1-linux-x86_64.tar.bz2
```

- 解压并创建软连接:

```
1 tar -xvzf phantomjs-2.1.1-linux-x86_64.tar.bz2
```

```
2 sudo cp -R phantomjs-2.1.1-linux-x86_64 /usr/local/share/
3 sudo ln -sf /usr/local/share/phantomjs-2.1.1-linux-x86_64/bin/phantomjs /usr/local/bin/
```

3-Chromedriver介绍

谷歌浏览器安装参考: <https://www.linuxidc.com/Linux/2017-12/149245.htm>

Chromedriver 也是一个能够被selenium驱动的浏览器, 但是和PhantomJS的区别在于它是有界面的

- 下载地址: <https://npm.taobao.org/mirrors/chromedriver>
- driver的安装, 最简单的安装方式是: 解压后把bin目录下的可执行文件移动到环境变量下, 比如/usr/bin 或者是/usr/local/bin 下面

```
1 ----- Linux 下安装方式
2 chromedriver_linux64.zip : 版本 ChromeDriver v2.22 (2016-06-06)
3 (支持 Chrome v49-52, 当前Ubuntu虚拟机下的Chrome版本为50.0)
4
5
6 1. unzip chromedriver_linux64.zip
7 2. sudo chmod +x chromedriver
8 3. sudo mv chromedriver /usr/local/bin/
9
10 ----- Windows 下安装方式
11 chromedriver_win32.zip : 版本 ChromeDriver v2.32 (2017-08-30)
12 (支持 Chrome v59-61)
13
14 1. 解压 chromedriver_win32.zip
15 2. 将 chromedriver 移动到 Python安装目录下
16 (其实放哪都可以, 但是需要配环境变量, 放Python安装目录下省事)
17
18 注意: Chromedriver和电脑上的chrome版本有对应关系, 建议使用最新的Chromedriver版本并且更新chrome
```