

數據採集| selenium實現自動翻頁的5個方法 轉載

mob60475703a599 2021-06-03 15:41:14

©著作權

文章標籤 爪哇 文章分類 爪哇 編程語言 閱讀數 1859

前言

採集某些網站的數據時，需要分頁獲取。

一般情況下，獲取總頁碼，然後構建分頁鏈接分別採集就可以。

但偶爾，請求地址並不會隨著頁碼變化而變化，這時就需要使用selenium 瀏覽器自動化測試工具來進行翻頁操作。

selenium 實現翻頁的前提

1. 下載瀏覽器驅動

我使用的是谷歌瀏覽器。

下載地址：

<http://chromedriver.storage.googleapis.com/index.html>

瀏覽器驅動需與瀏覽器版本一致。

下載後，將驅動所在地址添加到系統環境變量中。

2. 安裝並導入selenium 模塊

已配置環境變量：

```
1. from selenium import webdriver
2. driver = webdriver.Chrome()
3. driver.get(url)
```

未配置環境變量，需指定驅動所在地址：

```
1. from selenium import webdriver
2. driver = webdriver.Chrome("D:/chromedriver.exe")
3. driver.get(url)
```

selenium 實現翻頁的5個方法

1. 直接定位“下一頁”文本超鏈接並點擊

```
1. next = driver.find_element_by_link_text("下一頁")
2. next.click()
```

2. 按照標籤屬性逐步定位到“下一頁”並點擊

```
1. pagelist = driver.find_element_by_class_name("pagelist")
2. next = pagelist.find_elements_by_class_name("nextpage")
3. next.click()
```

3. 通過文本定位到“下一頁”標籤並點擊

```
1. next = driver.find_element_by_xpath("//*[text()='下一頁']")
2. next.click()
```

4. 在搜索框循環輸入頁碼，並回車

```
1. from selenium import webdriver
2. # 导入 Keys 用于模拟键盘操作
```



mob60475703a599

21 原創 36.3萬 人氣

0 翻譯 2645 轉載

[+ 關注](#)

近期評論

- 推薦學習 21天4周 更文挑戰 數據 對著樓主寫的博客，挺好
- Dubbo常用功能07--服務 哎呦，不錯哦
- java學習之為什麼要有h 現在工作中用到了，挺不錯
- Why Memory Barriers中 這位博主，你好可愛
- yaml 文件詳解 哎呦，不錯哦

近期文章

- 1.Prometheus新特性：分
- 2.通過Shared Informer擴
- 3.基於Raft構建大型分佈式
- 4.Kubernetes調度器101
- 5.使用Java開發Kubernete

2021年637篇	2020
2019年234篇	2018
2017年272篇	2016
2015年190篇	2014
2013年147篇	2012
2011年107篇	2010
2009年39篇	2008
2007年31篇	2006
2005年2篇	

熱評好文

- VisualSVN Server啟動錯
- 大數據之Zookeeper：zc

```
6. # 定位到页码输入框
7. input = driver.find_element_by_xpath('//*[@id="Page_input"]')
8. # 清空内容
9. input.clear()
10. # 输入页码
11. input.send_keys(page)
12. # 回车, 实现翻页
13. input.send_keys(Keys.ENTER)
14. # 执行JS脚本翻页
```

```
1. js = "document.getElementsByClassName('pagelist')[7].click()"
2. driver.execute_script(js)
```

--- 結尾 ---

相關閱讀

數據採集| 使用Python爬取51job



自在不思量

个人微信号: nanyinbeiyi

欢迎交流 欢迎指正 欢迎分享

<https://mp.weixin.qq.com/s/AGI0PpIF2-zQjJ9kzn5fZg>



贊



收藏



評論



分享



舉報

上一篇: 來感受Linux命令行的“真香定律”

下一篇: 迷茫之際, 有所放棄, 才能有所選擇



提問和評論都可以, 用心的回復會被更多人看到

評論

相關文章

selenium實現翻頁

selenium翻頁使用selenium的keys方法發送按鍵# 導包from selenium.webdriver.common.keys import Keys# 向下...



七日熱門

- Java入門_Java概述_Ja
- Java|Java面向對象
- Java-Java反射
- [Java] JAVA IO操作
- 【java】java的Enum
- 【Java】Java反射筆記
- 【Java基礎】Java拷貝
- java JAVA_HOME
- 【Java筆記】Java多態
- Java映射到Java bean

21天
更文挑戰

相關標籤

html5表單自動翻頁 idea
現的方法 java實現翻頁
頁 jquery自動翻頁 my
集 python selenium 方法
m翻頁 python實現翻頁
翻頁代碼

文章目錄

- 前言
- selenium 實現翻頁的前
- selenium 實現翻頁的5

基本思路: 首先用開發者工具找到需要提取數據的標籤列表: 利用xpath定位需要提取數據的列表然後再逐個提取相...

python爬蟲：利用selenium實現自動翻頁爬取某魚數據

本文的文字及圖片來源於網絡,僅供學習、交流使用,不具有任何商業用途,版權歸原作者所有,如有問題請及時聯繫...

Python數據採集案例(2)：Selenium實現的貓眼網播熱度採集

作者：長行時間：2020.05.26 Github原文：Week-04/Example-0402 實現目標本案例通過圖文詳細介紹使用Seleni...

自動翻頁

鼠標放到翻頁的地方，自動翻頁<LINK href="1/css.css" type=text/css rel=stylesheet> <TABLE cellSpacing=0 cellPadding=5 wid...

selenium 自動化的坑（5）

這次要說的自動化坑是關於<a>標籤的，話不多說，先上圖：這個表面上看起來是個輸入框，操作的時候需要鼠標...

PHP數據採集curl常用的5個例子

用php，curl主要是抓取數據，當然我們可以用其他的方法來抓取，比如fsockopen,file_get_contents等。但是只能...

數據採集需要的方法

使用 System；使用 System.Data；使用 System.Configuration；使用 System.Web；使用 System.Web.Security；使用 System...

Freemarker顯示一個page的數據，並實現翻頁功能

環境：Freemarker任務：顯示一個page的數據，並實現翻頁功能翻頁效果：說明：1、翻頁效果圖中總頁數只有1...

Selenium實現全自動打字

Selenium實現全自動打字任務使用selenium代替人工打字測試。效果思路模擬用戶登錄首先需要在https://dazi.kuk...

ETL數據採集方法

1、觸發器方式觸發器方式是普遍採取的一種增量抽取機制。該方式是根據抽取要求，在要被抽取的源表上建立插入、修改、刪...

使用Selenium實現微博爬蟲：預登錄、展開全文、翻頁

前言在CSDN發的第一篇文章，時隔兩年，終於實現了爬微博的自由！本文可以解決微博預登錄、識別"展開全文"...

數據採集--實現串口數據的上傳

寫在最前面華清遠見教育集團15年專注高端IT培訓做良心教育，做專業教育，做受人尊敬的職業教育創客學院官網...

Selenium3自動化測試【5】安裝Selenium

在線安裝(pip) 以管理員身份運行cmd，然後在cmd命令窗口中輸入【pip install Selenium】進行安裝。對於想安裝...

實現翻頁效果

實現翻頁效果 簡單翻頁://// RootViewController.m//// Copyright (c) 2014年YX All rights reserved.//import "RootVi...

7. 翻頁的實現

1. pylon的翻頁實現首先需要smarty模板引擎的支持：（不會Smarty的請點這裡）首先在src/apps/fnt/assembly.php系統裝配器中...

利用Selenium實現網站自動簽到

大家好，我是辰哥~今天給大家分享一下selenium自動簽到的小技巧。前情提要辰哥收藏了一些論壇網站，經常需...

python數據採集5-存儲數據

Google

已翻譯為 中文(繁體)

顯示原文

選項

51CTO 博客

代碼統治世界

首頁

關注

熱榜

訂閱專欄

大家都在搜索...

寫文章

51CTO 博客

技術成就夢想

Copyright © 2005-2022 51CTO.COM 版權所有京ICP證060544號

友情鏈接

51CTO鴻蒙社區

51CTO學堂

51CTO

關於我們

官方博客

意見反饋

了解我們

熱門標籤

在線客服

網站地圖

全

