

# Selenium robot framework 自動化爬蟲

專題編號：110-CSIE-S015

執行期限：109 年第 1 學期至 110 年第 1 學期

指導教授：陳碩漢

專題參與人員：107590018 韓宗穎

107590031 康紘郡

107590038 鄒承軒

## 一、摘要

本專題主要使用 Robot framework 並配合 Selenium 下尋找網頁中元素的函式與 RED 編輯器實作本專題，此外，在 keyword 命名方面也需遵循 upper camel style 的命名方式以便閱讀。

**關鍵詞:** Robot framework、Selenium、Robot Test、python

## 二、緣由與目的

現在科技日新月異，電商平台逐漸崛起，人們也逐漸習慣在網路上瀏覽商品以及購物，所以電商也順應潮流，會在網路平台上給予優惠來吸引消費者。但是降價以及優惠時段常常選在午夜，與正常人的生活習慣不太相符，所以我們希望能夠創造一個工具，能夠自動在降價的時候通知使用者。

而我們使用 Python 裡的 robot framework-seleniumlibrary 套件和 RED 的開發環境，模擬使用者操作網頁的行為，再配合 xpath 定位網頁上的元素，將我們想要獲得的產品資訊利用類似爬蟲的技術記錄下來。

## 三、研究範圍

在實作中，我們選擇使用 PChome 以及 MOMO 購物網，來作為我們 Demo 成果的範例，因此我們必須以 python 為底層，並利用 selenium 函式庫控制網頁，再搭配 Xpath 來將網站的任意一個 Element 做精準定位，並且設計 test step 來將我們的腳本撰寫精準。

Robot framework 提供使用者基本的 function，使得一些重複的行為，能夠透過高階語言來進行撰寫，而在一些額外的功能上，因為 robot framework 也是基於 python 來撰寫的，所以我們也可以使用 python 來擴充 robot framework 缺少的功能，能夠協助使用者輕鬆地使用簡單 keyword 來完成自動化測試。

## 四、使用技術方法

1-1. Selenium 是一個瀏覽器自動化的套件，可以利用 Python 撰寫自動化的腳本來執行各種的網站，包含開啟瀏覽器、填寫表單、點擊按鈕及取得網站內容等，多數用來執行網站功能的自動化測試，簡化繁瑣及耗時的網站測試工作。

1-2. Keyword 是基於 selenium library 中撰寫好的 function 來組合成更高階的語言，通常是為了增加可讀性。下表是一些在控制網頁時常使用的 keyword。

2.

Keyword	功能
Get Text	獲取網頁文字
Click Element	點擊網頁元素
Wait Until Element Is Visible	等待網頁元素出先現
Scroll Element Into View	捲動網頁到目標元素出現
Input Text	輸入文字
Open Browser	開起瀏覽器

3. XPath(XML Path Language) 是一個使用類似檔案路徑的語法，來定位 XML 文

件中特定節點的語言，因為能夠有效的尋找節點位置，所以也被廣泛的使用在 Python 網頁爬蟲的元素定位上。

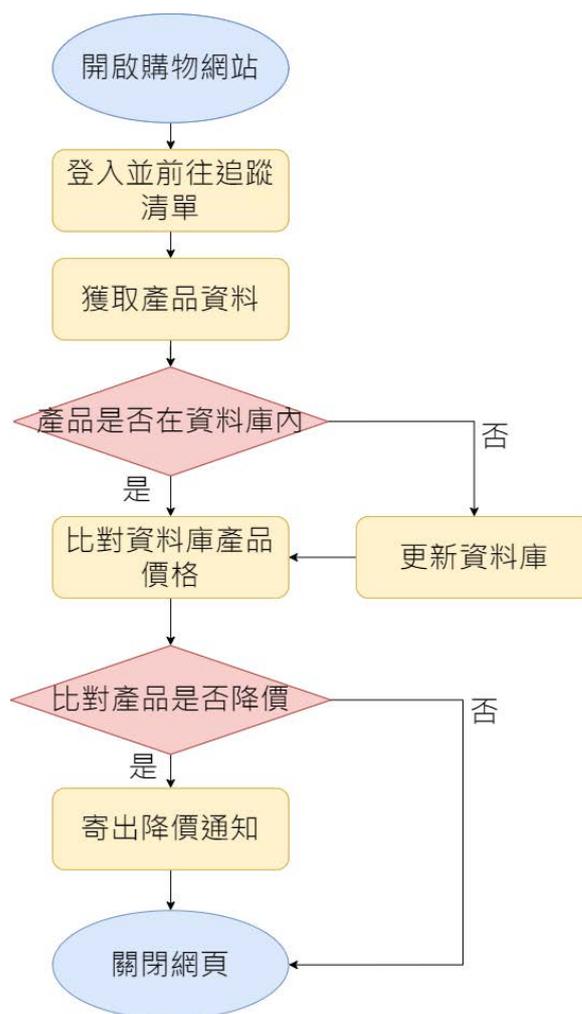
語法	用途
//*	選取文檔中所有元素
//<標籤>[@*]	取得所有含<標籤>屬性的元素。
child::*	取得下一層的所有子元素
descendant::*	取得所有下一層的元素
Ancestor::*	取得所有上一層的元素
preceding-sibling::*	選取當前節點之前的所有同級節點
following-sibling::*	選取當前節點之後的所有同級節點

4. WebDriver 是用來控制 Web 瀏覽器的行為，每一個瀏覽器都會有各自相對應的驅動程式(driver)，而 Selenium 透過 driver 作為後盾，使用 driver 所支援的瀏覽器進一步達到瀏覽器自動化。讓操作就像是真的使用者在操作一樣。

## 五、開發工具

1. Robot framework：劇本實做部分。透過程式碼模擬人工測試，達自動化測試。
2. Python:無法由 Selenium 原生 Keyword 完成，以 python 輔助。
3. RED, VScode: Integrated development environment

## 六、程式架構及流程



## 七、結論

藉由本次專題，能讓我們了解業界是如何來進行一件龐大的專案，也學習到有關網頁測試相關的技術及所使用的語言和工具，例如:利用 selenium來進行網頁控制並且用 Xpath對網頁的物件定位。  
**參考文獻**

- [1]<https://robotframework.org/SeleniumLibrary/SeleniumLibrary.html>
- [2]<https://chentsungyu.github.io/2019/09/03/Python/%5BPython%5D%20%E7%88%AC%E8%9F%B2%E7%AD%86%E8%A8%983-%20Selenium/>
- [3]<http://www.w3big.com/zh-TW/xpath/xpath-syntax.html>