



Selenium

By ThomasRed

What is Selenium

- ▶ *Selenium automates browsers.* That's it! What you do with that power is entirely up to you.



Selenium is a suite of tools to automate web browsers across many platforms.

什麼是駭客：自動化所有超過90秒的工作

- ▶ 這位老兄離開後 Alex 的前同事在檢查他的程式時驚訝地發現，原來這個傢伙幾乎從不自己幹活—只要任何工作需要佔用他 90 秒以上的時間，他就會編寫script，讓程式自動執行完。



- ▶ 向老婆大人彙報
- ▶ 自動修復客戶資料庫
- ▶ 編不上班藉口
- ▶ 自動煮咖啡

Python

- ▶ Python下載：

- ▶ <https://www.python.org/downloads/>

- ▶ 安裝方法：

- ▶ <https://www.foolegg.com/how-to-build-a-python-programming-environment-on-windows/>

PIP

- ▶ 套件管理工具
- ▶ 安裝
 - ▶ <https://pypi.python.org/pypi/pip>
 - ▶ 目前python好像都有預設安裝完成了
- ▶ **DRY (Don't Repeat Yourself)**



確認安裝完成

- ▶ python -V
- ▶ pip -V

```
C:\WINDOWS\system32\cmd.exe

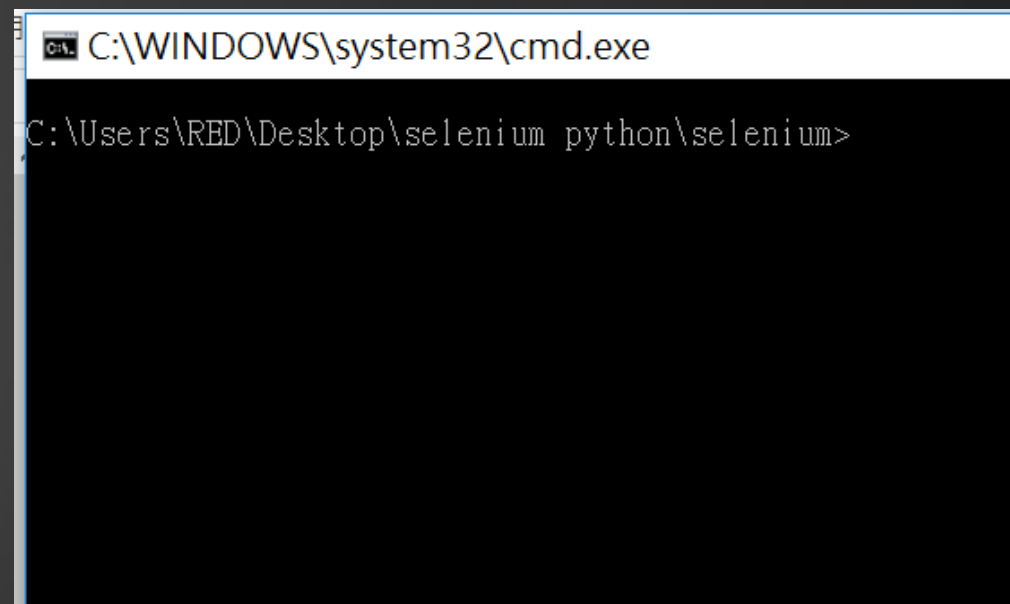
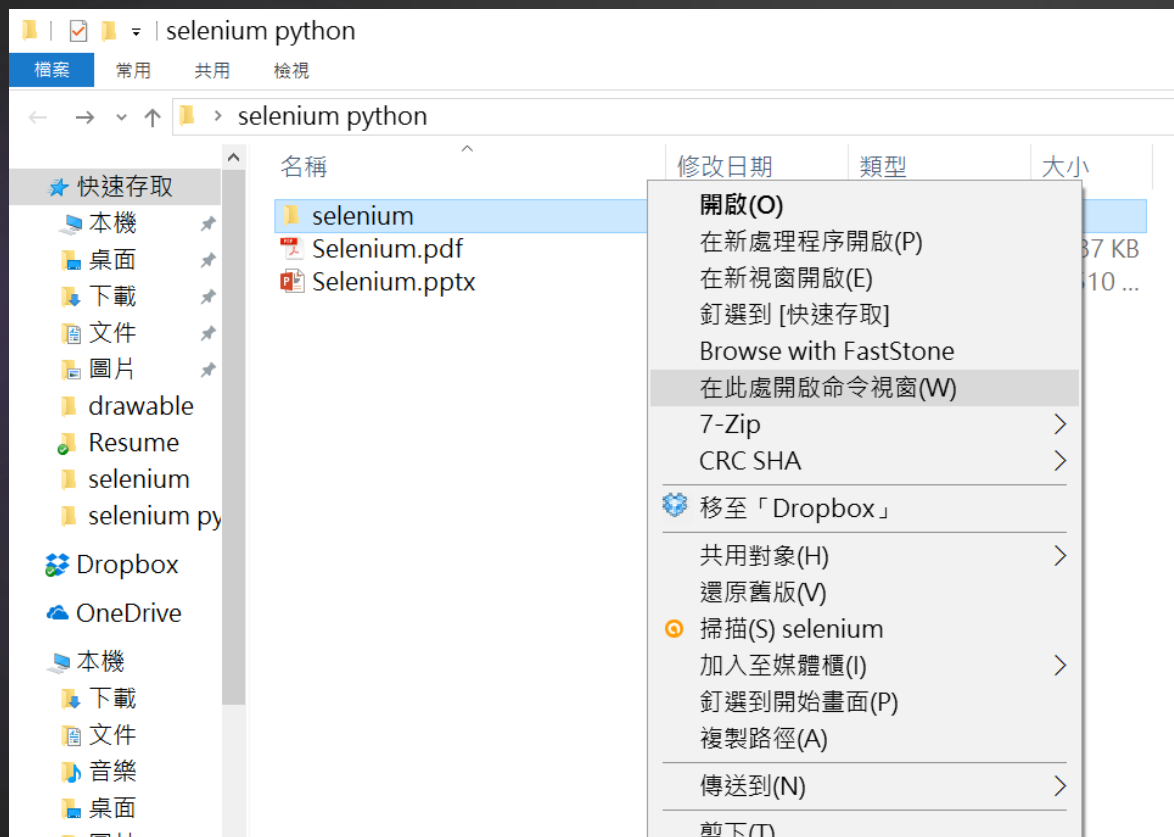
C:\Users\RED\Desktop\selenium python\selenium>python -V
Python 2.7.13

C:\Users\RED\Desktop\selenium python\selenium>pip -V
pip 9.0.1 from c:\python27\lib\site-packages (python 2.7)

C:\Users\RED\Desktop\selenium python\selenium>
```

友情小提示

- ▶ 在windows中對目標資料夾按 shift + 右鍵，可以在直接在目標資料夾中開啟CMD



Installing Selenium

- ▶ pip install selenium
- ▶ 以下是已經安裝過顯示的訊息

C:\WINDOWS\system32\cmd.exe

```
C:\Users\RED\Desktop\selenium python\selenium>pip install selenium  
Requirement already satisfied: selenium in c:\python27\lib\site-packages  
C:\Users\RED\Desktop\selenium python\selenium>
```



Chrome WebDriver

- ▶ 下載 Chrome WebDriver 並解壓縮
 - ▶ <https://sites.google.com/a/chromium.org/chromedriver/downloads>
- ▶ 建立一個資料夾，將第2步下載的 Chrome WebDriver 放進去
 - ▶ Your_Directory/chromedriver
- ▶ 程式碼當中記得要import

```
from selenium import webdriver
```



Locating Elements

- ▶ `find_element_by_id`
- ▶ `find_element_by_name`
- ▶ `find_element_by_xpath`
- ▶ `find_element_by_link_text`
- ▶ `find_element_by_partial_link_text`
- ▶ `find_element_by_tag_name`
- ▶ `find_element_by_class_name`
- ▶ `find_element_by_css_selector`
- ▶ `find_elements_by_name`
- ▶ `find_elements_by_xpath`
- ▶ `find_elements_by_link_text`
- ▶ `find_elements_by_partial_link_text`
- ▶ `find_elements_by_tag_name`
- ▶ `find_elements_by_class_name`
- ▶ `find_elements_by_css_selector`

```
import By
driver.find_element(By.XPATH, '//button[text()="Some text"]')
driver.find_elements(By.XPATH, '//button')
```



XML Path Language

- ▶ XPath 使用路徑表達式來選取文檔中的node
 - ▶ node-name
 - ▶ /
 - ▶ //
 - ▶ .
 - ▶ ..
 - ▶ @

Xpath



Waits

- ▶ Selenium 提供的等待

```
try:  
    element = WebDriverWait(driver, 10).until(  
        EC.presence_of_element_located((By.ID, "myDynamicElement"))  
    )  
finally:  
    driver.quit()
```

- ▶ Time sleep

```
time.sleep(random.uniform(1, 3))
```



擷取畫面

- ▶ Selenium Web Driver 提供這個函數 `save_screenshot`，只要給予檔名，就可以針對當下的畫面擷取，並且儲存。預設會儲存在與該執行 `python script` 相同路徑下。

```
driver.save_screenshot(file_name)
```

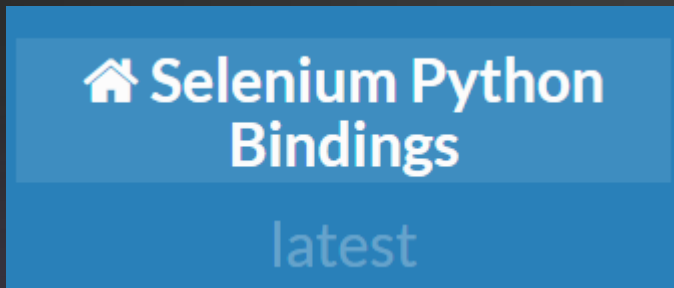


WebDriver

- ▶ install pyvirtualdisplay
 - ▶ (a wrapper for Xvfb)
- ▶ PhantomJs GhostDriver
 - ▶ <http://phantomjs.org/download.html>



參考



▶ <http://selenium-python.readthedocs.org>

▶ <http://chimerhapsody.blogspot.tw/2013/06/python-selenium-chrome.html>