# 反爬:登入授權模擬

- 了解「登入權限機制」的反爬蟲機制
- 「登入權限機制」反爬蟲的因應策略

## 作業目標

• 找一個需要登入的網站試試看,並說明思考流程(如果不知道要用哪個網站的話,可以試試看 https://github.com/new (https://github.com/new)網址,未登入時會被導向登入頁)

#### In [13]:

```
import requests
from bs4 import BeautifulSoup

r = requests.get('https://github.com/new')
soup = BeautifulSoup(r.text, 'lxml')
soup.title # 當我們直接存取 https://github.com/new 網址時,網頁會被導到登入頁
```

#### Out[13]:

<title>Sign in to GitHub · GitHub</title>

#### In [2]:

```
import requests
from bs4 import BeautifulSoup

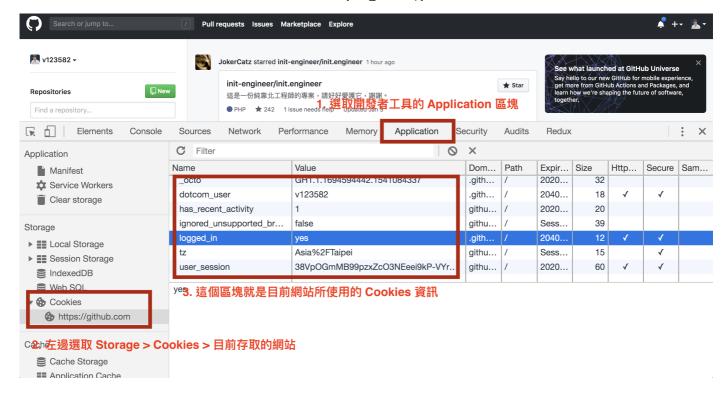
cookies = {
    'logged_in': 'yes',
    'user_session': '9-S3QlTWrtHybLYX0gYeGmdqlMdM_rTnfi3jbRLgdvYKJWoQ'
}

r = requests.get('https://github.com/new', cookies=cookies)
soup = BeautifulSoup(r.text, 'lxml')
soup.title # 加上檢查的 cookies ,就可以正確取得資料
```

#### Out[2]:

<title>Create a New Repository</title>

### 補充:從瀏覽器找 cookie 的方法



#### In [ ]: