

LAPORAN HASIL  
PERCOBAAN *SCRAP*  
WEBSITE SHOPEE.TW

## A. Langkah yang Dilakukan

### 1. Melakukan *Scrap* Secara Langsung

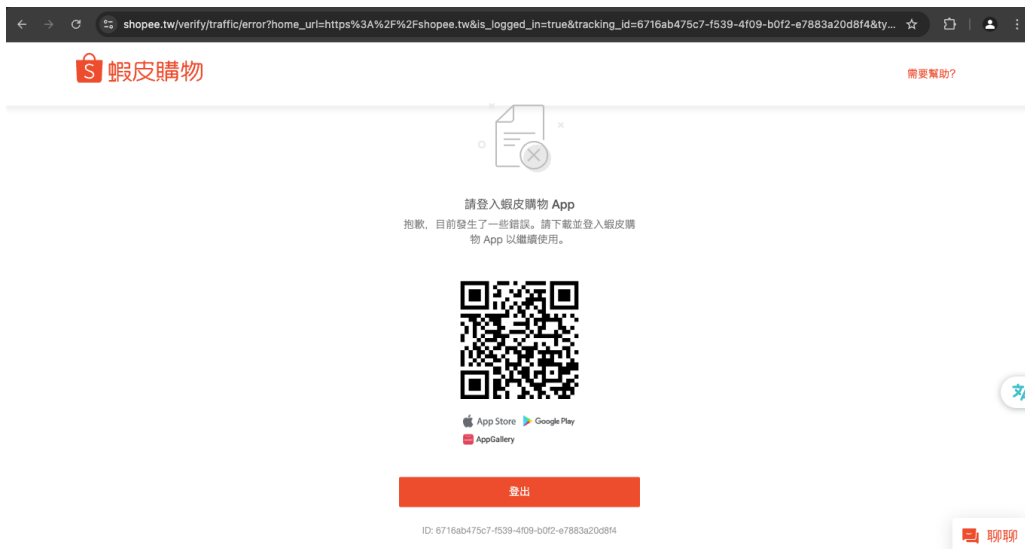
Pertama kali yang saya lakukan saat setelah memahami *challenge* yang didapat adalah mencoba untuk mengakses laman <https://shopee.tw> secara langsung, pada tahap ini saya belum melakukan *scrap*. Tetapi yang terjadi adalah setiap kali mengakses melalui peramban web, akan langsung *redirect* ke halaman <https://shopee.tw/buyer/login>.

Saya kemudian mencoba menggunakan kode program yang mana langsung menembak *url* ke salah satu alamat web yang ada pada *challenge*. Akan tetapi, hal yang terjadi adalah tidak bisa, karena ketika menggunakan webdriver, maka akan diarahkan ke halaman *login*. Disini, saya mencoba juga apakah apabila yang diakses merupakan *website* Shopee Indonesia akan berbeda, hal yang terjadi adalah sama yaitu apabila tidak melakukan *login*, maka akan *redirect* ke halaman *login*.

Pada tahap ini, saya masih belum menggunakan Python, tapi masih menggunakan Express. Setelah saya mendapat beberapa referensi, beberapa *website* memberi contoh untuk menggunakan Python.

### 2. Melakukan *Scrap* Menggunakan VPN & Cloudflare WARP

Ketika langkah pengaksesan secara langsung menemui jalan buntu, akhirnya saya membuat akun untuk [shopee.tw](https://shopee.tw) dengan tujuan agar dapat mengakses setidaknya halaman utama dari *website* tersebut. Langkah ini berhasil, saat saya bisa mengakses halaman utama dari *website* [shopee.tw](https://shopee.tw) termasuk tautan produk yang tertera pada *challenge*. Kondisi saya di titik ini adalah, saya belum mencoba untuk melakukan *scrap*, saya masih penasaran dan ingin tahu bahwa halaman utama dari [shopee.tw](https://shopee.tw) dapat diakses.



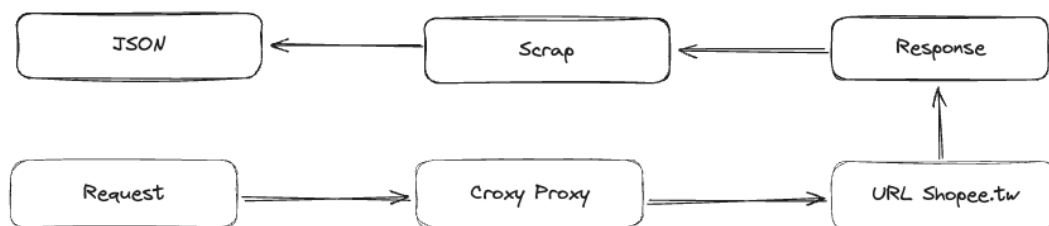
Namun, hal tersebut hanya terjadi mungkin sekitar 5-10 menit, sebelum kemudian ada *error* dimana jika diterjemahkan dari Bahasa Mandarin, saya diharuskan untuk mengunduh aplikasi Shopee. Saya kemudian terpikir, apakah jika menggunakan Cloudflare WARP atau VPN server Taiwan maupun global akan dapat masuk kembali mengakses halaman utama Shopee. Beberapa *provider* yang saya coba antara lain Urban VPN dan Free VPN for Chrome - VPN Proxy VeePN. Akan tetapi masalah yang saya alami sama, tidak dapat mengakses halaman utama dari Shopee. Setelah itu, saya mencoba melakukan *scrap* tapi dengan tetap menyalakan VPN dan Cloudflare WARP. Namun, hal yang saya dapatkan adalah sama yaitu data tidak dapat dilakukan *scrap*.

### 3. Melakukan *Scrap* Menggunakan Proxy

Setelah dua cara pertama gagal saya lakukan, saya kemudian mencoba menerapkan tips dari Pak Cahyo yaitu dengan menggunakan *proxy*. Cara saya tetap seperti sebelumnya, yaitu pengaksesan langsung menggunakan peramban web. Hasil yang saya dapatkan adalah halaman utama dapat diakses jika menggunakan croxy proxy meskipun proses memuat halamannya lama.



Setelah dapat mengakses halaman utama dari halaman utama shopee.tw, saya kemudian mencoba menduplikasi mekanisme dari croxy proxy pada kode program *scrap* dengan skema seperti pada gambar berikut.



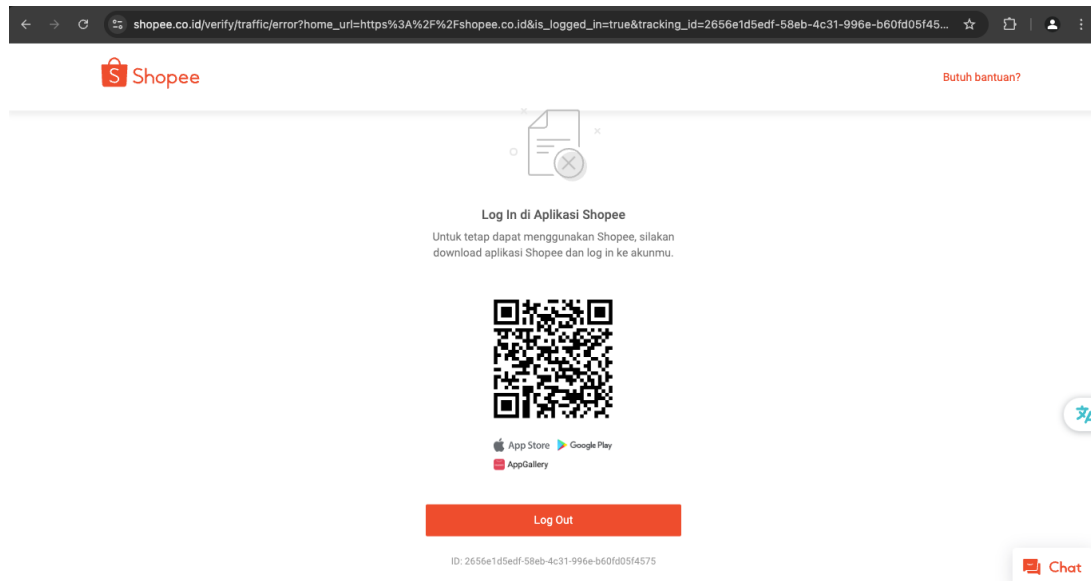
Akan tetapi, hal yang saya dapat yaitu sama, masih belum dapat mengambil data dari *website* shopee alias proses *scrap* yang saya lakukan gagal. Saya juga menggunakan beberapa *provider proxy* yang tentu saja saya mencari *free trial* seperti Smartproxy, Webshare, dan tentu saja Croxy Proxy. Namun, proses *scrap* masih belum dapat saya lakukan.

#### 4. Melakukan *Scrap* Menggunakan Cookies

Cara terakhir yang saya coba adalah menggunakan *cookies*, cara ini sebenarnya cukup ampuh untuk berhasil mengakses halaman detail produk shopee. Jadi saya memanfaatkan *extension* editthiscookies yang dapat mengeksport *cookies* login kita dari suatu *website* dan disimpan dalam bentuk JSON.

Pada tahap ini, beberapa kali saya bisa mengakses halaman detail produk, tapi saya masih belum berhasil mengambil data dari *website* untuk proses *scrap*. Hal yang tidak terduga terjadi, yaitu muncul pesan error agar melakukan proses mengunduh aplikasi Shopee. Alhasil saya masih belum bisa melakukan proses *scrap* pada *website*

Shopee. Pada kasus ini, saya menggunakan alamat Shopee Indonesia untuk mencoba memahami paradigma apakah ada perbedaan antara Shopee Taiwan dan Shopee Indonesia.



## B. Kendala yang Dihadapi

1. Tidak dapat mengakses halaman utama shopee.tw secara penuh dan nyaman. Meskipun beberapa kali berhasil mengakses halaman utama shopee.tw, tapi durasi pengaksesannya sangat singkat jika tidak menggunakan *proxy*.
2. Akun Shopee saya sempat terkena *restricted* karena diduga memiliki aktivitas yang tidak wajar, jika saya membaca pedoman Shopee dari *website* resminya. Apabila ada akun yang terbukti terlibat dalam aktivitas *scraping* akan dibatasi.
3. Mekanisme penggunaan *proxy* yang masih *tricky* bagi saya terutama bagaimana cara untuk melakukan sinkronisasi antara *request* kode program agar diterima oleh *proxy* yang kemudian akan meneruskan permintaan dari kode program yang dijalankan.
4. Sepertinya, shopee.tw menerapkan *region lock* sehingga hanya *user* dari wilayah sekitar Taiwan saja yang dapat mengakses secara langsung dan penuh.

## C. Hipotesis Mengenai Pemecahan Masalah Terkait *Scrap* Shopee

1. Sepertinya, penggunaan *residential proxy* seperti yang disarankan Pak Cahyo bisa menjadi opsi. Awalnya saya ingin mengakali dengan menggunakan WARP, VPN, maupun Proxy biasa karena biasanya dengan menggunakan salah satu dari 3 *tools* tersebut, *website* yang tidak dapat diakses secara konvensional jadi dapat diakses.
2. Tapi saya masih ada pertanyaan, jika dengan menggunakan *residential proxy* ternyata bisa mengakses *website* shopee.tw dan tidak diarahkan untuk mengunduh aplikasi

*mobile* Shopee. Harusnya ini juga berlaku untuk Shopee Indonesia, yang mana apabila diakses terus menerus, tidak akan ada permintaan untuk mengunduh aplikasi Shopee.