



## Understand Your Users

### Easy Setup, Quick Results

Treejack Helps You Design Your Information Architecture With Data. Try For Free Today.

[optimalworkshop.com](https://optimalworkshop.com)

OPEN

# How to avoid Selenium webdriver from being detected as bot or web spider

26

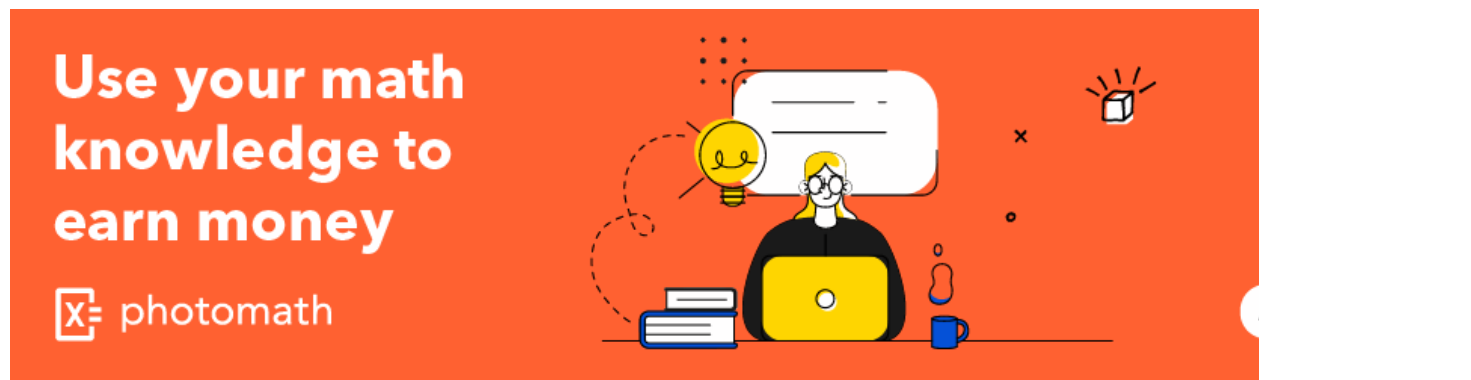
JAN

26 JANUARY 2021 | WRITTEN BY [CHIN-HOCK TAN](#) [PRINT](#) [EMAIL](#)



Before we start to use [php-webdriver and Selenium](#) for web scraping and social media auto posting, we need to do some settings in code or file modifications to avoid our script from being detected as web bot or spider. I have listed some ways to hide our automation using Selenium. The methods can be used for any programming languages as well. Please note that this is not a complete list and from time to time web servers companies can find new methods to detect and block our Selenium automation. Anyway, we just have to factor in all known methods in our scripts to reduce chances of detection.

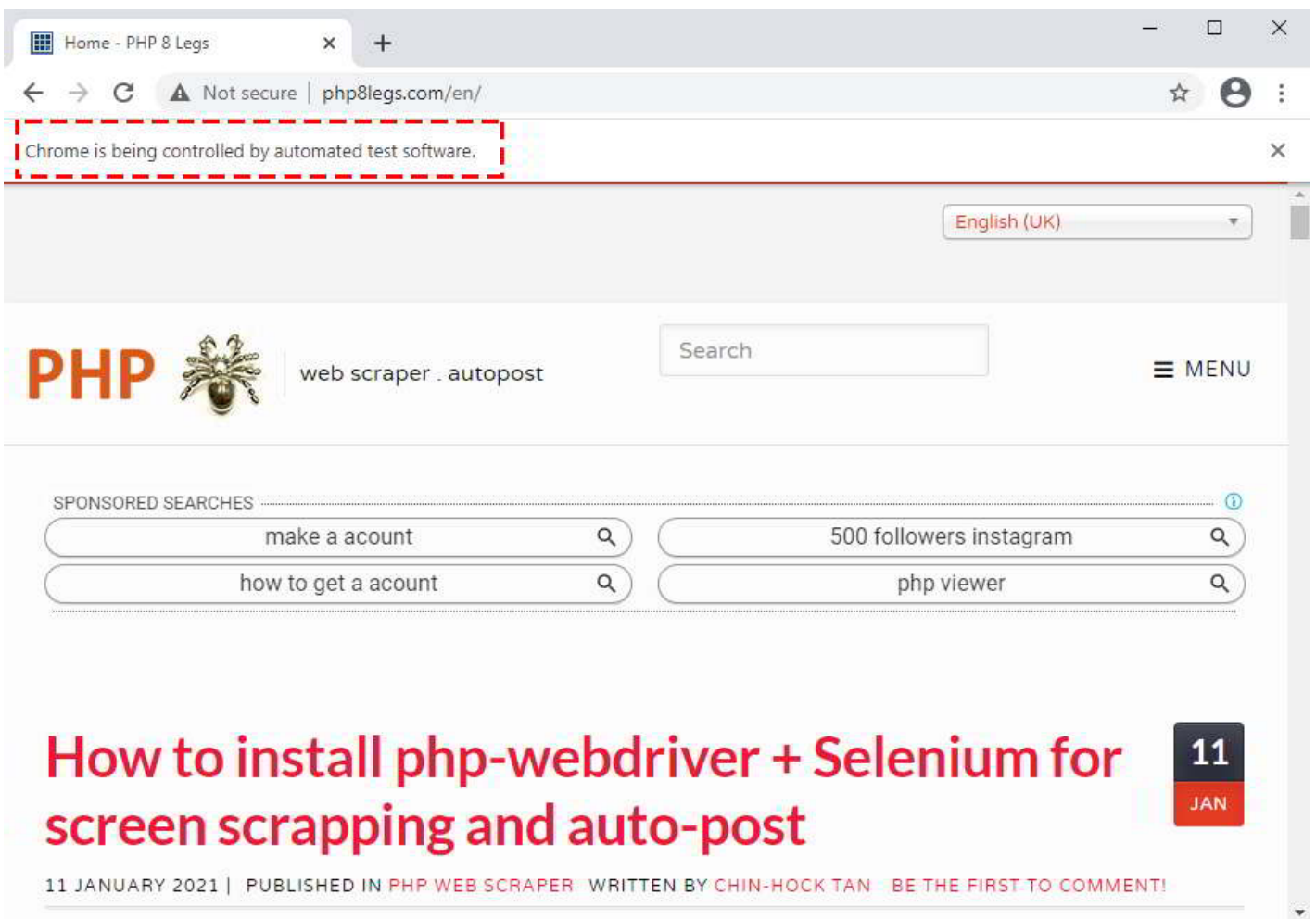
1. Remove browser control flag
2. Remove signature in javascript
3. Set User-Agent
4. Avoid using headless browser
5. Use maximum resolution
6. Follow page flow
7. Use proxy or VPN
8. Insert random delay
9. Use cookies to login



Previous article : [How to install php-webdriver + Selenium for screen scrapping and auto-post](#)

### 1. Remove browser control flag

If you run Selenium with default settings, you will see a line of notification "Chrome is being controlled by automated test software" at the top of Chrome browser. I am not sure can web server see this notification or detect the flag that turn this on. But since it is there in front of screen, we can turn the flag off.

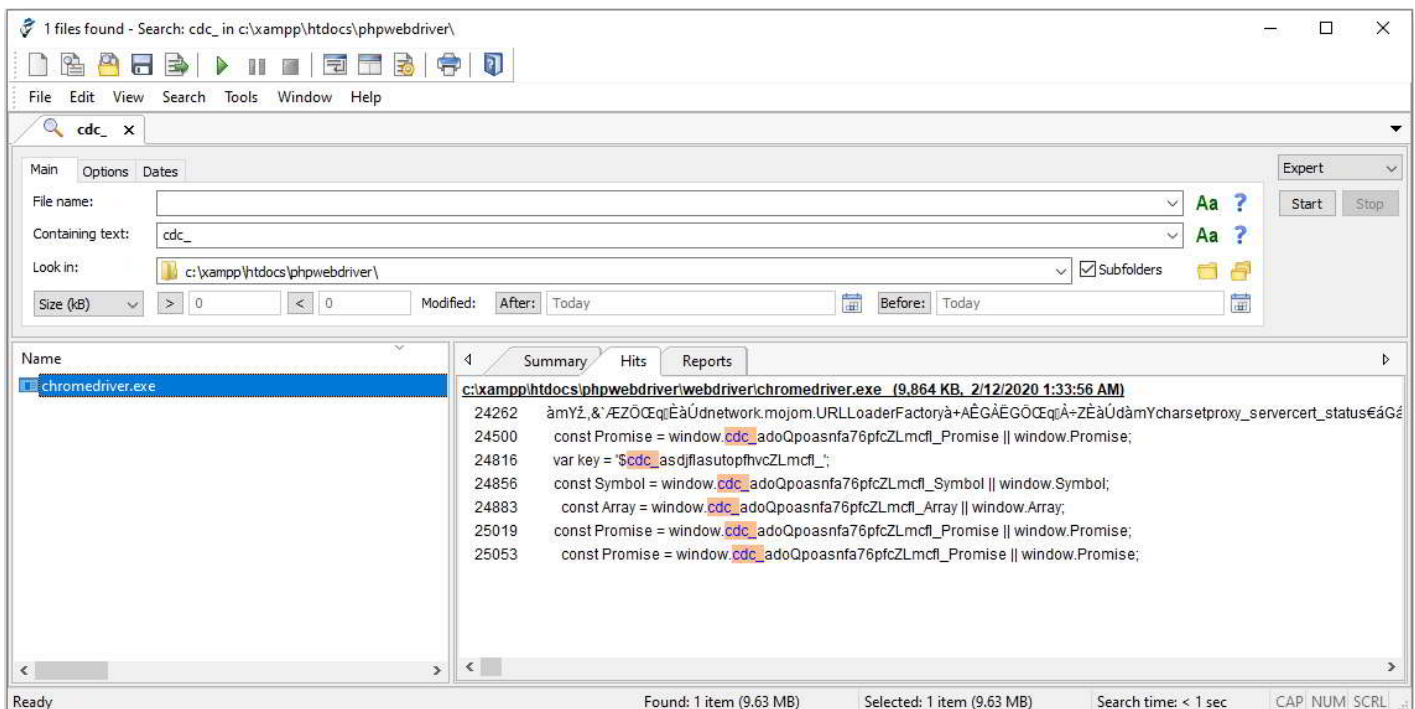


In php-webdriver, use `setExperimentalOption` to disable automation flag in `ChromeOptions` object. You will not see the notification again after this.

```
1 | $ops = new ChromeOptions();  
2 | $ops->setExperimentalOption("excludeSwitches", array("enable-automation"));  
3 | $capabilities = DesiredCapabilities::chrome();  
4 | $capabilities->setCapability( ChromeOptions::CAPABILITY, $ops ); <br>$driver = |
```

## 2. Remove signature in javascript

Inside `chromedriver.exe` (same for `geckodriver` (firefox) and `edgedriver` (edge)) there is a javascript signature that used by bot detection software such as [FingerprintJS](#), [Imperva](#) or even Google's Captcha. I use [Agent Ransack](#) to search for "cdc\_" signature in `chromedriver.exe` binary file.



The signature is "\$cdc\_asdfilasutopfhvcZLmcfl\_". What we can do is change "cdc" to string of same length. For example, I change "cdc" to "tch". You can change to anything like "abc", "xyz" etc.

Since this is a binary file, we can not edit that signature with normal text editor. I am using "vim" for this purpose.



Go to <https://www.vim.org/download.php> and download self-installing executable file.

vim.org/download.php

SPONSOR

Vim development

VOTE

for features

the editor

BUY

the Vim book

HELP

Uganda

LEARN

Vim

not logged in (login)

ENHANCED BY Google

Search

Home

Advanced search

About Vim

Community

News

Sponsoring

Trivia

Documentation

Downloaded

Vim from GitHub

Vim from Mercurial

List of Mirrors

Sources

Patches

Development

Runtime files

Script links

Translations

Old stuff

Scripts

Tips

My Account

Site Help

Note

Vim 8.1 has been released. Many people are happily using Vim without any problems. Nevertheless, the usual disclaimers apply. We take no responsibility for any trouble that is caused by Vim. For known problems type "help todo" in Vim.

If you find a problem, please report it on the [vim-dev mailing](#). Or send it to bugs AT vim DOT org.

Vim and ICCF Holland supporters: [Collectivizaray.com](#) is happy to help with the ongoing development of the VIM project and their efforts to help and support students in developing countries.

Downloading Vim

Vim is available for many different systems and there are several versions. This page will help you decide what to download.

Most popular:

Click this link to download the [self-installing executable using ftp](#).

Signed MS-Windows files are available on the [vim-vim82-installer site](#) ([gvim\\_8.2.0012\\_x86\\_signed.exe](#) is recommended)

Unix:

See the [GitHub](#) page, or [Mercurial](#), if you prefer that. There is also an [AppImage](#) which is build daily and runs on many Linux systems.

Mac:

See the [MacVim](#) project for a GUI version and [Homebrew](#) for a terminal version

Details and options for:

Unix

PC: MS-DOS and MS-Windows

Amiga

OS/2

Macintosh

Others

Mirrors

Alternative sites to download Vim files from.

Sources

Build Vim yourself and/or make changes.

GitHub

Obtain Vim sources with a git client (**recommended**).

Mercurial

Obtain Vim sources with a Mercurial client (recommended if you don't like git).

Patches

Include the latest improvements (requires sources and rebuilding).

Runtime

Get the latest syntax files, documentation, etc..

Script links

Links to individual syntax, indent, color, compiler and ftplugin scripts.

Translations

Non-English documentation packages.

Versions before 7.3 can also be obtained with [Subversion](#) and [CVS](#).

Version

Vim 8.2 is the latest stable version. It is highly recommended, many bugs have been fixed since previous versions. If you have a problem with it (e.g., when it's too big for your system), you could try version 6.4 or 5.8 instead.

To avoid having to update this page for every new version, there are links to the directories. From there select the files you want to download. In the file names ## stands for the version number. For example, [vim##src.zip](#) with version 8.2 is [vim82src.zip](#) and [vim##src.tar.gz](#) for version 8.2 is [vim-8.2-src.tar.gz](#). Links are provided for quick access to the latest version. Note that the links point to the latest version (currently 8.2) to avoid that caching causes you to get an older version.

Unix

The best way to install Vim on Unix is to use the sources. This requires a compiler and its support files. Compiling Vim isn't difficult at all. You can simply type "make install" when you are happy with the default features. Edit the Makefile in the "src" directory to select specific features.

BANK ON MORE PRIZES

Spent With AmBank Debit Card

Win \$850,000 Cash

10 September 2021 - 31 March 2022

Learn More

Help Uganda

kindle fire

\$199

Buy at Amazon

After installation, go to C:\Program Files (x86)\Vim\vim82> (because the installation not set env path automatically, but you can do it yourself if you want).

Run command "vim.exe <path to>\chromedriver.exe

```

cmd Command Prompt
24/01/2021 04:40 PM <DIR> macros
08/12/2019 02:20 AM 20,445 makemenu.vim
11/12/2019 05:06 AM 40,648 menu.vim
08/12/2018 03:15 AM 3,430 mswin.vim
08/11/2019 04:34 AM 60,895 optwin.vim
24/01/2021 04:40 PM <DIR> pack
24/01/2021 04:40 PM <DIR> plugin
12/12/2019 07:43 PM 4,900 README.txt
20/07/2016 02:44 AM 17,780 rgb.txt
26/06/2019 03:29 AM 11,348 scripts.vim
24/01/2021 04:40 PM <DIR> spell
08/12/2019 02:21 AM 39,770 synmenu.vim
24/01/2021 04:40 PM <DIR> syntax
12/12/2019 08:19 PM 108,544 tee.exe
24/01/2021 04:40 PM <DIR> tools
24/01/2021 04:40 PM <DIR> tutor
24/01/2021 04:40 PM 104,593 uninstall-gui.exe
12/12/2019 08:19 PM 160,256 uninstall.exe
27/09/2019 06:57 PM 3,851 uninstall.txt
12/12/2019 08:19 PM 2,566,656 vim.exe
27/01/2019 03:53 AM 1,333 vimrc_example.vim
12/12/2019 08:30 PM 117,248 vimrun.exe
27/01/2018 05:20 AM 1,812 vimtutor.bat
28/07/2017 11:11 PM 581,632 winpty-agent.exe
28/07/2017 11:11 PM 508,928 winpty32.dll
12/12/2019 08:19 PM 126,976 xxd.exe
37 File(s) 10,254,924 bytes
19 Dir(s) 474,683,981,824 bytes free

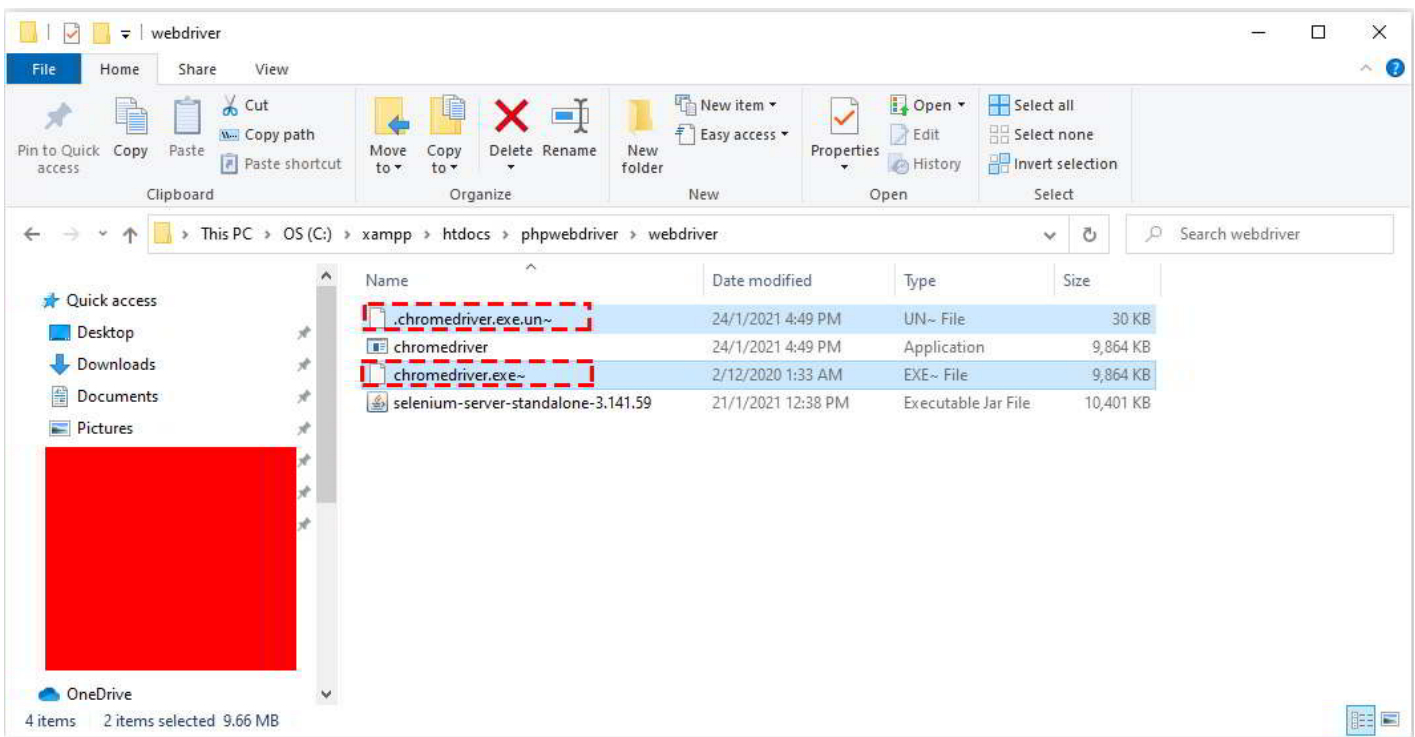
C:\Program Files (x86)\Vim\vim82>vim.exe c:\xampp\htdocs\php8legs\webdriver\chromedriver.exe_

```

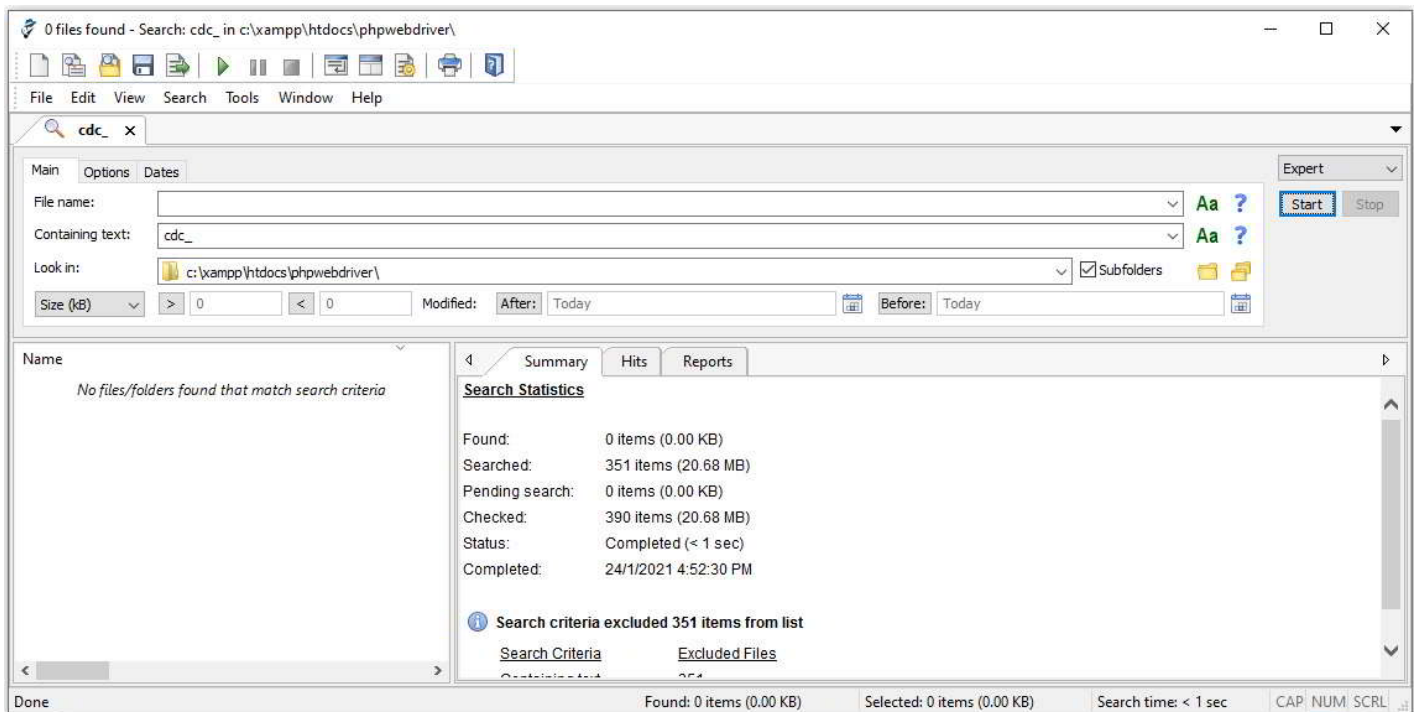
The binary file displayed in vim looks like this. Type ":<%=cdc\_/tch\_/g" to make a global change of string "cdc\_" to "tch\_". Enter to execute the command.







Now if we do the same search, it is gone. You can also verify by searching the new string that you changed. So now the signature has changed!



### 3. Set User-Agent

Social medias most likely keep track of our internet IP address and user-agent when we use browser to create account in social media or updating new post. So it is good to use the same user-agent during Selenium automation.

From the browser that we use for social media activates, go to <https://www.whatismybrowser.com/detect/what-is-my-user-agent>.

← → ↻ [whatismybrowser.com/detect/what-is-my-user-agent](https://whatismybrowser.com/detect/what-is-my-user-agent) ☆

**WhatIsMyBrowser.com** Search Guides Detect my settings Tools Developers

[Homepage](#) > [Detect my settings](#) > [What is my user agent?](#)

## What is my user agent?

Your User Agent is:

**Mozilla/5.0 (Windows NT 10.0; Win64; x64)**

**AppleWebKit/537.36 (KHTML, like Gecko)**

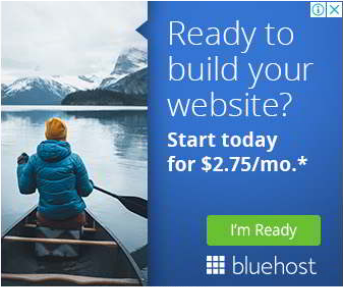
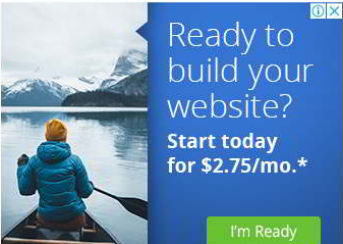
**Chrome/88.0.4324.104 Safari/537.36**

This page shows you what your web browser is sending in the "User-Agent" header for your HTTP requests. Want to know [what headers your browser is sending?](#)

### What does all of that mean?!

Each bit of your user agent indicates something particular about your system.

There's no "standard" way of writing a user agent string, so different web browsers use different formats (some are *wildly* different), and many web browsers cram loads of information into their user agents. That's where WhatIsMyBrowser.com steps in - we decode your user agent string to figure out everything it's saying.

Cut and past the user-agent string and set in Selenium. Example:

```

1 | $chrome_options = array( '--user-agent=Mozilla/5.0 (Windows NT 10.0; Win64; x64);
2 | $ops = new ChromeOptions();
3 | $ops->addArguments( chrome_options );
4 | $capabilities = DesiredCapabilities::chrome();
5 | $capabilities->setCapability( ChromeOptions::CAPABILITY, $ops );
6 | $driver = RemoteWebDriver::create( host, $capabilities );

```

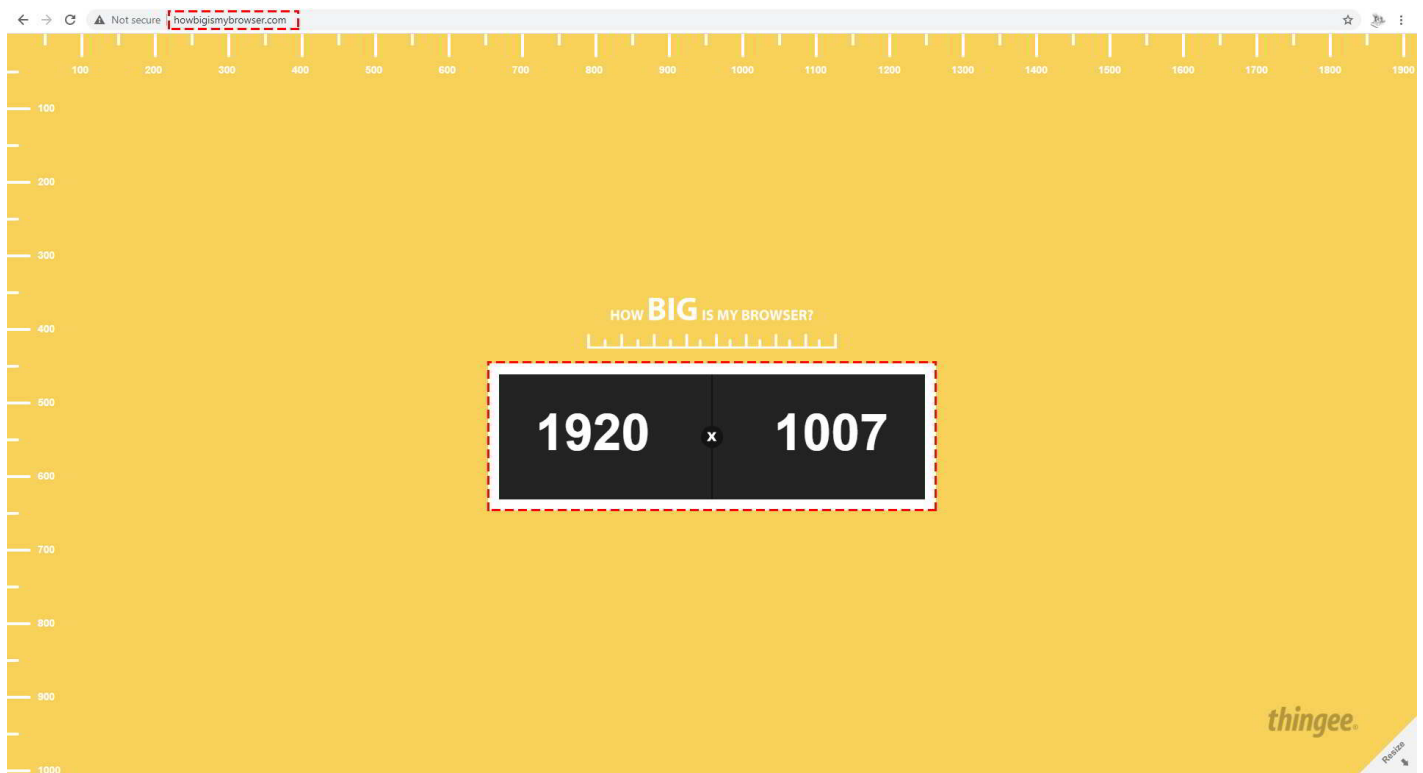
#### 4. Avoid using headless browser

Chrome browser introduced the ability to run in headless mode, that is the ability to run Chrome without creating visible browser window and other benefits such as greater testing reach, improve speed and test performance and multitasking. However, in real world, we perform social media activities through web browser. So it is good to leave the browser window remain opened during autoposting.

#### 5. Use maximum resolution

Since we want to leave the browser window opened, we can set it to a reasonable size. To check your browser window size, go to <http://howbigismybrowser.com/>.





You can set to maximum screen size or to the nearest number of pixel in Selenium.

```
1 | $chrome_options = array( '--start-maximized' ) // set to max screen size
2 | $chrome_options = array( 'window-size=1400,900' ) // set to 1400x900
```

## 6. Follow page flow

Unlike scraping using cURL, it is better to follow the page flow when using webdriver. For example, if you can only go to page C after browsing page A, then B, don't direct go to page C. Try to imitate human user browsing actions.

## 7. Use proxy or VPN

Never use your own IP address for scraping or auto posting. Website like Amazon will block your IP address if the server detect unusual activities. Always use VPN or proxy. Even that, do not use the same VPN or proxy to continuously scraping the same website. Always change to new VPN address or proxy to avoid detection.

## 8. Insert random delay

Always insert random delay between two actions. For example, after login, insert random delay of 5 to 10 seconds before go to next page. So server will see different delay time from one action to another for every cycles.

```
1 | sleep( rand ( 5, 10 ) );
```

## 9. Use cookies to login

This is important for social media posting. Store the cookies file after login with username/password for the first time. Then use cookies for subsequence logins. Login with username/password and without cookies too frequently in a day might get your account blocked.

In the next article, I will write about how to use php-webdriver + Selenium to auto post to Pinterest.

# Use your math knowledge to earn money

 photomath



Last modified on Tuesday, 16 February 2021 19:31

Rate this item

(4 votes)

Read **2483** times

Published in [PHP Web Scraper](#)

Tagged under

Selenium

WebDriver

Chromedriver

phpwebdriver

bot

spider

[Tweet](#)

[Like](#) [Sign Up](#) to see what your friends like.

## Related items

[Pinterest bot - How to auto-login with php-webdriver + Selenium](#)

[How to install php-webdriver + Selenium for screen scrapping and auto-post](#)

More in this category: [« How to install php-webdriver + Selenium for screen scrapping and auto-post](#)

[back to top](#)



## Tags

auto login

auto update

Chromedriver

cURL

Email Extractor

Email Harvesting

Facebook

Graph API v22

Harvesting Emails

instagram

Installation

invisibility

Joomla

login form

MySQL

PHP

PHP SDK V4

phpwebdriver

PHP web scrapping

Price Comparison Engine

proxy

Regular Expression

screen scrapping

Selenium

Spam

Sphinx Search

spider

user agent

WebDriver

XAMPP



Get SiteGround hosting at low price and fast server. Score high on PageSpeed test like our website!



## Comments

---

selvam

ON SATURDAY, 12 JUNE 2021 12:19

Hi team , I am looking to scrap a website via VPN using Curl with PHP (php version 5.6). Can...

[HTTP Get Request via PHP/cURL To Request Web Page Source File](#)

hema

ON THURSDAY, 01 OCTOBER 2020 08:58

how can i register

[How to auto-follow, auto-comment and auto-like a post on Instagram Page with PHP/cURL \(2\)](#)

---

[Subscribe to this RSS feed](#)

## Most Viewed Articles

---

[The Art of Invisibility for Ninja Web Spider](#)

03 January 2014 Read 46021 times

[Create MySQL Database for PHP Web Spider Extracted Emails Addresses \(4\)](#)

19 November 2013 Read 42398 times

[Facebook Remote Status Update with PHP/cURL Bot](#)

04 February 2014 Read 37594 times

[Subscribe to this RSS feed](#)

## Get IPVanish VPN Today!!

---



[Home](#)   [About Me](#)   [Privacy](#)   [Terms of Service](#)

Copyright © 2013-2021 php8legs.com. All Rights Reserved. Framework by php8legs.com.



















