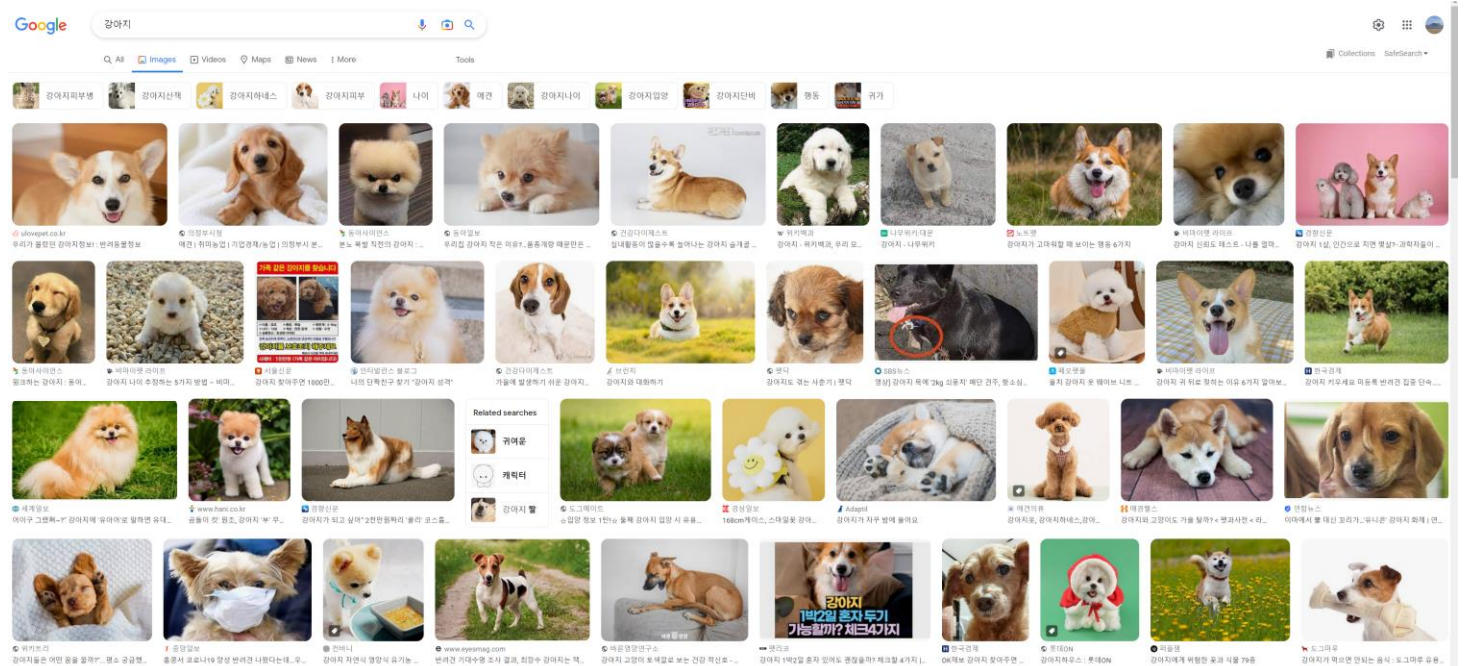


Web crawling with Selenium

Dynamic Rendering of Web pages



Authentification

NAVER 카페홈 이웃 가입카페 새글 586 내소식 7 채팅 6 삼미슈퍼스..

대한민국 1등 중고거래 **중고나라** 1900만 회원 달성

(주)중고나라는 통신판매를 허가받지 않은 판매 혹은 구매 회원 간의 상을 거래 정보 및 거래에 관여하지 않으며 어떠한 의무와 책임도 부담하지 않습니다.

중고나라 위키 · 농치치마!중나앱 꿀팁 · 이벤트 · 공지사항 · 이용정책 · 꿀딜특구

★ 카페정보 나의활동

매니저 **중나운영자**
since 2003.12.10.
카페소개

숲 19,031,403 초대하기
★ 즐겨찾는 멤버 3,514,465명
게시판 구독수 2,514,599회
우리카페앱 수 153,695회

주제 생활 > 중고용품

NAVER 대표카페 2021년 대표 인기카페

카페 글쓰기

카페 채팅

검색

★ 즐겨찾는 게시판

전체글보기 183,115,013

★중고나라 앱 바로가기★
★카페>앱 실시간 연동신청

새 것 같은 중고폰 최상급 폰도 중고나라에서!

IT기기 정보가 가득! 게시판 구경하러 가기!

내 폰 팔기 오토바이 팔기 반값요금제 자전거 팔기 보험가입 중고폰 특가

Best 중고나라 Weekly Pick

SAVE 67% Galaxy Tab S4 +무선키보드+S펜 폴셋
정상판매가 ₩899,000
298,000원

kt skylife TV+인터넷+WiFi 평생 **30%할인**
다이렉트상당 월 19,800원

마켓팅/광고/제휴 문의

넘 새 러 정 없 이 블랙 프라이데이 프로 슬림 9,900원
+ 특별 경품 증정! 구매하기 >

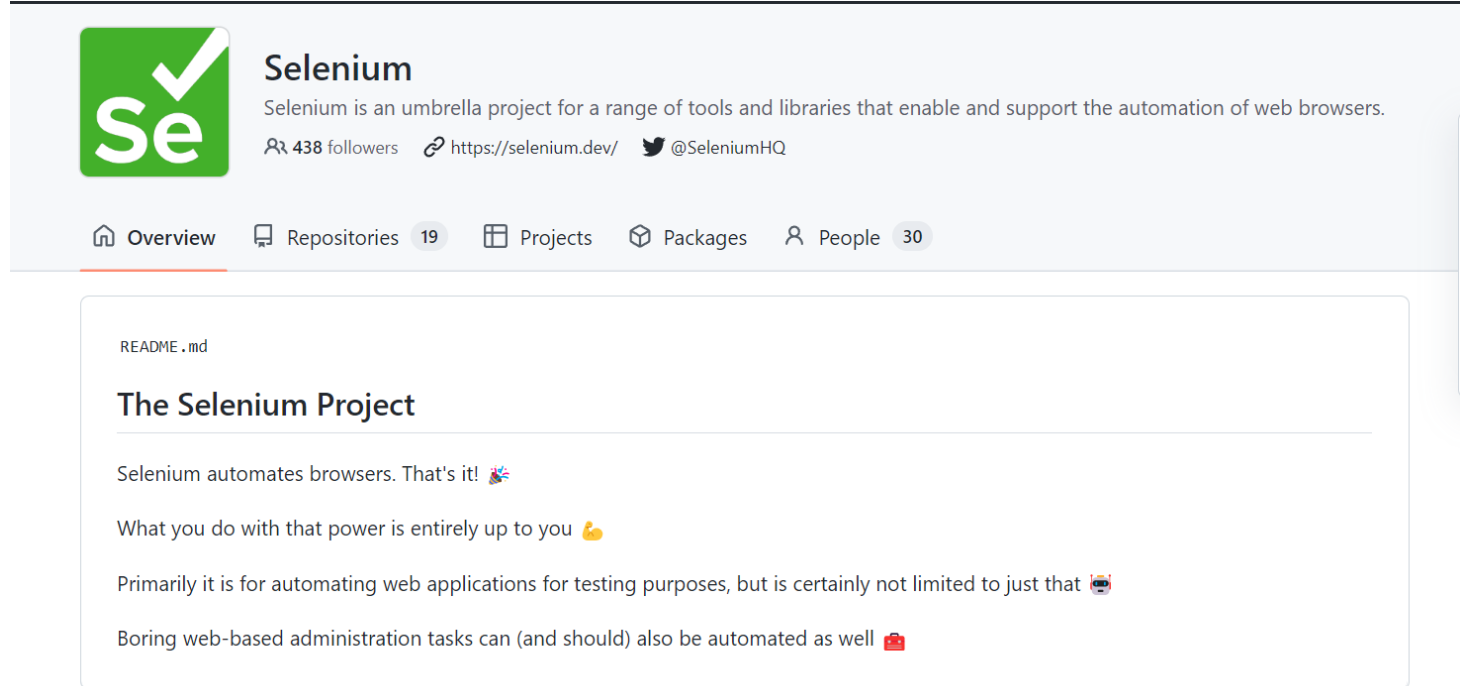
Best 중고나라 추천

단, 3일 사용! 신품급! 갤럭시 워치5 & 버즈2 프로 사전예약 **169,000원**

올해 이사한 서울 청년 최대 **40만원 드림**
부동산 중개수수료부터 이사비까지 청년몰방정보통 4만 원 상당

Selenium

- Selenium is an open source umbrella project for a range of tools and libraries aimed at supporting browser automation.



The screenshot shows the Selenium GitHub repository page. At the top is the Selenium logo (a green square with a white checkmark and the letters 'Se') and the repository name 'Selenium'. Below the logo, a description states: 'Selenium is an umbrella project for a range of tools and libraries that enable and support the automation of web browsers.' It also shows '438 followers', the website 'https://selenium.dev/', and the Twitter handle '@SeleniumHQ'. A navigation bar includes links for 'Overview' (selected), 'Repositories' (19), 'Projects', 'Packages', 'People' (30), and a search icon. The main content area displays the 'README.md' file, titled 'The Selenium Project'. The text in the README reads: 'Selenium automates browsers. That's it! 🚀', 'What you do with that power is entirely up to you 💪', 'Primarily it is for automating web applications for testing purposes, but is certainly not limited to just that 🤖', and 'Boring web-based administration tasks can (and should) also be automated as well 📦'.

Selenium
Selenium is an umbrella project for a range of tools and libraries that enable and support the automation of web browsers.
438 followers <https://selenium.dev/> @SeleniumHQ

Overview Repositories 19 Projects Packages People 30

README.md

The Selenium Project

Selenium automates browsers. That's it! 🚀

What you do with that power is entirely up to you 💪

Primarily it is for automating web applications for testing purposes, but is certainly not limited to just that 🤖

Boring web-based administration tasks can (and should) also be automated as well 📦

Selenium Install

- Python bindings for Selenium

selenium 3.4.3

Python bindings for Selenium

Downloads ↓

Introduction

Python language bindings for Selenium WebDriver.

The *selenium* package is used to automate web browser interaction from Python.

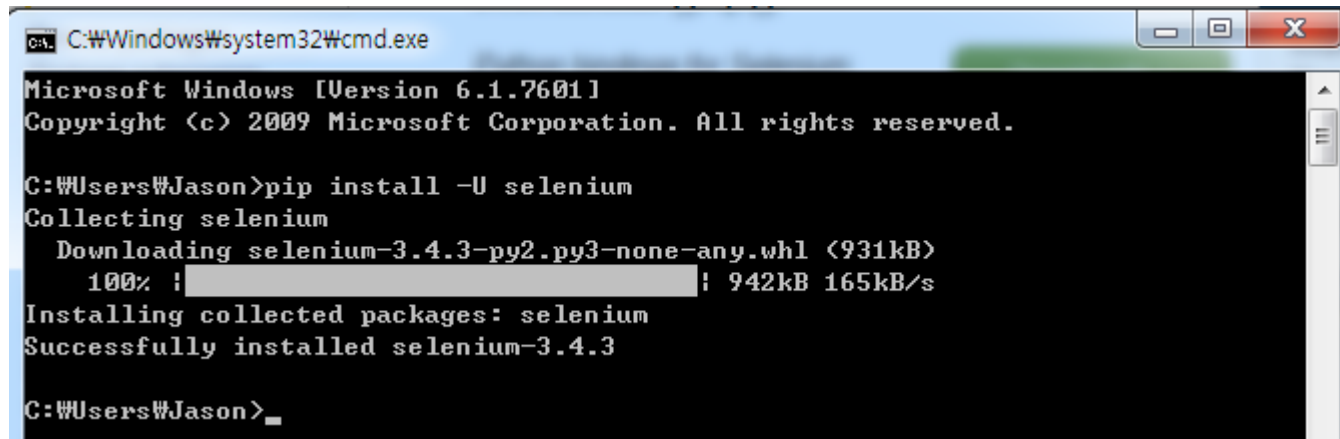
Home: <http://www.seleniumhq.org>

Docs: [selenium package API](#)

Dev: <https://github.com/SeleniumHQ/Selenium>

PyPI: <https://pypi.python.org/pypi/selenium>

IRC: **#selenium** channel on freenode



```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

C:\Users\Jason>pip install -U selenium
Collecting selenium
  Downloading selenium-3.4.3-py2.py3-none-any.whl (931kB)
    100% |#####| 942kB 165kB/s
Installing collected packages: selenium
Successfully installed selenium-3.4.3

C:\Users\Jason>
```

윈도우 명령 프롬프트를 실행한 후 다음 명령어를 입력

`pip install -U selenium`

ChromeDriver - WebDriver for Chrome

- <https://chromedriver.chromium.org/downloads>

Downloads

Current Releases

- If you are using Chrome version 108, please download [ChromeDriver 108.0.5359.22](#)
- If you are using Chrome version 107, please download [ChromeDriver 107.0.5304.62](#)
- If you are using Chrome version 106, please download [ChromeDriver 106.0.5249.61](#)
- For older version of Chrome, please see below for the version of ChromeDriver that supports it.

If you are using Chrome from Dev or Canary channel, please following instructions on the [ChromeDriver Canary](#) page.

For more information on selecting the right version of ChromeDriver, please see the [Version Selection](#) page.

ChromeDriver 108.0.5359.22

Supports Chrome version 108

- Resolved issue 4204: Fix the failing WPT test_top_level_and_one_child [Pri-]
- Resolved issue 4215: Rename chromedriver_mac64_m1 build to chromedriver_mac_arm64 [Pri-]
- Resolved issue 4221: Crash when a click command closes the browser [Pri-]

For more details, please see the [release notes](#).

ChromeDriver 107.0.5304.62

Supports Chrome version 107

For more details, please see the [release notes](#).







ChromeDriver 107.0.5304.18

Supports Chrome version 107

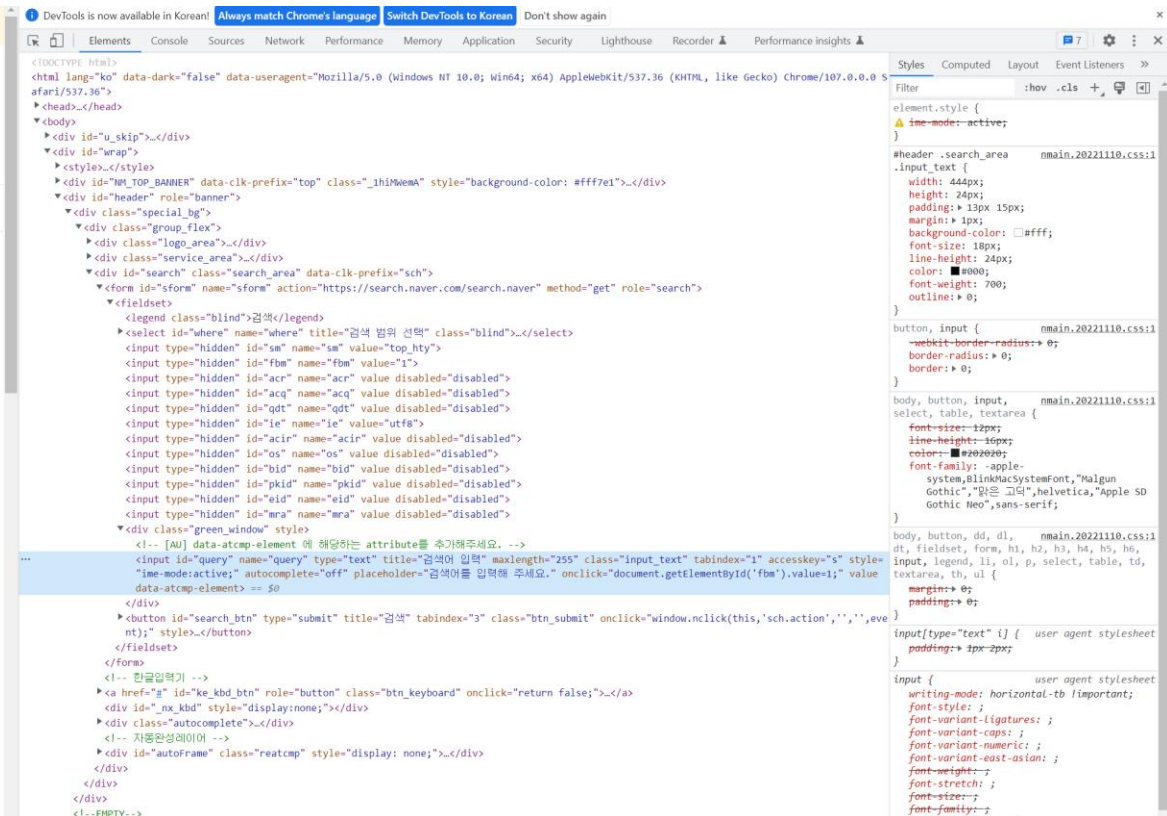
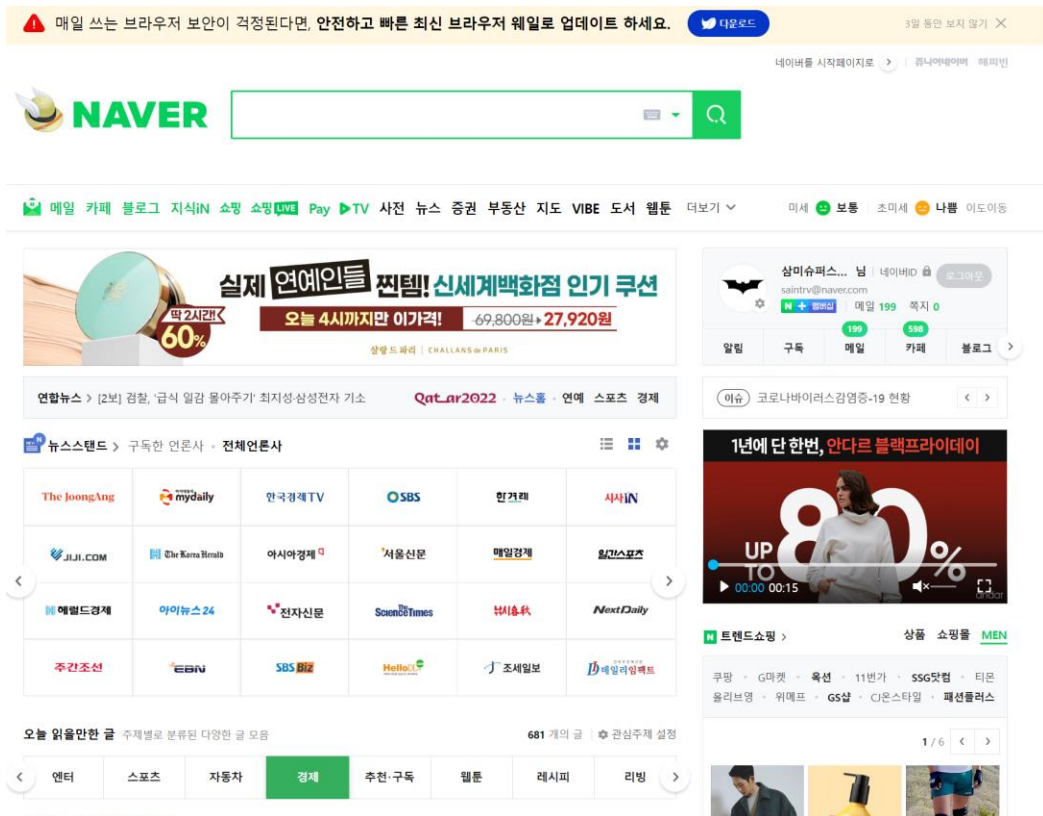
- Resolved issue 4180: Prohibit any BiDi commands handling until notified that BiDiMapper has started [Pri-]
- Resolved issue 4198: ChromeDriver BiDi should not assume that BiDiMapper is always in the first tab [Pri-]
- Resolved issue 4207: WebDriver error when running WPT [Pri-1]

For more details, please see the [release notes](#).

Index of /94.0.4606.41/

	Name	Last modified	Size	ETag
	Parent Directory		-	
	chromedriver linux64.zip	2021-09-09 09:57:33	9.42MB	9a00e20a5e5646c009c0ad74a446df10
	chromedriver mac64.zip	2021-09-09 09:57:36	7.80MB	f89f291ccc799a9de32e832807ce7e42
	chromedriver mac64_m1.zip	2021-09-09 09:57:38	7.15MB	69b92932297c2ce2764aa4c7744e1c17
	chromedriver win32.zip	2021-09-09 09:57:41	5.72MB	c853398ecdb35290e60e57a833e61752
	notes.txt	2021-09-09 09:57:45	0.00MB	b88b7fe40b1af6891a9f57da13e08914

Naver Search



Naver Search

```
from selenium import webdriver  
from selenium.webdriver.common.by import By  
import time
```

```
browser = webdriver.Chrome("chromedriver.exe")  
browser.get("http://www.naver.com")
```

#네이버 검색창 (id를 통해 찾거나 class를 통해 찾을 수 있습니다.)

```
input = browser.find_element(By.CLASS_NAME, "input_text")  
input.send_keys("미래에셋대우")  
input.submit()  
time.sleep(5)  
print(browser.page_source)
```


Chrome version error

```
File "C:\Users\saint\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.10_qbz5n2kfra8p0\LocalCache\local-packages\Python310\site-packages\selenium\webdriver\remote\webdriver.py", line 444, in execute
    self.error_handler.check_response(response)
File "C:\Users\saint\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.10_qbz5n2kfra8p0\LocalCache\local-packages\Python310\site-packages\selenium\webdriver\remote\errorhandler.py", line 249, in check_response
    raise exception_class(message, screen, stacktrace)
selenium.common.exceptions.SessionNotCreatedException: Message: session not created: This version of ChromeDriver only supports Chrome version 108
Current browser version is 107.0.5304.107 with binary path C:\Program Files\Google\Chrome\Application\chrome.exe
Stacktrace:
Backtrace:
    (No symbol) [0x00A1E563]
    (No symbol) [0x009A7FC1]
    (No symbol) [0x0089D04D]
    (No symbol) [0x008BF6B4]
    (No symbol) [0x008BA97C]
```

Naver Login

- Use Chrome Developer Tool to verify the ID of your ID/password

한국어

NAVER

PC방 등 공용PC라면 QR코드 로그인이 더 안전해요 X

ID 로그인 | 일회용 번호 | QR코드

☒ 로그인 상태 유지 ☒ IP보안

로그인

[비밀번호 찾기](#) | [아이디 찾기](#) | [회원가입](#)

네이버앱 200% 더 똑똑하게 쓰는 법
네이버앱 활용백서
[활용법 총정리 바로가기 >](#)

[이용약관](#) | [개인정보처리방침](#) | [책임의 한계와 법적고지](#) | [회원정보 고객센터](#)

NAVER Copyright © NAVER Corp. All Rights Reserved.

DevTools is now available in Korean! Always match Chrome's language Switch DevTools to Korean Don't show again

Elements Console Sources Network Performance Memory Application Security Lighthouse Recorder Performance insights

```
<!DOCTYPE html>
<html lang="ko">
<head>...</head>
<body>
  <div id="wrap" class="wrap">
    <div class="u_skip">...</div>
    <header class="header" role="banner">...</header>
    <div id="container" class="container">
      <!-- content -->
      <div class="content">
        <div class="login_wrap">
          <ul class="menu_wrap" role="tablist">...</ul>
          <form id="frmNIDLogin" name="frmNIDLogin" target="_top" autocomplete="off" action="https://nid.naver.com/nidlogin.login" method="POST">
            <input type="hidden" id="localechange" name="localechange" value="<!-- ... -->
            <input type="hidden" name="dynamicKey" id="dynamicKey" value="1A2ns5KeXao1StaAgJaOm3QgavQgsX0s1NiDD1IHd6bez66Rasuft_ubG5LTq_z2Ta_aVV9m-Z1oLEBybT58Jbk9g9F6YzjGU3pJxrELY">
            <input type="hidden" name="encpw" id="encpw" value="<!-- ... -->
            <input type="hidden" name="encpt" id="encpt" value="1">
            <input type="hidden" name="svctype" id="svctype" value="1">
            <input type="hidden" name="smart_LEVEL" id="smart_LEVEL" value="1">
            <input type="hidden" name="bvdsd" id="bvdsd" value="<!-- ... -->
            <input type="hidden" name="encnm" id="encnm" value="<!-- ... -->
            <input type="hidden" name="locale" id="locale" value="ko_KR">
            <input type="hidden" name="url" id="url" value="https://www.naver.com">
          <ul class="panel_wrap">
            <li class="panel_item" style="display: block;">
              <div class="panel_inner" role="tabpanel" aria-controls="loinid">
                <div class="id_pw_wrap">
                  <div class="input_row" id="id_line">...</div>
                  <div class="input_row" id="pw_line">
                    <div class="icon_cell" id="pw_cell">...</div>
                    <input type="password" id="pw" name="pw" placeholder="비밀번호" title="비밀번호" class="input_text" maxlength="16">...</div>
                    <span role="button" class="btn_delete" id="pw_clear" style="display: block;">...</span>
                  </div>
                </div>
                <div class="login_keep_wrap" id="login_keep_wrap">...</div>
                <div class="login_error_wrap" id="err_capslock" style="display: none;">...</div>
                <div class="login_error_wrap" id="err_empty_id" style="display: none;">...</div>
                <div class="login_error_wrap" id="err_empty_pw" style="display: none;">...</div>
                <div class="login_error_wrap" id="err_common" style="display: none;">...</div>
                <div class="btn_login_wrap">...</div>
              </div>
            </li>
          </ul>
        </form>
      </div>
    <ul class="find_wrap" id="find_wrap">...</ul>
  </div>
</body>
```

Naver Login

```
from selenium import webdriver
from selenium.webdriver.common.by import By
import time

browser = webdriver.Chrome("chromedriver.exe")
browser.get("https://nid.naver.com/nidlogin.login?mode=form&url=https%3A%2F%2Fwww.naver.com")

id = browser.find_element(By.ID, "id")
id.send_keys("myid")

passwd = browser.find_element(By.ID, "pw")
passwd.send_keys("mypassword")
passwd.submit()
time.sleep(5)
```

Google Image Search

- excute_script
 - Excute java script

```
# Function for scrolling to the bottom of Google
# Images results
def scroll_to_bottom():

    last_height = driver.execute_script('\
    return document.body.scrollHeight')

    while True:
        driver.execute_script('\
        window.scrollTo(0,document.body.scrollHeight)')

        # waiting for the results to load
        # Increase the sleep time if your internet is slow
        time.sleep(3)

        new_height = driver.execute_script('\
        return document.body.scrollHeight')

        # checking if we have reached the bottom of the page
        if new_height == last_height:
            break

        last_height = new_height
```

Google Image Search

- Screenshot(save_file_path)
 - Save image of node to save_file_path

```
# Loop to capture and save each image
for i in range(1, 50):

    # range(1, 50) will capture images 1 to 49 of the search results
    # You can change the range as per your need.
    try:
        print(i)
        # XPath of each image
        img = driver.find_element(By.XPATH,
            '//*[@id="islrq"]/div[0]/div[' +
            str(i) + ']/a[1]/div[1]/img')

        # Enter the location of folder in which
        # the images will be saved
        img.screenshot('Download-Location' +
            query + ' (' + str(i) + ').png')
        # Each new screenshot will automatically
        # have its name updated

        # Just to avoid unwanted errors
        time.sleep(0.2)

    except:

        # if we can't find the XPath of an image,
        # we skip to the next image
        continue
```

Google Image Search

- Use XPATH to find node

```
# What you enter here will be searched for in
# Google Images
query = "dogs"

# Creating a webdriver instance
driver = webdriver.Chrome("chromedriver.exe")

# Maximize the screen
driver.maximize_window()

# Open Google Images in the browser
driver.get('https://images.google.com/')

💡 Finding the search box
box = driver.find_element(By.XPATH, '//*[@maxlength="2048"]')
|

# Type the search query in the search box
box.send_keys(query)

# Pressing enter
box.send_keys(Keys.ENTER)
```


XPATH

- XPath (short for XML Path) is a query language used for navigating and selecting portions of an XML document.

```
<bib>
  <book price="30"> <publisher> Addison-Wesley </publisher>
    <author> Serge Abiteboul </author>
    <author> <first-name> Rick </first-name>
      <last-name> Hull </last-name>
    </author>
    <author> Victor Vianu </author>
    <title> Foundations of Databases </title>
    <year> 1995 </year>
  </book>
  <book price="55">
    <publisher> Freeman </publisher>
    <author> Jeffrey D. Ullman </author>
    <title> Principles of Database and Knowledge Base Systems </title>
    <year> 1998 </year>
  </book>
</bib>
```

Xpath: Summary

bib	matches a bib element
*	matches any element
/	matches the root element
/bib	matches a bib element under root
bib/paper	matches a paper in bib
bib//paper	matches a paper in bib , at any depth
//paper	matches a paper at any depth
paper book	matches a paper or a book
@price	matches a price attribute
bib/book/@price	matches price attribute in book , in bib
bib/book[@price<"55"]/author/lastname	matches...

XPATH

- `box = driver.find_element(By.XPATH, '//*[@maxlength="2048"]')`

```
▶ <div class="iblp" jsname="uFMOof">...</div> flex
▼ <div jscontroller="vZr2rb" class="a4bIc" jsname="gLfyf" jsaction="h5M12e;input:d3sQLd;blur:jI3wzf">
  ▶ <style data-impl="1668580592071">...</style>
  <div class="YacQv gsfi" jsname="vdLsw"></div>
  <input class="gLfyf gsfi" jsaction="paste:puy29d;" maxlength="2048" name="q" type="text" aria-auto
  </div>
▶ <div class="dRVYxd">...</div> flex
```

- *Selection of nodes by position*

`//a[3]` selects the third link in the document

`//table[last()]` selects the last table in the document

`//a[position() < 3]` selects the first three links in the document

- *Asterisks (*) match any set of characters or nodes, and can be used in a variety of situations*

`//table/tr/*` selects all children of `tr` tags in all tables (this is good for selecting cells using both `th` and `td` tags)

`//div[@*]` selects all `div` tags that have any attributes

- *Root nodes versus non-root nodes*

`/div` will select the `div` node only if it is at the root of the document

`//div` selects all `div`s anywhere in the document

- *Attribute selection*

`//@href` selects any nodes with the attribute `href`

`//a[@href='http://google.com']` selects all links in the document that point to Google

Google Image Search

- Use XPATH to find node

```
# What you enter here will be searched for in
# Google Images
query = "dogs"

# Creating a webdriver instance
driver = webdriver.Chrome("chromedriver.exe")

# Maximize the screen
driver.maximize_window()

# Open Google Images in the browser
driver.get('https://images.google.com/')

💡 Finding the search box
box = driver.find_element(By.XPATH, '//*[@maxlength="2048"]')

# Type the search query in the search box
box.send_keys(query)

# Pressing enter
box.send_keys(Keys.ENTER)
```