

# Ask Django

카페에 파이썬으로 글 쓰기

# 파이썬 웹요청 라이브러리

- MechanicalSoup [#github](#)
  - 자그마치 3개의 파이썬 라이브러리가 합쳐진 콤비네이션
  - requests, beautifulsoup, [mechanize](#)
  - 웹요청에 requests를 사용하기 때문에, 자바스크립트 처리 불가
- Selenium [#doc](#)
  - 웹브라우저 자동화 툴
  - 기존 웹브라우저를 사용하기에 자바스크립트 처리 가능

# 네이버 카페

- iframe과 JavaScript로 처리되는 부분이 상당히 많으며,
  - MechanicalSoup로 처리하면 좋으나, 너무 복잡 ;;;

실행성능은 낮아도, 일반적인 브라우저를 사용하는 *selenium*을 써봅시다.

# Selenium #Platforms

```
pip3 install selenium
```

- 지원하는 브라우저 : IE6이상, Chrome, Firefox, Safari, Opera 등
- 지원언어 : Python, Java, C#, JavaScript, Objective-C, Perl, PHP, R, Ruby, Haskell
- 지원OS : 윈도우, 맥, 리눅스

# WebDriver API

- 여러 브라우저에 대해, 단일 API 로서 접근
- 지원 종류
  - Internet Explorer Driver
  - PhantomJS Driver
  - Firefox Driver
  - Chrome Driver 등

# Selenium을 통한 네이버 검색 수행

```
from selenium import webdriver

driver = webdriver.Chrome('드라이버경로')

try:
    driver.get('http://naver.com') # 네이버 웹사이트 방문

    tag_query = driver.find_element_by_name('query')
    tag_query.send_keys('AskDjango')
    tag_query.submit()

    for tag in driver.find_elements_by_css_selector('.sh_blog_title'):
        print(tag.text, tag.get_attribute('href'))

finally:
    driver.close()
```

# BeautifulSoup 써보기

```
from bs4 import BeautifulSoup
from selenium import webdriver

driver = webdriver.Chrome('드라이버경로')
try:
    driver.get('http://naver.com') # 네이버 웹페이지 방문

    tag_query = driver.find_element_by_name('query')
    tag_query.send_keys('AskDjango')
    tag_query.submit()

    soup = BeautifulSoup(driver.page_source, 'html.parser')
    for tag in soup.select('.sh_blog_title'):
        print(tag.text, tag['href'])
finally:
    driver.close()
```

# Selenium을 통한 네이버 로그인 수행

```
driver.get('https://nid.naver.com/nidlogin.login') # 웹페이지 방문
```

```
tag_id = driver.find_element_by_name('id') # 태그 name속성으로 태그 찾기  
tag_id.send_keys('네이버아이디')
```

```
tag_pw = driver.find_element_by_name('pw') # 태그 name속성으로 태그 찾기  
tag_pw.send_keys('네이버암호')
```

```
tag_pw.submit()
```



# 카페글 쓰기

로그인을 먼저 수행

```
from selenium.webdriver.support.ui import Select
```

```
# 카페 모바일 페이지로 이동
```

```
# 모바일 페이지가 쓰기 편해요. Desktop용은 iframe :()
```

```
driver.get('https://m.cafe.naver.com/askdjango')
```

```
# 글쓰기 폼으로 이동
```

```
driver.find_elements_by_css_selector('a[class*=WriteArticle]')[0].click()
```

```
## 메뉴 선택
```

```
select = Select(driver.find_element_by_name('menuid'))
```

```
# select.select_by_value('1')
```

```
select.select_by_visible_text('파이썬 자동화 연습 게시판')
```

# 제목 쓰기

```
field = driver.find_elements_by_css_selector('input[name=subject]')[0]
field.clear()
field.send_keys('테스트 제목 by selenium')
```

# 내용쓰기

```
iframe = driver.find_elements_by_css_selector('iframe')[0]
driver.switch_to_frame(iframe)
```

```
body = driver.find_element_by_id('body')
body.click()
```

```
body.send_keys("""
hello world 1 - https://nomade.kr/vod/
hello world 2 - https://nomade.kr/vod/
hello world 3
hello world 4
""")
```

# iframe 빠져나오기

```
driver.switch_to_default_content()
```

# 완료

```
driver.find_element_by_css_selector('[class*=Submit]').click()
```

# Selenium 연습문제 (1)

- 각자 네이버 계정에 로그인 한 후에, 네이버 받은메일함의 첫 페이지 제목을 크롤링해보세요.

# Selenium 연습문제 (2)

- 네이버 웹툰. 오늘자 업데이트 목록을 크롤링해서, 네이버 카페 게시판에 업데이트 현황을 남겨보세요.

AskDjango 네이버카페는 자동화 연습을 위한 테스트카페입니다. 자유롭게 글을 남겨보세요. :D

*Life is short,  
use Python3/Django.*