

## <Python Web Crawler>

우선 `urllib.request` 라이브러리에서 부터 `urlopen` 함수를 가져오고  
`bs4` 라이브러리에서 `BeautifulSoup` 함수를 가져온다

그 후 `urlopen`이라는 함수를 통해 내가 원하는 사이트(크롤링 하고자 하는 사이트)의 주소를 가져온다.

그렇게 호출하게 되면 해당 웹페이지의 html 태그들을 응답하며 html 내용들을 받는다.

그 후 html 내용을 받은 다음, 그 내용에 접근할 수 있는 함수를 사용한다.

```
bsObject = BeautifulSoup(html, "html.parser")
```

그 후 `for` 반복문을 통해 해당 웹페이지의 object를 찾는다

내 코드의 경우에 네이버 뉴스의 기사를 가져오기 위해 '`a`' 와 '`href`' 를 사용하였지만, 이미지 소스나, 다른 것들도한 가져올 수 있다.

(요소검사를 통해 태그를 확인하며 원하는 object를 가져올 수 있다)

따라서 내가 만들고자 하는 목적에 따라 필요한 소스들을 뽑아낼 수 있다.