

urllib 사용법 및 기본 스크래핑

[11]:

```
1 import urllib.request as req
```

[12]:

```
1 # 파일 URL
2 img_url = "http://post.phinf.naver.net/20160621_169/14664824680681mSHj"
3 html_url = "http://google.com"
```

[13]:

```
1 # 다운받을 경로
2 save_path1 = "C:/Myexam/test1.jpg"
3 save_path2 = "C:/Myexam/index.html"
```

[14]:

```
1 # 예외 처리
2 try:
3     file1, header1 = req.urlretrieve(img_url, save_path1)
4     file2, header2 = req.urlretrieve(html_url, save_path2)
5 except Exception as e:
6     print("Download failed.")
7     print(e)
8 else:
9     # Header 정보 출력
10    print(header1)
11    print(header2)
```

```

Date: Fri, 13 Jan 2023 09:48:12 GMT
Accept-Ranges: bytes
Cache-Control: max-age=15552000
Content-Length: 797094
Content-Type: image/jpeg
Last-Modified: Fri, 13 Jan 2023 09:46:57 GMT
p3p: CP="ALL CURa ADMA DEVa TAIa OUR BUS IND PHY ONL UNI PUR FIN
COM NAV INT DEM CNT STA POL HEA PRE LOC OTC"
Age: 76
Expires: Wed, 12 Jul 2023 09:46:57 GMT
Connection: close

```

```

Date: Fri, 13 Jan 2023 09:48:13 GMT
Expires: -1
Cache-Control: private, max-age=0
Content-Type: text/html; charset=ISO-8859-1
Cross-Origin-Opener-Policy-Report-Only: same-origin-allow-popups; report-to="gws"
Report-To: {"group":"gws","max_age":2592000,"endpoints":[{"url":"https://csp.withgoogle.com/csp/report-to/gws/other"}]}
P3P: CP="This is not a P3P policy! See g.co/p3phelp for more info."
Server: gws
X-XSS-Protection: 0
X-Frame-Options: SAMEORIGIN
Set-Cookie: 1P_JAR=2023-01-13-09; expires=Sun, 12-Feb-2023 09:48:13 GMT; path=/; domain=.google.com; Secure
Set-Cookie: AEC=ARSKqsJgz8pAQiNY8AtMe3fT2ieXGVxdwW9lWkERmGGZUIr0tgKC5A3u4w; expires=Wed, 12-Jul-2023 09:48:13 GMT; path=/; domain=.google.com; Secure; HttpOnly; SameSite=lax
Set-Cookie: NID=511=IPHmzANtU0uS0jOfckVon1l8ndl5Farw7Q6Tin8bfuMt7hX54L0JKn7pDBXs2gYdiIHWtOrR_MHmAd86jS9fA-C2UaU3dXcBCS3KtdYeBFP2GCng-6w0tuhsI6w2JeZUU8lOYkB28sbuhKfVpjnDgfHFCkDbEt5BfjJyyldlpYA; expires=Sat, 15-Jul-2023 09:48:13 GMT; path=/; domain=.google.com; HttpOnly
Accept-Ranges: none
Vary: Accept-Encoding
Connection: close
Transfer-Encoding: chunked

```

[15]:

```

1 # 다운로드 파일 정보
2 print("Filename1 {}".format(file1))
3 print("Filename2 {}".format(file2))
4 print()

```

Filename1 C:/Myexam/test1.jpg

Filename2 C:/Myexam/index.html

[16]:

```
1 # 성공
2 print("Download Succeed.")
```

Download Succeed.

[]:

```
1
```

[]:

```
1
```