

urllib 사용법 및 기본 스크래핑

```
In [1]: import urllib.request as req
```

```
In [2]: # 파일 URL
img_url = "https://sesac.seoul.kr/static/common/images/www/common/img_ssac_intro.png"
html_url = "http://google.com"
```

```
In [3]: # 다운받을 경로
save_path1 = "C:/Myexam/test1.jpg"
save_path2 = "C:/Myexam/index.html"
```

```
In [4]: # 예외 처리
try:
    file1, header1 = req.urlretrieve(img_url, save_path1)
    file2, header2 = req.urlretrieve(html_url, save_path2)
except Exception as e:
    print("Download failed.")
    print(e)
else:
    # Header 정보 출력
    print(header1)
    print(header2)
```

```
Date: Tue, 25 Nov 2025 08:28:56 GMT
Content-Type: image/png
Content-Length: 31431
Connection: close
Accept-Ranges: bytes
ETag: W/"31431-1677198054000"
Last-Modified: Fri, 24 Feb 2023 00:20:54 GMT
```

```
Date: Tue, 25 Nov 2025 08:10:19 GMT
Expires: -1
Cache-Control: private, max-age=0
Content-Type: text/html; charset=ISO-8859-1
Content-Security-Policy-Report-Only: object-src 'none';base-uri 'self';script-src 'nonce-r8kV6KZ29myX4aC6bUuLDQ' 'strict-dynami
c' 'report-sample' 'unsafe-eval' 'unsafe-inline' https: http;;report-uri https://csp.withgoogle.com/csp/gws/other-hp
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: AEC=AaJma5sFFI0htyZQ9lVaub5uzt2_EIqSBi8pVjPR_FgFBT46k2PtWQaoYmk; expires=Sun, 24-May-2026 08:10:19 GMT; path=/; dom
ain=.google.com; Secure; HttpOnly; SameSite=lax
Set-Cookie: NID=526=h8p235ISUc2_YogGp3MQdC1DClmdpDZQR7vFEoMvp6zq1wKt0Lcsn4AW7AunFvOHhZdIJD1XGEjt-VI-Q3Z7bhEeLXqA1ChFYzKRb-qS1A9
69k5k9U4cddF10PLlgceEtDh1vfKOGmnJ0wiKCDuFBlawWijptq_t_QPjDjxwY9zSvUqr9Y1C0gouqB2yJKfaCU5KebORH5sje8upPntBDT9GAXf0w; expires=We
d, 27-May-2026 08:10:19 GMT; path=/; domain=.google.com; HttpOnly
Accept-Ranges: none
Vary: Accept-Encoding
Connection: close
Transfer-Encoding: chunked
```

In [5]:

```
# 다운로드 파일 정보
print("Filename1 {}".format(file1))
print("Filename2 {}".format(file2))
print()
```

```
Filename1 C:/Myexam/test1.jpg
Filename2 C:/Myexam/index.html
```

```
In [6]: # 성공  
        print("Download Succeed.")
```

Download Succeed.

```
In [ ]:
```