

# Lecture du code source d'une page web via la librairie urllib

urllib-read-01

```
#!/usr/bin/env python
# -*- coding: utf-8 -*-
"""
exemple de programme pour obtenir le code d'une page avec la librairie
urllib
source :
https://realpython.com/python-web-scraping-practical-introduction/
"""
from urllib.request import urlopen

site_url = 'https://dvillers.umons.ac.be/wiki/page_simple'
page = urlopen(site_url)
print(page)
# page est un objet urllib

html_bytes = page.read()
html = html_bytes.decode("utf-8")
# html est une chaîne de caractères contenant la page web, après
décodage

lines = html.splitlines()
print(lines)
# lines contient les lignes de la page web

print('*****')

# affichage de toutes les lignes
for line in lines:
    print(line)
```

From:

<https://dvillers.umons.ac.be/wiki/> - **Didier Villers, UMONS - wiki**

Permanent link:

<https://dvillers.umons.ac.be/wiki/teaching:progappchim:urllib>

Last update: **2022/03/21 15:46**



