

Forum:[Hack Forums](#)

Thread: [proxy grabber \[what do you guys think?\]](#)

ExoticModzV1 - February 7, 2016, 7:39 am

i only wanted to share this code because i see no one use BeautifulSoup to grab proxies!!!
coded in 25minutes.

this is an example of how to use bs4 to grab data off tables ^_^

i think more people need to use BeautifulSoup, its a brilliant module, and so easy to scrape data with.

Code '''

coded By ExoticModsV1

'''

```
from bs4 import BeautifulSoup
import urllib2
```

```
def GrabProxies():
```

```
    wfile = open('proxies.txt', 'w')
```

```
    url = " https:// free-proxy-list .net/"
```

```
    req = urllib2.Request(url, headers={'User-Agent' : "Magic Browser"})
```

```
    con = urllib2.urlopen(req)
```

```
    counter = 0
```

```
    soup = BeautifulSoup(con)
```

```
    table = soup.find("table")
```

```
    for row in table.findAll('tr')[1:100]:
```

```
        counter = counter + 1
```

```
        col = row.findAll('td')
```

```
        col1 = col[0].getText()
```

```
        col2 = col[1].getText()
```

```
        f = (col1 + ":" + col2)
```

```
        wfile.write("\n" + str(col1) + ":" + str(col2))
```

```
        print(f)
```

```
    url = " https:// www .socks -proxy .net/"
```

```
    req = urllib2.Request(url, headers={'User-Agent' : "Magic Browser"})
```

```
    con = urllib2.urlopen(req)
```

```
    counter = 0
```

```
    soup = BeautifulSoup(con)
```

```
    table = soup.find("table")
```

```
    for row in table.findAll('tr')[1:50]:
```

```
        counter = counter + 1
```

```
        col = row.findAll('td')
```

```
        col3 = col[1].getText()
```

```
col2 = col[-8].getText()
print(col2 + ":" + col3)
f = (col2 + ":" + col3)
```

```
wfile.write("\n" + str(col2) + ":" + str(col3))
```

```
url = " https :/ /www .us-proxy .org/"
```

```
req = urllib2.Request(url, headers={'User-Agent' : "Magic Browser"})
con = urllib2.urlopen(req)
counter = 0
soup = BeautifulSoup(con)
table = soup.find("table")
for row in table.findAll('tr')[1:100]:
```

```
    counter = counter + 1
    col = row.findAll('td')
    col1 = col[0].getText()
    col2 = col[1].getText()
    print(col1 + ":" + col2)
    f = (col1 + ":" + col2)
```

emadaldin1 - February 7, 2016, 7:59 am

Very nice coding indeed it grab proxies from 3 diff sites :D

Subrosia - February 13, 2016, 8:45 am

I get this

Quote:line 64

PM Find Quote Quote this post Report

^

SyntaxError: invalid syntax

```
sudo python3 ~/Desktop/proxygrabberHF.py
(yes I created proxies.txt in the same directory)
```

Also, take a look at this code. Very similar to what youre doing. I get an error on line 40 with that one though.

[github.com /OffensivePython/Proxist/blob/master/proxist.py](https://github.com/OffensivePython/Proxist/blob/master/proxist.py)

ExoticModzV1 - February 15, 2016, 12:53 pm

(Feb 13, 2016 09:45 AM)Subrosia Wrote: [\[CITE\]](#) I get this

```
sudo python3 ~/Desktop/proxygrabberHF.py
(yes I created proxies.txt in the same directory)
```

Also, take a look at this code. Very similar to what youre doing. I get an error on line 40 with that one though.

github.com /OffensivePython/Proxist/blob/master/proxist.py
whats the error, i think the website i was scraping off has changedd the tables so it actually is not
scraping anything :)

(Feb 13, 2016 09:45 AM)Subrosia Wrote: [\[CITE\]](#) I get this

```
sudo python3 ~/Desktop/proxygrabberHF.py  
(yes I created proxies.txt in the same directory)
```

Also, take a look at this code. Very similar to what youre doing. I get an error on line 40 with that
one though.

github.com /OffensivePython/Proxist/blob/master/proxist.py
yes, its simular, but more complicated, and not as messy as my code, haha.

Wireless Network - February 16, 2016, 2:05 pm

I've never used bs4, but i'll definitely give it a shot.

ExoticModzV1 - February 16, 2016, 5:46 pm

(Feb 16, 2016 03:05 PM)Wireless Network Wrote: [\[CITE\]](#) I've never used bs4, but i'll definitely give it
a shot.

thank you, and good luck! to grab proxies you only need 3-10 lines of code. its easy!

emadaldin1 - February 19, 2016, 11:11 pm

I've never used bs4, but thanks a lot :D

CyberError - February 20, 2016, 7:44 am

So its grab proxy from different proxy site? So its grab working proxy ?

AuntJason - February 21, 2016, 7:06 am

I'd suggest a less obvious UA than Magic Browser

adramax- - February 22, 2016, 2:47 am

Awesome Info and great help , thanks a ton mate !!! :)

ExoticModzV1 - February 27, 2016, 1:56 pm

(Feb 21, 2016 08:06 AM)AuntJason Wrote: [\[CITE\]](#) I'd suggest a less obvious UA than Magic Browser
yers, buts i coded it very short time, and its just an example to help people

EwilDave - February 27, 2016, 9:29 pm

I'd probably just tell people to use VPNGate and OpenVPN
mostly because it's very easy to scrape VPNs from vpngate.com/api/phone
which is exactly what the python vpngate project does!

This is the actual gist on Github: gist.github.com/Lazza/bbc15561b65c16db8ca8

Sophocles - March 1, 2016, 3:28 am

Why not use requests just curious ?

ExoticModzV1 - March 2, 2016, 11:30 am

(Mar 1, 2016 04:28 AM)Drays Wrote: [\[CITE\]](#) Why not use requests just curious ?
prefer urllib2, people find that weird ;P

bet4r00t - March 3, 2016, 8:53 am

Nice work :) I'll send a link to my friend he was talking about this in school but didn't know how to make one haha