

Forum:[Hack Forums](#)

Thread: [How to parse js with python](#)

Turnikman - October 16, 2018, 2:40 pm

Hi i have a website

When i try to parse it with bs4 and search only for paragraphs it shows: <p id = "test"></p>

But there is jscript: document.getElementById("test").innerHTML = "It works";

and should show: It works

i have a code that "should simulate the browser" and show the right result, the problem is that the code doesnt work.

code:

```
from bs4 import BeautifulSoup as soup
import sys
import requests
from PyQt5.QtWebEngineWidgets import QWebEnginePage
from PyQt5.QtWebEngineWidgets import QWebEngineView
from PyQt5.QtWidgets import QApplication
from PyQt5.QtCore import QUrl
class Client(QWebEnginePage):
    def __init__(self, url):
        self.app = QApplication(sys.argv)
        QWebEnginePage.__init__(self)
        self.loadFinished.connect(self.on_page_load)
        self.mainFrame().load(QUrl(url))
        self.app.exec()

    def on_page_load(self):
        self.app.quit()
url = "test"
client_response = Client(url)
file = client_response.mainFrame().toHtml()

sp = soup(file, "html.parser")
print(sp.find_all("p"))
```

and it gives me this error:

```
client_response = Client(url)
```

TypeError: arguments did not match any overloaded call:

```
QWebEnginePage(parent: QObject = None): argument 1 has unexpected type 'str'
```

```
QWebEnginePage(QWebEngineProfile, parent: QObject = None): argument 1 has unexpected type 'str'
```

i dont want to use selenium, because you should specify the browser

thanks and sorry for my bad english ?

Infamia - October 16, 2018, 2:50 pm

Code def _init_(self, url):
double under scores, like this:

Code def __init__(self, url):

Turnikman - October 17, 2018, 12:59 am

Thanks, i made it work but i have another problem

it parse it in the beggining, but when i change the value and run the code again it doesnt work

i have a button in the html, and when i press it i want to change the output, but the python code doesnt catch it

this is working

```
<p class = "jstest" id = "isactive" >Doesnt work</p>  
<script> document.getElementById("isactive").innerHTML = "it works";  
</script>
```

but this part - not :

```
function toggle(){  
    if(!toggled){  
  
        toggled = true;  
        st.status = "On";  
        document.getElementById("isactive").innerHTML = st.status;  
        return;  
  
    }  
.  
.  
.
```