

Trying to get Title and URL into csv from top 3 articles. Getting 'NoneType' object has no attribute 'get_text'

Asked today Modified today Viewed 32 times



Trying to get Title and URL into csv from top 3 articles. Getting 'NoneType' object has no attribute 'get_text'.

0



I have tried multiple different lines of code and looked through the html lines to try different ones there also. Been doing this for like 2 hours and I am tired of being stuck.



I would like to run a list of URLs from the same website to web scrape for PRs from different companies. I would like to be able to copy and paste this code and duplicate it easily for different websites also.

```
from bs4 import BeautifulSoup
import requests
import pyperclip
```

```
urls = ['https://www.businesswire.com/portal/site/home/search/?
searchType=all&searchTerm=lobe%20sciences&searchPage=1',
        'https://www.businesswire.com/portal/site/home/search/?
searchType=all&searchTerm=enveric&searchPage=1',
        'https://www.businesswire.com/portal/site/home/search/?
searchType=all&searchTerm=cybin&searchPage=1',
        'https://www.businesswire.com/portal/site/home/search/?
searchType=all&searchTerm=delix&searchPage=1'
]
#scrape elements
for url in urls:
    response = requests.get(url)
    soup = BeautifulSoup(response.content, "html.parser")
```

```
#print titles only
h3 = soup.find("h3", class_="bw-news-list")
print(h3.get_text())
```

I was expecting it to print the URLs and titles for the articles at each URL above. Then I want to be able to append this to a csv

[python](#) [html](#) [web-scraping](#)

Share Edit Follow Flag

edited 1 hour ago



[John Gordon](#)

31.1k 8 39 61

asked 2 hours ago



[Steve](#)

21 3

New contributor

▲ soup.find("h3", class_="bw-news-list") did not find any matching element, so it returned None.
🚩 – [John Gordon](#) 1 hour ago

▲ I only see one <h3> element in that screenshot, and it doesn't have class="bw-news-list" .
🚩 – [John Gordon](#) 1 hour ago

▲ @JohnGordon when I change the element to ul, that is with that class, it still gives me the same error.
🚩 – [Steve](#) 1 hour ago

1 Answer

Sorted by: Highest score (default) ▾



0



You could try using Selenium instead of BeautifulSoup. I noticed some JavaScript and Ajax calls on the pages you're attempting to scrape. BeautifulSoup can scrape data from static webpages but Selenium can execute JavaScript. You can inspect the DOM structure of your `soup` variable to confirm whether it lacks any `<h3>` tags with the class name you're targeting.

Here is my Selenium code that works, but be aware that it's memory-intensive and requires a web driver to operate.



```
from selenium import webdriver
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.by import By
from selenium.webdriver.support.wait import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC

def initialize_webdriver(headless=False):
    chrome_options = Options()
    if headless:
        chrome_options.add_argument('--headless')
    driver = webdriver.Chrome('../chromedriver-win64/chromedriver.exe',
options=chrome_options)
    return driver
```

```

if __name__ == '__main__':
    chrome_driver = initialize_webdriver(headless=True)

    urls = [
        'https://www.businesswire.com/portal/site/home/search/?searchType=all&searchTerm=lobe%20sciences&searchPage=1',
        'https://www.businesswire.com/portal/site/home/search/?searchType=all&searchTerm=enveric&searchPage=1',
        'https://www.businesswire.com/portal/site/home/search/?searchType=all&searchTerm=cybin&searchPage=1',
        'https://www.businesswire.com/portal/site/home/search/?searchType=all&searchTerm=delix&searchPage=1']

    try:
        for url in urls:
            # loading the webpage
            chrome_driver.get(url)

            # wait for the list of items to load
            WebDriverWait(chrome_driver, 10).until(
                EC.presence_of_element_located((By.XPATH, "//ul[@class='bw-news-
list']"))))

            # scraping top 3 sites
            list_item = chrome_driver.find_element(By.XPATH, "//ul[@class='bw-news-
list']")

            top_sites = list_item.find_elements(By.XPATH, "//li/h3/a")
            for index, item in enumerate(top_sites[:3], 0):
                print(f'Title {index}: {item.text}')
                print(f"URL {index}: {item.get_attribute('href')}")
                print('-----')

    finally:
        chrome_driver.quit()

```

Prints:

```

Title 0: Lobe Sciences Provides Company Review of 2023
URL 0: http://www.businesswire.com/news/home/20240108837220/en/Lobe-Sciences-Provides-Company-Review-of-2023
Title 1: Patient-Reported Outcomes from the CARTITUDE-4 Study Showed Clinically Meaningful Improvements in Health-Related Quality of Life and Reductions in Multiple Myeloma Symptoms Following Treatment with CARVYKTI® (ciltacabtagene autoleucel)
URL 1: http://www.businesswire.com/news/home/20231211889964/en/Patient-Reported-Outcomes-from-the-CARTITUDE-4-Study-Showed-Clinically-Meaningful-Improvements-in-Health-Related-Quality-of-Life-and-Reductions-in-Multiple-Myeloma-Symptoms-Following-Treatment-with-CARVYKTI%C2%AE-ciltacabtagene-autoleucel
Title 2: Lobe Sciences Provides Update on L-130 Clinical Program
URL 2: http://www.businesswire.com/news/home/20231207780834/en/Lobe-Sciences-Provides-Update-on-L-130-Clinical-Program
-----
Title 0: Membership_App_French_Jun08rev
URL 0:
http://www.businesswire.com/portal/binary/com.epicentric.contentmanagement.servlet.Content
Title 1: Enveric Biosciences Updates Key Data Underpinning Selection of EB-003 as Lead Drug Candidate from EVM301 Series
URL 1: http://www.businesswire.com/news/home/20240103821215/en/Enveric-Biosciences-

```

Updates-Key-Data-Underpinning-Selection-of-EB-003-as-Lead-Drug-Candidate-from-EVM301-Series

Title 2: Enveric Biosciences Announces Exercise of Warrants and Issuance of New Warrants in a Private Placement for \$1.8 Million Gross Proceeds Priced At-the-Market

URL 2: <http://www.businesswire.com/news/home/20231229768169/en/Enveric-Biosciences-Announces-Exercise-of-Warrants-and-Issuance-of-New-Warrants-in-a-Private-Placement-for-1.8-Million-Gross-Proceeds-Priced-At-the-Market>

Title 0: Cybin Announces Grant of Two Additional Patents in Japan in Support of its DMT Program

URL 0: <http://www.businesswire.com/news/home/20240207711728/en/Cybin-Announces-Grant-of-Two-Additional-Patents-in-Japan-in-Support-of-its-DMT-Program>

Title 1: Cybin Announces FDA Clearance to Initiate a Phase 2a Study of CYB004 in Generalized Anxiety Disorder

URL 1: <http://www.businesswire.com/news/home/20240123732289/en/Cybin-Announces-FDA-Clearance-to-Initiate-a-Phase-2a-Study-of-CYB004-in-Generalized-Anxiety-Disorder>

Title 2: Cybin Announces Positive Topline Data from Phase 1 Studies of Proprietary Deuterated DMT Molecules CYB004 and SPL028

URL 2: <http://www.businesswire.com/news/home/20240108181308/en/Cybin-Announces-Positive-Topline-Data-from-Phase-1-Studies-of-Proprietary-Deuterated-DMT-Molecules-CYB004-and-SPL028>

Title 0: Delix Therapeutics Awarded National Institutes of Health Grant to Advance Vital Research of Novel Neuroplastogen for Substance Use Disorders

URL 0: <http://www.businesswire.com/news/home/20231213417680/en/Delix-Therapeutics-Awarded-National-Institutes-of-Health-Grant-to-Advance-Vital-Research-of-Novel-Neuroplastogen-for-Substance-Use-Disorders>

Title 1: Delix Presents Interim Data From Phase I Trial of Novel Neuroplastogen at ACNP Annual Meeting

URL 1: <http://www.businesswire.com/news/home/20231206561331/en/Delix-Presents-Interim-Data-From-Phase-I-Trial-of-Novel-Neuroplastogen-at-ACNP-Annual-Meeting>

Title 2: Freedom Biosciences Emerges from Stealth with \$10.5 Million Seed Financing to Develop Next-Generation Ketamine and Psychedelic-based Mental Health Treatments

URL 2: <http://www.businesswire.com/news/home/20220823005203/en/Freedom-Biosciences-Emerges-from-Stealth-with-10.5-Million-Seed-Financing-to-Develop-Next-Generation-Ketamine-and-Psychedelic-based-Mental-Health-Treatments>

Share Edit Delete Flag

answered 28 mins ago



Chaithra KC

31 4