

Python Web Scraper that Prints out the Most Recent Articles from Google News

```
scraper.py > ...
1  import bs4
2  import sys
3  from bs4 import BeautifulSoup as soup
4  from urllib.request import urlopen
5
6  news_url = "https://news.google.com/news/rss"
7  Client = urlopen(news_url)
8  xml_page = Client.read()
9  Client.close()
10
11  soup_page = soup(xml_page, "xml")
12  news_list = soup_page.findAll("item")
13
14  sys.stdout = open("Articles.txt", "w")
15
16  #print news title, url and publication date
17
18  for id, news in enumerate(news_list, 1):
19      print("\n")
20      print(id, " ", news.title.text)
21      print(news.link.text)
22      print(news.pubDate.text)
23      print("-"*60)
24
25  sys.stdout.close()
```

```
3  1  Fauci wants a national coronavirus stay-at-home order, but Trump doesn't - Newsweek
4  https://news.google.com/_i/rss/rd/articles/CBMieGh0dHBzOi8vd3d3Lm51d3N3ZWVrLmNvbS9mYXVjaS1zYX1zLWFsbC1zdGF0ZXMt2hvdWxkL
5  Fri, 03 Apr 2020 04:13:22 GMT
6  -----
7
8
9  2  Today's coronavirus update: 10M out of work, NYC nurses finally get help - New York Post
10 https://news.google.com/_i/rss/rd/articles/CBMiZGh0dHBzOi8vb3d3Lm51d3N3ZWVrLmNvbS9mYXVjaS1zYX1zLWFsbC1zdGF0ZXMt2hvdWxkL
11 Thu, 02 Apr 2020 21:38:48 GMT
12 -----
13
14
15 3  Trump rails against 'witch hunts' after Pelosi announces committee to oversee coronavirus response - USA TODAY
16 https://news.google.com/_i/rss/rd/articles/CBMiIAFodHRwczovL3d3dy51c2F0b2RheS5jb20vc3RvcnkvbWV3cy9wb2xpdG1jcy8yMDIwLzA0L
17 Fri, 03 Apr 2020 01:51:38 GMT
18 -----
19
20
21 4  ♦The holiday of a lifetime♦: An oral history of the infected, rejected Zaandam cruise ship - The Washington Post
22 https://news.google.com/_i/rss/rd/articles/CBMiIAFodHRwczovL3d3dy53YXNoaW5ndG9ucG9zdC5jb20vbGlmZXN0eWx1L3N0eWx1L3RoZS1ob
23 Fri, 03 Apr 2020 03:23:38 GMT
24 -----
```