## RSS Feeds

Al Jazeera provides official RSS feeds for several sections of the website, including:  
- All News: https://www.aljazeera.com/xml/rss/all.xml  
- Economy: https://www.aljazeera.com/xml/rss/economy.xml  
- Opinion: https://www.aljazeera.com/xml/rss/opinion.xml  
- Sports: https://www.aljazeera.com/xml/rss/sport.xml  
  
These feeds are in standard XML format and include elements like <title>, <link>, and <pubDate> for each article.

import requests

import xml.etree.ElementTree as ET

rss\_url = "https://www.aljazeera.com/xml/rss/all.xml"

response = requests.get(rss\_url)

rss\_content = response.content

root = ET.fromstring(rss\_content)

channel = root.find('channel')

items = channel.findall('item')

for item in items[:5]:

title = item.find('title').text

link = item.find('link').text

pub\_date = item.find('pubDate').text

print(f"Title: {title}")

print(f"Link: {link}")

print(f"Published: {pub\_date}")

print("-" \* 50)

output:

A screenshot of a computer program

AI-generated content may be incorrect.

<https://www.aljazeera.com/video/inside-story/2025/5/17/how-to-achieve-a-lasting-ceasefire-between-russia-and-ukraine?traffic_source=rss>

<https://www.aljazeera.com/news/2025/5/17/one-killed-after-explosion-near-fertility-clinic-in-palm-springs?traffic_source=rss>

## API Access :

While browsing Al Jazeera’s website, several internal requests returning JSON data were observed via XHR and Fetch operations. These requests resemble API behavior but are **not officially documented** or intended for public use.

Moreover, the website’s robots.txt file explicitly states:  
Disallow: /api  
This clearly prohibits any form of automated access to the API endpoints.

As a result, **automated tools and crawlers must not interact with these API routes**.

Instead, Al Jazeera provides a safer and officially supported alternative:  
**RSS feeds**, which offer structured and up-to-date content retrieval across various sections of the site.

A screenshot of a computer

AI-generated content may be incorrect.



### API get Code :

import requests

url = "https://www.aljazeera.com/api"

#url = "https://jsonplaceholder.typicode.com/posts"

headers = {

"User-Agent": "MyCrawler/1.0",

"Accept": "application/json"

}

response = requests.get(url, headers=headers)

if response.status\_code == 200:

data = response.json()

for post in data[:5]:

print(f"Post ID: {post['id']}, Title: {post['title']}")

else:

print(f"Failed to get data, status code: {response.status\_code}")