Africa news Web scraping using Python, Beautiful soup and Selenium

Prepared By: Muhammad Raheeq ur Rehman

Role: Interne

Date:28th July, 2025

# Table of Contents

1. 1. Introduction
2. 2. Objective
3. 3. Tools and Technologies
4. 4. System Design / Workflow
5. 5. Code Overview
6. 6. Sample Output (CSV)
7. 7. Conclusion

# 1. Introduction

This project is aimed at scraping articles from the Africanews website using Python. It automates the extraction of news titles, links, and other relevant data, storing the results in a CSV file.

# 2. Objective

The primary objective of this project is to practice and demonstrate web scraping techniques using Python libraries such as BeautifulSoup and Selenium, and to store the extracted data for analysis.

# 3. Tools and Technologies

- Python  
- BeautifulSoup  
- Selenium  
- Pandas  
- CSV File Handling  
- HTML (for display purposes)

# 4. System Design / Workflow

1. Load the Africanews homepage  
2. Use Selenium to load dynamic content (if necessary)  
3. Use BeautifulSoup to parse HTML and extract article titles and links  
4. Store the extracted data in a structured CSV format  
5. Optionally visualize or display the output using HTML

# 5. Code Overview

The project consists of two main Python scripts:  
- main.py: Initializes scraping logic and handles data writing  
- Pythonproject.py: Contains reusable functions and data extraction methods  
Both files use exception handling to avoid runtime errors, and rely on page structure for data retrieval.

# 6. Sample Output (CSV)

Below is a preview of the first few rows of the scraped data:

|  |  |  |  |
| --- | --- | --- | --- |
| **HEADLINE** | **DATE** | **LINK** | **DESCRIPTION** |
| Russia scales down celebrations honoring its navy as Ukraine launches more drone attacks | 16 minutes ago | https://www.africanews.com/2025/07/28/russia-scales-down-celebrations-honoring-its-navy-as-ukraine-launches-more-drone-attacks/ | Russia on Sunday scaled down festivities honoring its navy, citing security concerns as continuing Ukrainian drone attacks posed a challenge to ... |
| Relief and frustration in Europe as US-EU trade deal agrees 15% tariff on most European exports | 1 hour ago | https://www.africanews.com/2025/07/28/relief-and-frustration-in-europe-as-us-eu-trade-deal-agrees-15-tariff-on-most-european-exp/ | The US and EU have agreed a framework for a trade deal, bringing relief that a transatlantic trade war has been averted but frustration for many ... |
| Coalition led by paramilitary RSF forms parallel government in Sudan | 27/07 - 12:25 | https://www.africanews.com/2025/07/27/coalition-led-by-paramilitary-rsf-forms-parallel-government-in-sudan/ | The news raises concerns of further divisions in the country that could prolong the devastating civil war. |
| Increasing number of African Americans relocating to Africa | 26/07 - 16:32 | https://www.africanews.com/2025/07/26/increasing-number-of-african-americans-relocating-to-africa/ | Those who've made the move describe the need to connect with their history and find a better life. |
| New law to punish Russians who search online for 'extremist' content | 26/07 - 14:35 | https://www.africanews.com/2025/07/26/new-law-to-punish-russians-who-search-online-for-extremist-content/ | The official definition of extremist activity is extremely broad and includes opposition groups and the 'international LGBT movement'. |

# 7. Conclusion

This project successfully demonstrates a real-world application of web scraping using Python. The extracted data can be useful for news aggregation, sentiment analysis, or content research.