

Task: Developing a News Website Using Web Scraping

Project Overview: The goal of this project is to develop a news website that aggregates and displays news articles from various sources using web scraping. The website should provide a user-friendly interface, categorize news articles, and update automatically at regular intervals.

Project Requirements:

1. Web Scraping Component

- Identify reliable news source (e.g., BBC, CNN, The Guardian, etc.).
- Use web scraping tools such as BeautifulSoup, Scrapy, or Selenium to extract relevant news data in real-time.
- Extract essential information:
 - Headline
 - Author (if available)
 - Published date
 - Summary
 - Article content
 - Image URL
 - Source link
- Implement error handling to manage failed scrapes or website structure changes.

2. Backend Development

- Develop backend using Flask/Django/Node.js to fetch news data in real-time.

3. Frontend Development

- Use React/Angular/Vue.js for the frontend.
- Design a user-friendly UI with the following pages:

- Homepage: Display top news articles with images and summaries.
- Categories: Categorized sections (Politics, Technology, Sports, Entertainment, etc.).
- Search: Implement a search functionality to find specific news articles.
- Detailed News Page: Show full article details when clicked.
- Ensure responsiveness and cross-browser compatibility.

4. Features & Future Enhancements (Optional)

- Implement pagination.
- Add filters (date, source, category, etc.).
- Enable bookmarking of articles.
- Implement user authentication for personalized experiences.
- Provide dark mode support.
- Add a news API fallback mechanism in case scraping fails.

5. Deployment

- Deploy the application on free cloud service provider (Ex. Netlify, PythonAnyWhere, etc.)

Deliverables:

1. Web scraper script with documentation.
2. Backend API serving real-time news data.
3. Fully functional frontend with proper UI/UX.
4. Deployed news website with live updates.
5. Project report covering methodology, challenges, and solutions.

Timeline:

- Two Weeks