

2 MINUTE READS - ARTICLE SUMMARIZER

OBJECTIVE:

To create a tool that condenses lengthy blog posts and articles into summaries.

FEATURES:

- Web-crawlers to gather content from articles
- NLP for summarization.
- User interface for input and displaying results.

HOW IT WOULD WORK:

1. DEVELOPING THE WEB CRAWLER
 - Set Up Crawler: Use BeautifulSoup to create a crawler that can search the web based on the input and collect URLs of relevant articles.
 - Parse Content: Extract the text content from the collected URLs
2. IMPLEMENT THE SUMMARIZATION ALGORITHM
 - Implement a basic text summarization model using a pre-trained T5 model from the Hugging Face Transformers library.
 - Fine-Tuning: Fine-tune the model on a dataset of articles and their summaries if necessary.
3. BACKEND/FRONTEND:
 - Backend: Set up a server using Flask to handle the input, trigger the web-crawler, and process the summarization.
 - Frontend: Create a simple UI using HTML, CSS, and JavaScript or a framework like React to allow users to enter a search string and view the summarized results.
4. TESTING AND OPTIMIZATION

Repository: <https://github.com/shravanisaraf/two-minute-reads-summarizer.git>