TEAM DETAILS AND PROBLEM STATEMENT

- Problem Statement : **Global Article Summarization and Link Analysis Platform(Problem 1)**
- Team Name : **Quest.ai**
- Team Leader Details (Name, Phone Number, Email) : **Lakshya Raj Vijay, 9602524537, lrvijay2003@gmail.com**
- Institution Name : **SRM Institute of Science and Technology, Kattankulathur**
- Course Enrolled : **B.TECH & CSE**

OPEN HACK

SHAPE TOMORROW

Problem Statement

In today's fast-paced world, we are bombarded with an overwhelming number of articles and updates that significantly impact our lives. However, **identifying the most important** ones and effectively understanding them can be quite challenging. Even when we find relevant articles, **extracting meaningful insights** can be difficult. This problem affects millions of people daily. Even if they understand a particular article, they often lack **related references**, forcing them to embark on another search for additional information or articles. This process can be further complicated by the possibility of finding crucial information in **different languages**.

Objectives

01

Develop an intelligent system that identifies and aggregates the most important articles based on individual needs and preferences.

02

Integrate natural language processing (NLP) tools to summarize key points and extract main insights from articles.

03

Provide related references and additional resources for deeper understanding and exploration of topics.

04

Offer multilingual translation capabilities to overcome language barriers and access information in different languages.

05

Design a user-friendly interface that facilitates easy navigation, search, and information retrieval.

Methodologies

1. **Data Acquisition:**
   - Utilised **newspaper3k** and **urllib3** to scrape articles from various news websites based on user queries or topics.
2. **Text Extraction and Processing:**
   - Extracted text from the acquired articles using **newspaper3k** for online articles and **pdfplumber** for PDF documents.
   - Preprocess the text data by removing noise, such as HTML tags or irrelevant content, using **beautifulsoup4** library.
   - Implemented language processing techniques using **langchain** and **google-generativeai** for keyword extraction and summarisation.
3. **Article Filtering and Summarization:**
   - Developed a language processing model using **langchain** and **google-generativeai** leverages NLP techniques to understand user inputs and match them with relevant articles.
   - Used summarisation algorithms from **langchain** and **google-generativeai** to generate concise summaries of the filtered articles.

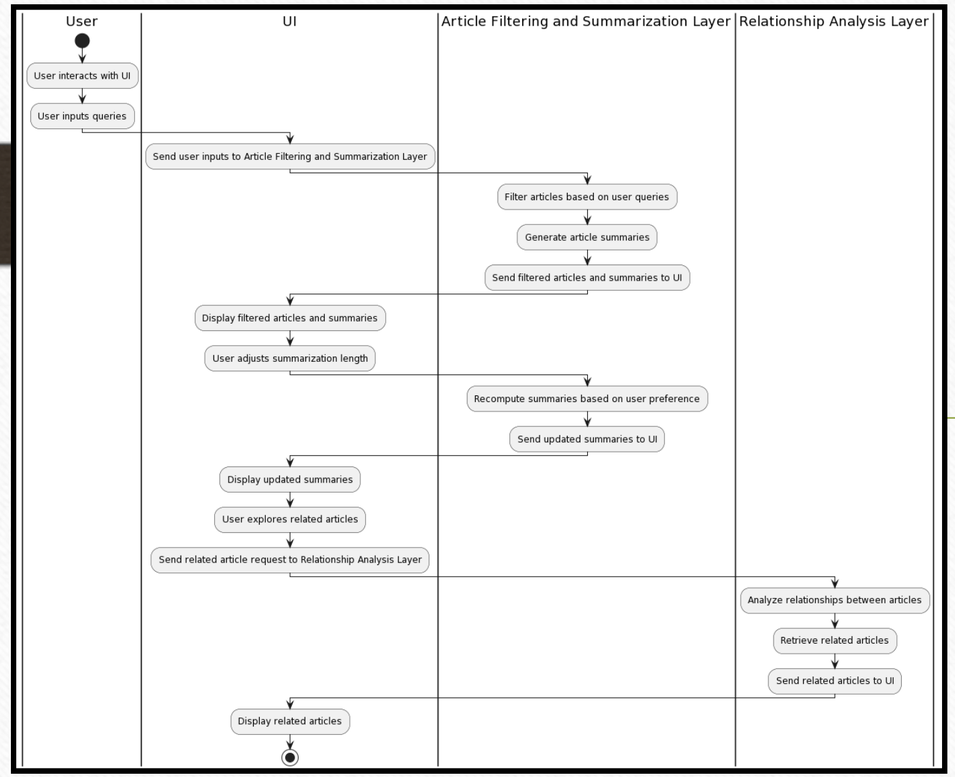
Methodologies

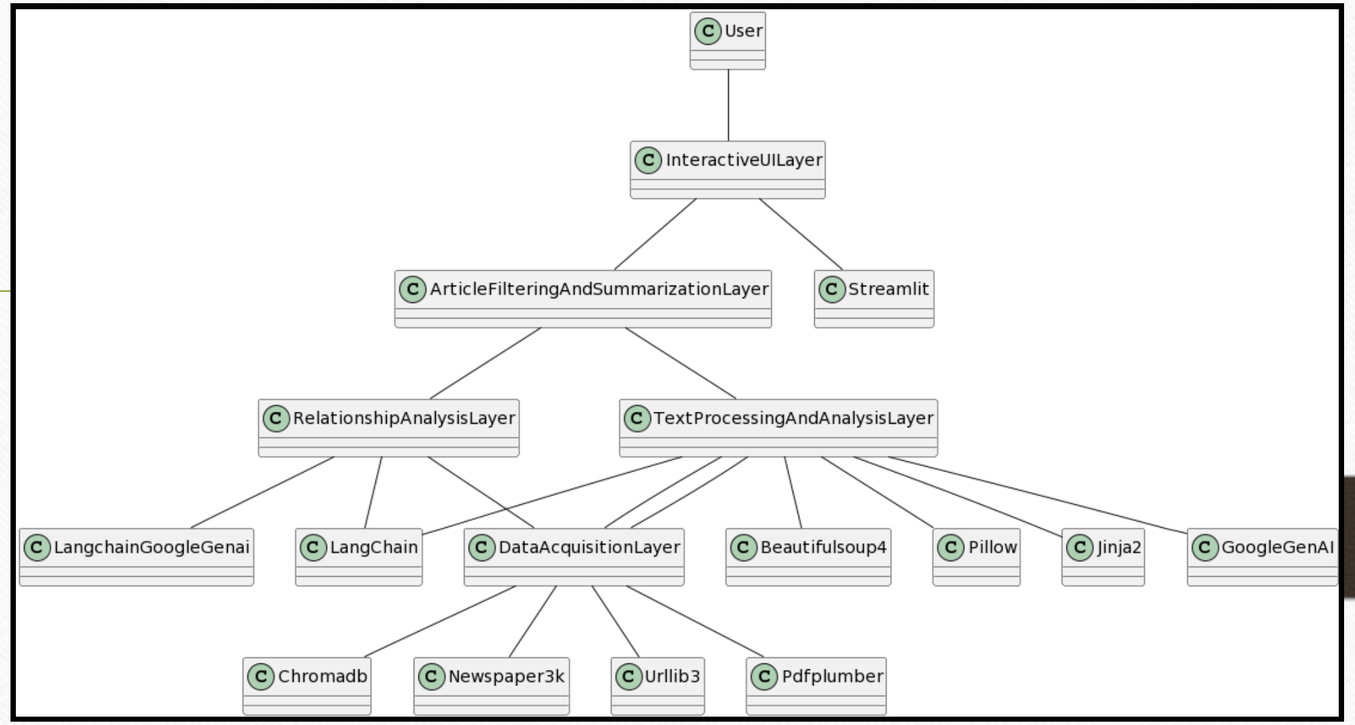
4. **Relationship and Language Analysis:**
   - Utilized **Chromadb** for storing and querying relationships between articles, enabling efficient retrieval and analysis.
   - Used **langchain-google-gena**i and **langchainhub** for additional language processing capabilities

5. **Interactive UI Development:**
   - Develop an interactive user interface using **streamlit**, leveraging its simplicity and flexibility for building web applications.
   - Incorporated features for users to input queries and adjust no of related articles using sliders or text inputs.

6. **Deployment and Integration:**
   - Utilized **streamlit** for deployment, allowing users to access the platform via a web browser without requiring additional installations.
   - Integrated **streamlit** with other components of the technology stack, such as the language processing model and relationship analysis modules, to provide a seamless user experience.

Work Flow





Architecture

Future Work

- Make the platform **accessible** to users with **diverse needs and abilities**.
- Analyze the **sentiment** of the article to provide users with an understanding of the **author's tone and potential biases**.
- Integrate **fact-checking mechanisms** to ensure the **accuracy and credibility** of the information presented in the articles.
- Enhance the **translation** capabilities to provide summaries and insights in a **broader range of languages.**

THANK YOU

TEAM - **QUEST.AI**
TEAM MEMBERS-

- **LAKSHYA RAJ VIJAY**
- **MAYANK PUVVALA**
- **DALAYI TEJA**

GITHUB LINK- InsightDigest

VIDEO LINK- LINK