AI tool for automated scanning and understanding of new scientific or technical posts and alerting about new information relevant to users' interests

Team members: Kyle Vinod, Katie Martinez, Ryan Ta, Alexander Larios | **Faculty adviser:** Tom Arodz, Ph.D. | **Sponsor:** VCU College of Engineering | **Mentor:** Tom Arodz

Overview and Background

With so many new scientific articles being published every year—over 2.5 million—it's nearly impossible for people to stay updated on what's important to them. On average, it takes around 17 years for major research findings to actually make an impact in real-world applications. This delay can hold back progress in areas like healthcare, engineering, and policy-making, where outdated information can lead to mistakes or missed opportunities. Current tools like Google Scholar and PubMed are helpful but overwhelming. They provide tons of information but don't filter or personalize it to individual needs, meaning users spend way too much time searching through irrelevant results. Our project aims to fix this with an AI tool that scans and understands new scientific or technical posts, then alerts users about the ones that match their interests. The tool uses AI to deliver personalized updates in real time, saving users time and effort while making it easier to stay informed. By simplifying complex research and tailoring it to what people care about, we hope this tool will help professionals make better decisions, spark innovation, and even encourage more collaboration across different fields.

Methods

For our testing, the main focus was to find how well the large language model (LLM) could interpret a set of interests and return scientific articles that best match those interests. Multiple LLMs were tested to determine which one is able to most accurately find similar scientific articles. We are interested in testing which LLM yielded the best results so it could be used in the final design. We chose to test some of the most popular LLMs currently: GPT-4, Gemini, Claude 3 Opus, and Llama 3. The set of interests provided were the same for all LLMs to keep testing consistent. We used arXiv and its API to access the scientific articles in its database. The database was kept the same for all the LLM designs to keep the access to information consistent between testing designs.

Analysis

ArXiv's API was used to retrieve multiple scientific articles for the LLM to analyze. The LLM was given a prompt containing the list of interests and instructions to find the scientific articles that match the best with the list of interests and return them. The LLM would go through a list of scientific articles and interpret each of their abstracts to determine which one is the closest match to the interests given. Once the list of closest matches were retrieved, they were reviewed by the tester to see how accurately the abstracts match with the interests. After review, modifications could be made on the LLM to obtain more accurate results that align with how the tester thinks the interests should be interpreted. Each of the LLM results were then compared with one another to see which one found the best matching articles.

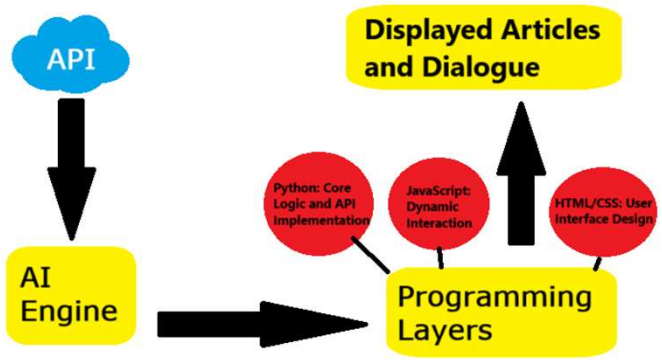

Figure 1: AI System Architecture

Results

- All APIs successfully connected to the AI design
- Each AI system fetched and displayed articles tailored to user queries
- Some APIs included unique dialogue, enhancing user interaction

Discussion

- AI systems demonstrated basic functionality:
  - Retrieved and formatted articles successfully
  - Generated user-specific dialogue
- The accuracy of article recommendations has not been evaluated yet
- Future plans include:
  - Implementing metric to measure recommendation accuracy
  - Refining the system for improved relevance and precision

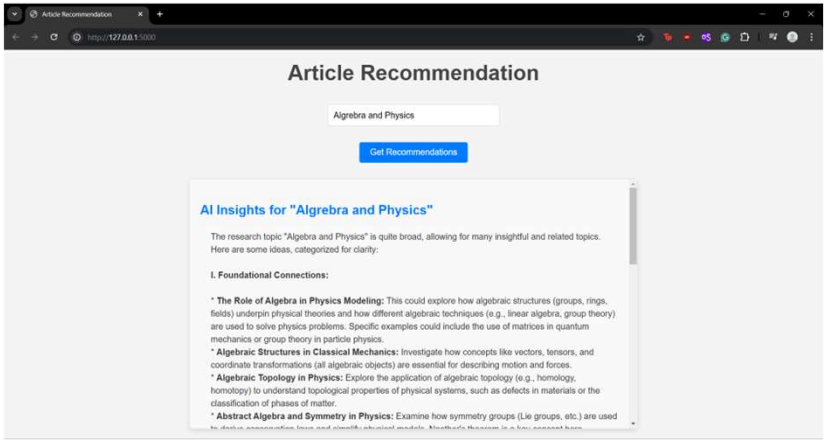

Figure 2: AI Output Example

VCU College of Engineering