vahan

Research Paper Summarization Multi-Agent System

Candidate Description

Build a multi-agent system that can find, analyze, and summarize research papers from various sources, organize them by topic, and generate audio podcasts discussing the findings.

Problem Statement: Staying updated on research across multiple fields is challenging and time-consuming. Create a system that can search for relevant research papers, process them, organize and summarize them by topic, and create accessible audio summaries.

Requirements:

1. Build a system that can:
    - Search for research articles based on topics with filtering options (relevance, recency, etc.)
    - Accept research papers from file uploads (PDF, academic papers)
    - Process papers from URLs to academic repositories
    - Handle DOI references
2. Implement topic classification based on user-provided topic list
3. Create a multi-agent system with specialized agents for:
    - Paper search and discovery
    - Paper processing and information extraction
    - Topic classification
    - Summary generation
    - Cross-paper synthesis
    - Audio generation
4. Generate both individual paper summaries and cross-paper topic syntheses
5. Create audio podcast versions of the summaries
6. Include a citation system for tracing information to sources

Expected Deliverables:

1. Working prototype with search, document processing, and summary generation
2. Source code with clear documentation
3. Sample papers and generated summaries for demonstration
4. README file explaining:
    - Setup instructions
    - System architecture
    - Multi-agent design and coordination approach
    - Paper processing methodology
    - Audio generation implementation
    - Limitations and future improvements

General Submission Guidelines

1. **Code Repository:** Submit your code through a GitHub/GitLab repository
2. **Documentation:** Include clear README and inline code documentation
3. **Setup Instructions:** The application should be able to set up and launch in no more than 2 commands - dockerization is recommended
4. **Demo:** Include screenshots or a brief video demonstration if possible
5. **Time Management:** The assignment is designed to be completed in approximately 12 hours
6. **Focus Areas:** Prioritize core functionality over visual polish or edge cases
7. **Technology Choices:** You may use technologies of your choice but be prepared to justify your selections
8. **Sample input output :** Add sample input and output in the submissions

Remember that this assignment is meant to showcase your:

- Problem-solving approach
- Engineering implementation skills
- LLM integration capabilities
- Software design decisions
- Technical communication

We value working solutions with thoughtful architecture over perfect implementations. Good luck!