**Syllabus Chatbot using Retrieval Augmented Generation (RAG)**

This project will provide students with an innovative way to interact with course syllabi using natural languages. By integrating Retrieval Augmented Generation (RAG) techniques, students will be able to ask questions about course syllabi in a conversational manner. Below are the key requirements and features of the system.

**Requirements:**

1. **Web-Based System:** The system will be a user-friendly web platform that allows easy access for both students and professors.
2. **Syllabus Upload by Professors:** Professors will have the ability to upload their syllabi in PDF format. During the upload process, they will be prompted to specify important metadata such as “department,” “course number,” and “course name” to ensure accurate categorization.
3. **Restricted Upload Access:** Only professors can upload syllabi to the system, and they must log in to do so. The system will not permit unauthorized users to upload documents.
4. **Student Access:** Students can search for a course by selecting the appropriate department and course number or name, and ask questions related to that syllabus. No login is required for students, making the system accessible to everyone.
5. **Choice of Language Models:**

* **In-House Open-Source Model:** You can use the open-source large language model deployed in my lab. While this model may be slower, it is free to use.

**Additional Considerations:**

* **API:** The teams will elect liaisons to design an API that allows your chatbot to communicate with the LLM. Attached please find a sample, which only considers "prompt" and "max\_tokens". You will need to decide what else you will need to control the behavior the LLM.
* **Search Optimization:** Ensure the system can accurately retrieve syllabus information based on both specific and vague queries.
* **User Experience:** Prioritize clear navigation and response accuracy to enhance the student’s learning experience.
* **Data Privacy:** Include measures to protect the privacy and integrity of the uploaded syllabi, ensuring that only authorized individuals can modify them.

**Tools to Use for Design and Development**

* You must use some UML CASE tools to make your analysis and design, such as StarUML and IBM Rational Rose.
* You must use Microsoft Project to work out your workplan and use MS Project to keep track of the progress of your project.
* You may use Microsoft Visio to draw Entity-Relation diagrams (ERDs), and data flow diagrams (DFDs).
* You must use some Unit Test tools (e.g., Pytest, JUnit, CppUnit, etc.) to perform formal testing.