Technical Assessment Assignment: Develop a RAG-based AI Application that answers questions about yourself!

Objective:

Develop a Retrieval-Augmented Generation (RAG) based AI system capable of answering questions about yourself. The system should handle inquiries in English and manage follow-up questions effectively.

Core Requirements:

Data Ingestion:

- Implement functionality to ingest personal data from various sources such as text files or PDFs. Examples of personal data include the applicant's resume, personal website content, etc.
- The ingested data should then be indexed into a basic vector database (e.g., ChromeDB) to facilitate efficient retrieval.

RAG System Integration:

- Utilize a Retrieval-Augmented Generation (RAG) approach for generating answers. This involves combining the retrieval of relevant information from the indexed database with the generative capabilities of large language models (LLMs) to produce accurate and contextually relevant responses.

Orchestration:

- You can use langchain or llamaindex for orchestrating the interaction between the AI components, including data retrieval from the vector database and the generative model's input/output processing. You can also choose to build your own functionality if keen!

Follow-Up Questions:

- Implement logic to handle follow-up questions, maintaining the context of the conversation to provide coherent and contextually appropriate responses.

Interface Options:

- You are free to choose between developing API endpoints or integrating a basic frontend (like Gradio or Streamlit) for user interaction with the AI system.

Model and Tools Flexibility:

- Applicants can use any large language models, including but not limited to OpenAI's models or other open-source options.

Deliverables:

- Code Repository:
    - Submit a well-organized code repository containing all the source code developed for this project. Include clear instructions for setting up and running the application.
- Documentation:
    - Provide documentation covering the system architecture, data ingestion process, RAG integration, possible methods of evaluating the quality of the application, and any other relevant aspects of your solution.
- Demonstration:
    - Include a video or live demonstration link showcasing the system's capabilities, specifically highlighting the handling of follow-up questions, quality of question answering etc.

Evaluation Criteria:

- Functionality: Accuracy and relevance of the AI system's responses, handling of follow-up questions.
- Innovation: Creativity in system design, integration of components, and problem-solving approaches.
- Code Quality: Clarity, organization, and documentation of the codebase.
- Usability: User experience design and ease of interaction with the AI system, via API or Gradio frontend.

This assignment offers an opportunity to showcase your expertise in AI, machine learning, NLP, and your ability to build complex, innovative systems. We look forward to reviewing your solution and understanding how you approached this challenge.