SE/ML technical challenge

Problem:

We have a collection of internal documents, and we want to implement a document-based GPT system that can answer questions based on their content. The primary goal is to enable users to ask questions about these documents and receive short, precise, and well-sourced answers. Each answer should include a link to the specific part of the document where the information was found, allowing users to verify the information directly.

System Requirements:

- Users should receive answers that include a link to the relevant document and the specific paragraph where the information was found. This helps ensure the answers can be easily verified.
- The system should only answer questions that are relevant to the content of the documents. If a question is not relevant, the system should inform the user that the question is not applicable to the documents.
- The system should be designed to allow easy updates with new documents, ensuring the content remains current and comprehensive.
- Implement guardrails to prevent the system from returning inappropriate or irrelevant answers.

Submission Guidelines

- Provide a solution that meets the requirements of this challenge. The implementation should be in Python and expose a REST API for interaction. **Bonus points** if it comes with a simple frontend interface to interact with the API.
- Ensure the solution is production-ready, adhering to best practices for software development and deployment.
- The solution should be containerized.
- Include a README file with detailed instructions on how to run the solution.
- Provide a small technical report explaining the approach taken, the challenges faced, and the solutions implemented.
- Push the complete solution to a private GitHub repository and share access with the user `baddoub`.

Evaluation Criteria

- The quality of the code, adherence to best practices.
- The accuracy of the answers provided by the system.
- The design and functionality of the REST API.
- The clarity and completeness of the provided documentation.