**AI Chatbot Project Documentation**

**1. Mục tiêu chính (Main Objective)**

**Phát triển một chatbot có thể xử lý các truy vấn của người dùng về tài liệu đính kèm, phục vụ cho mục đích học tập (learning) và đọc hiểu tài liệu (document comprehension).**

**2. Chức năng chính (Key Functions)**

* **Trả lời câu hỏi về tài liệu:** Cung cấp câu trả lời chính xác, phù hợp dựa trên tài liệu đính kèm do người dùng cung cấp.
* **Tóm tắt tài liệu:** Xử lý các văn bản dài (PDF, DOCX, TXT…) và cung cấp bản tóm tắt dễ hiểu, giúp người dùng đọc nhanh và nắm bắt thông tin cốt lõi.
* **Phục vụ học tập & nghiên cứu:** Hỗ trợ người dùng học tập, nghiên cứu hiệu quả hơn thông qua việc hỏi đáp và tóm tắt tài liệu.

**3. Các điểm nổi bật của sản phẩm (Selling Points)**

| **Điểm nổi bật** | **Ý nghĩa/Tác dụng** |
| --- | --- |
| Seamless AI Crawling | Tìm kiếm nhanh chóng trong các tài liệu do người dùng cung cấp. |
| Context-Aware Responses | AI hiểu và trả lời theo ngữ cảnh, tăng độ chính xác. |
| Efficiency | Đáp ứng tức thì mọi câu hỏi, tăng hiệu quả học tập và nghiên cứu. |
| Accuracy | Câu trả lời chính xác, không phụ thuộc định dạng đầu vào. |
| Time-Saving | Tự động hóa trả lời, giảm thao tác thủ công, tiết kiệm thời gian cho người dùng. |

**4. Danh sách công cụ và giấy phép sử dụng (Tools & Licenses)**

| **Chức năng mở rộng** | **Công cụ / Công nghệ chính** | **Giấy phép sử dụng** |
| --- | --- | --- |
| Tích hợp Knowledge Base | S3, Lambda, Kendra, AI Model (Hugging Face, OpenAI) | AWS License, OpenAI API License |
| Context-Based Guardrails | OpenAI models | OpenAI API License |
| UI Refinement | React, Material UI | N/A |
| AI Integration | Python, OpenAI, Hugging Face | OpenAI API License |
| Database Integration | PostgreSQL | AWS/GCP Cloud Provider License |
| Authentication & Authorization | AWS Cognito | AWS Cognito License |

**5. Kiến trúc hệ thống (System Architecture)**

**Các thành phần chính:**

* **Frontend:**
  + Người dùng nhập câu hỏi và tải file tài liệu.
  + Hiển thị câu trả lời, tóm tắt và lịch sử hỏi đáp.
* **Backend:**
  + Lưu file và lịch sử hội thoại.
  + Kiểm tra định dạng, dung lượng file (File Validator).
  + Trích xuất và xử lý nội dung file (PDF, DOCX, TXT).
  + Tiền xử lý văn bản gửi lên AI.
* **API Layer:**
  + Giao tiếp với Azure OpenAI để xử lý câu hỏi và nội dung tài liệu.
* **Storage Layer:**
  + Lưu trữ lịch sử hỏi đáp, file đã upload và dữ liệu liên quan.

**6. Tính năng nổi bật & Đề xuất cải tiến**

**6.1. Tính năng nổi bật mới cập nhật**

1. **Tối ưu UI/UX**
   * Thêm hướng dẫn sử dụng rõ ràng trên giao diện.
   * Hiển thị progress bar khi upload file để nâng cao trải nghiệm người dùng.
2. **Xác thực dữ liệu đầu vào**
   * Kiểm tra định dạng file (PDF, DOCX, TXT), dung lượng file (≤10MB).
   * Đảm bảo nội dung câu hỏi hợp lệ trước khi gửi đi xử lý.
3. **Hiển thị & tìm kiếm lịch sử hỏi đáp**
   * Cho phép người dùng xem lại lịch sử hỏi đáp theo từng file.
   * Tìm kiếm và lọc lịch sử theo từ khóa hoặc tên file.
4. **Tóm tắt nội dung file trước khi hỏi**
   * Sau khi upload, hệ thống tự động tóm tắt nội dung tài liệu và hiển thị cho người dùng kiểm tra trước khi đặt câu hỏi.
5. **Hỗ trợ đa ngôn ngữ**
   * Giao diện và câu trả lời chatbot hỗ trợ nhiều ngôn ngữ, phù hợp với người dùng quốc tế, học tập hoặc môi trường đa văn hóa.

**7. Tổng kết**

Dự án xây dựng chatbot AI hỗ trợ người dùng khai thác tri thức từ tài liệu đính kèm, tối ưu cho mục đích học tập, nghiên cứu và đọc hiểu. Hệ thống sử dụng các công nghệ AI hiện đại, tích hợp bảo mật, hiệu năng cao, giao diện thân thiện, đáp ứng nhu cầu thực tiễn của cá nhân và doanh nghiệp. Các tính năng mới giúp tăng trải nghiệm, đảm bảo độ chính xác, tiết kiệm thời gian và đáp ứng linh hoạt các kịch bản sử dụng thực tế.