1. **Multi-Language Support:**

* Add multilingual embedding models to support documents and queries in non-English languages.

2. **Scalability:**

* Replace pickle storage with a distributed vector database like Pinecone, Weaviate, or Elasticsearch for scalable indexing and retrieval.

3. **OCR Integration:**

* Include OCR tools (e.g., Tesseract) for processing scanned documents.

4. **Real-Time Collaboration:**

* Allow simultaneous user uploads and queries by integrating session-based storage (e.g., Redis).