

  ChatGPT 4o 

acting as a business systems analyst, write detailed user requirements for the following high level business requirements:

Create a web application for Document Loading and Retrieval-Augmented Generation (RAG)

The application should have a user-friendly interface designed to load documents from a pre-defined folder called "data", process these documents to extract their content, including text contained within images within documents, and utilize a local LLM to create a Retrieval-Augmented Generation (RAG) solution. The main functionalities of the application include:

Document Loading:

At launch, the system will check to see if any files in the data folder are new or have changed. If there are files that are new or have changed, they are loaded into the persistent vector DB. Otherwise, the application loads normally. Documents in the data folder include PDF, PDF with images, TXT, CSV, DOC, DOCX, PPT, PPTX, HTM, and HTML.

When extracting content from files, text is split into manageable chunks and embedded in a vector store, which allows for efficient similarity searches.



Message ChatGPT

ChatGPT can make mistakes. Check important info.

