

Retrieval Augmented Generation (RAG)



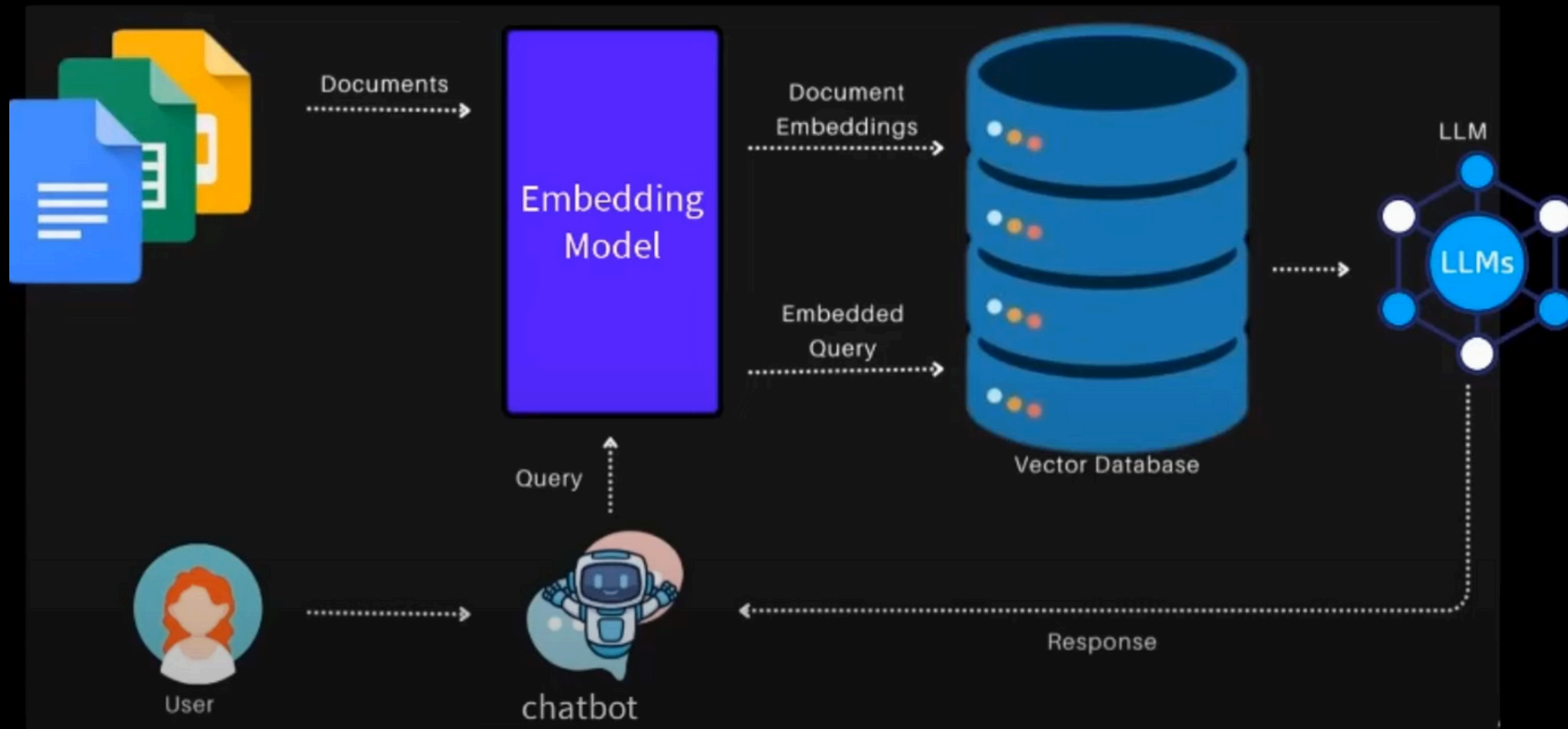
Retrieval Augmented Generation (RAG)

Limitation of LLMs:

- LLMs can't answer correctly to private data
- LLMs can't provide the most current information

Retrieval Augmented Generation (RAG)

RAG allows LLMs to use external sources for better responses



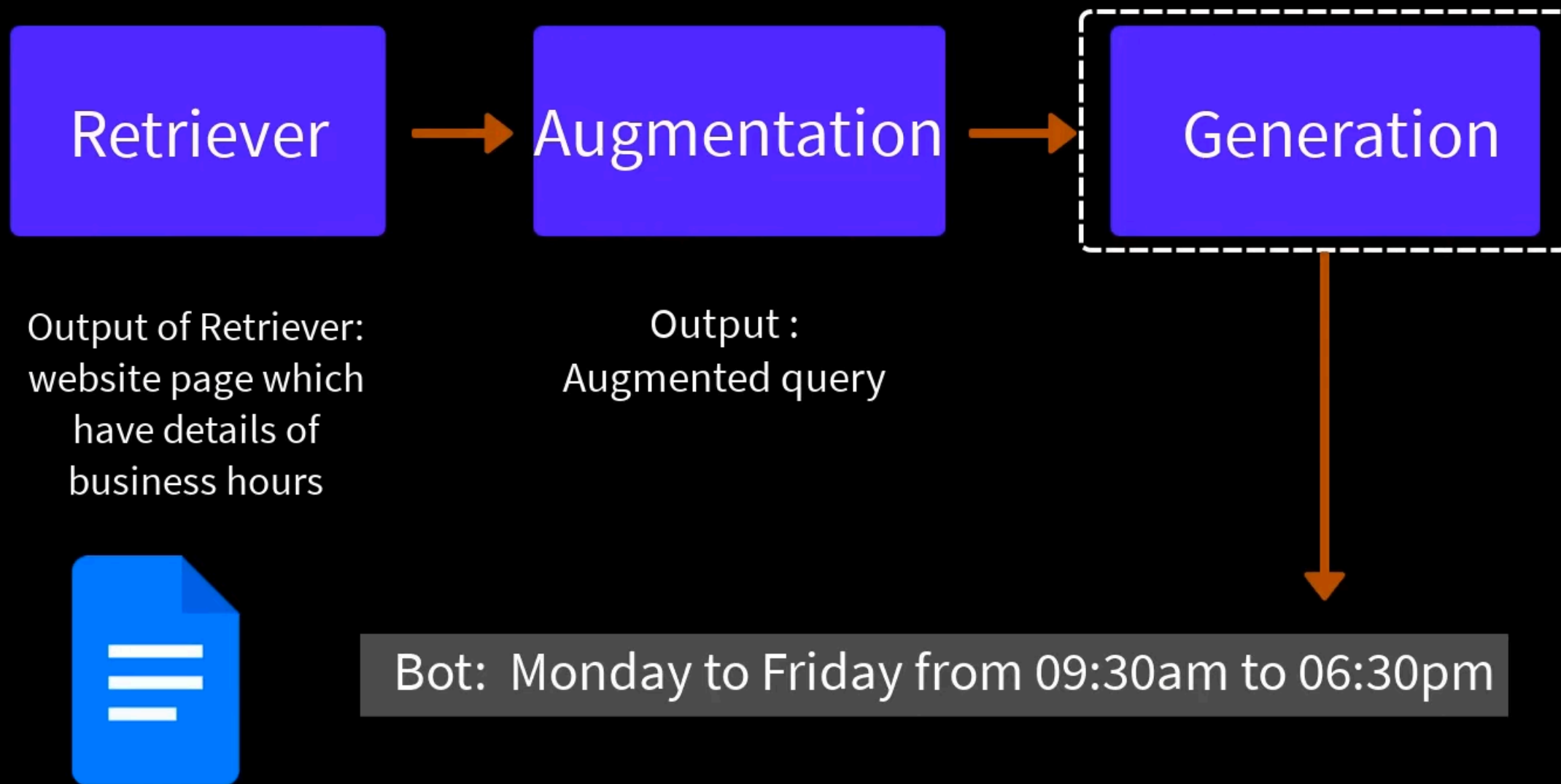
Retrieval Augmented Generation (RAG)

RAG allows LLMs to use external sources for better responses

- More accurate ans on private data
- Stay updated with new information

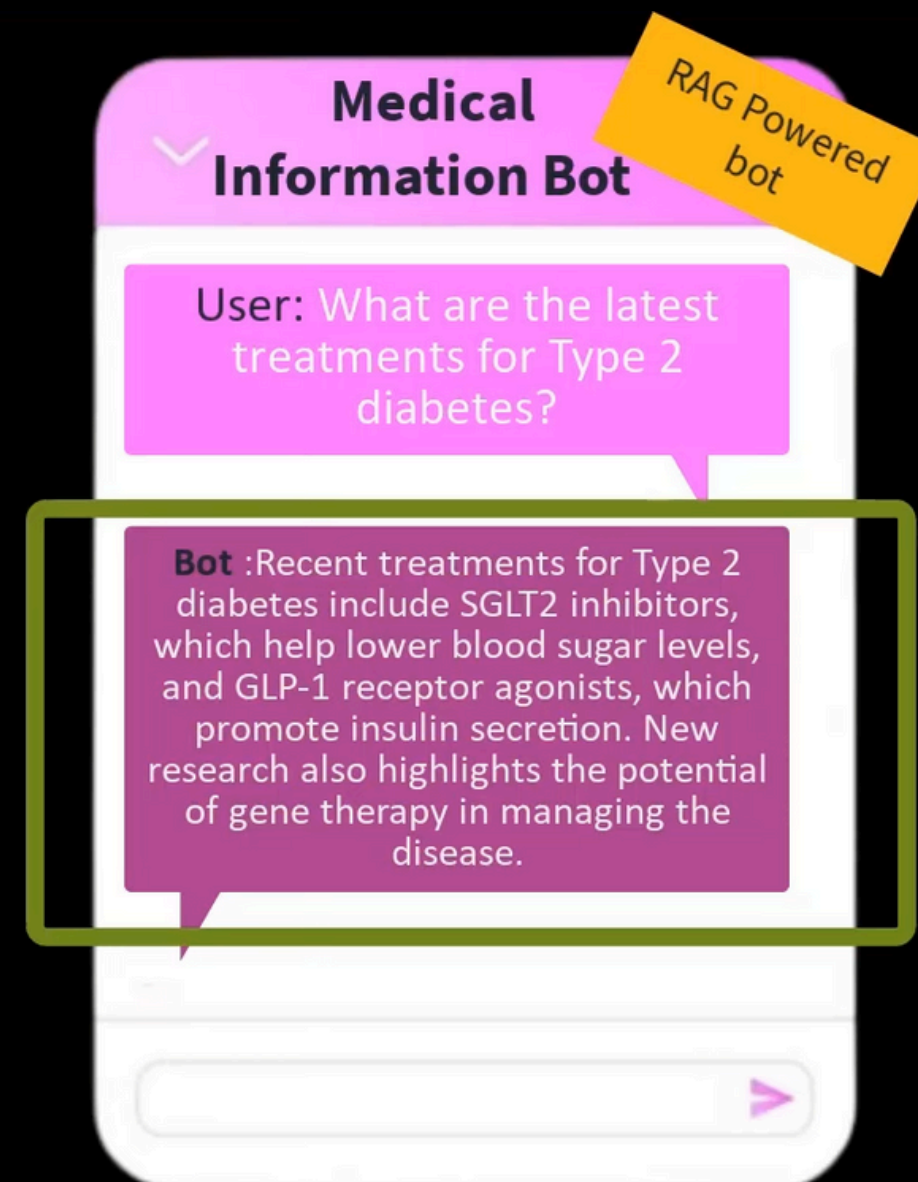
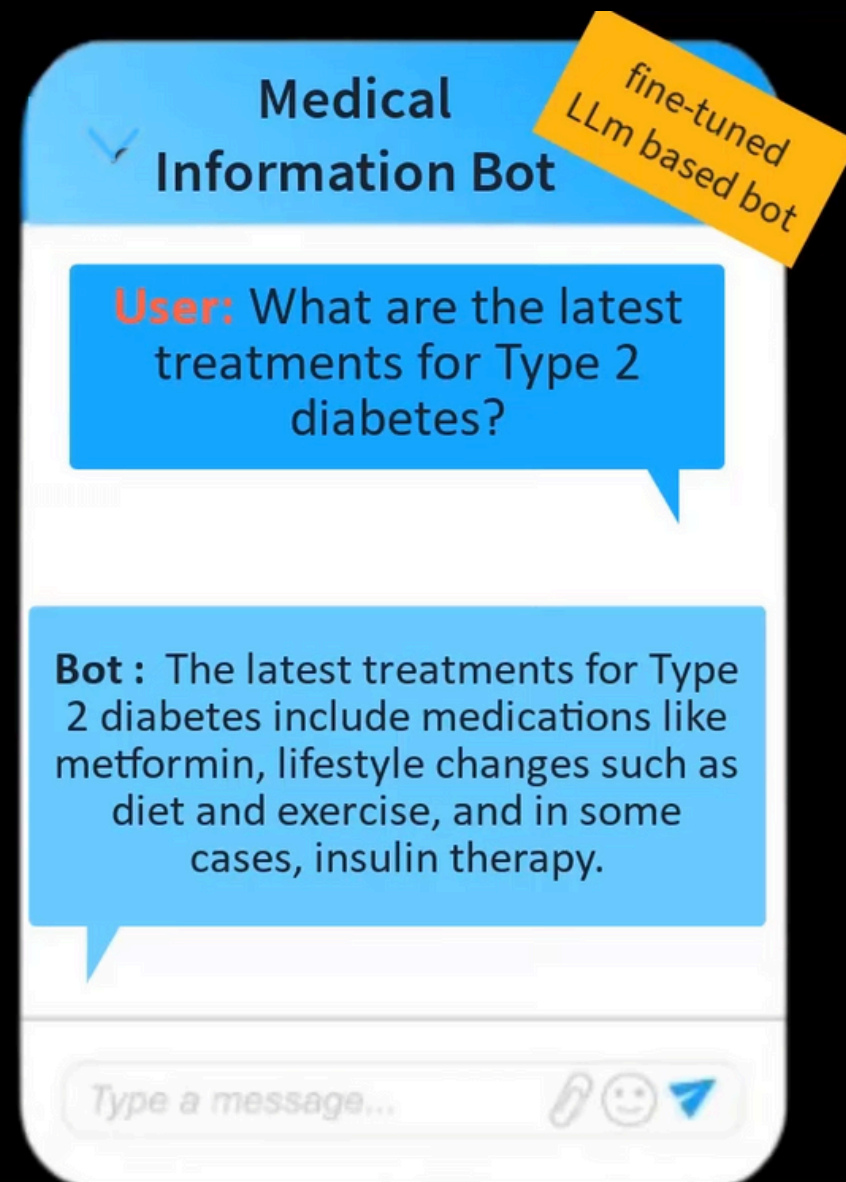
Retrieval Augmented Generation (RAG)

User: What are your business hours?



RAG vs. Fine-tuning

Both extend the abilities of language models



RAG Projects

LangChain provides document loaders for CSVs, file directories, markdown, HTML, JSON, and PDFs. This example uses the Python PDF loader.

