

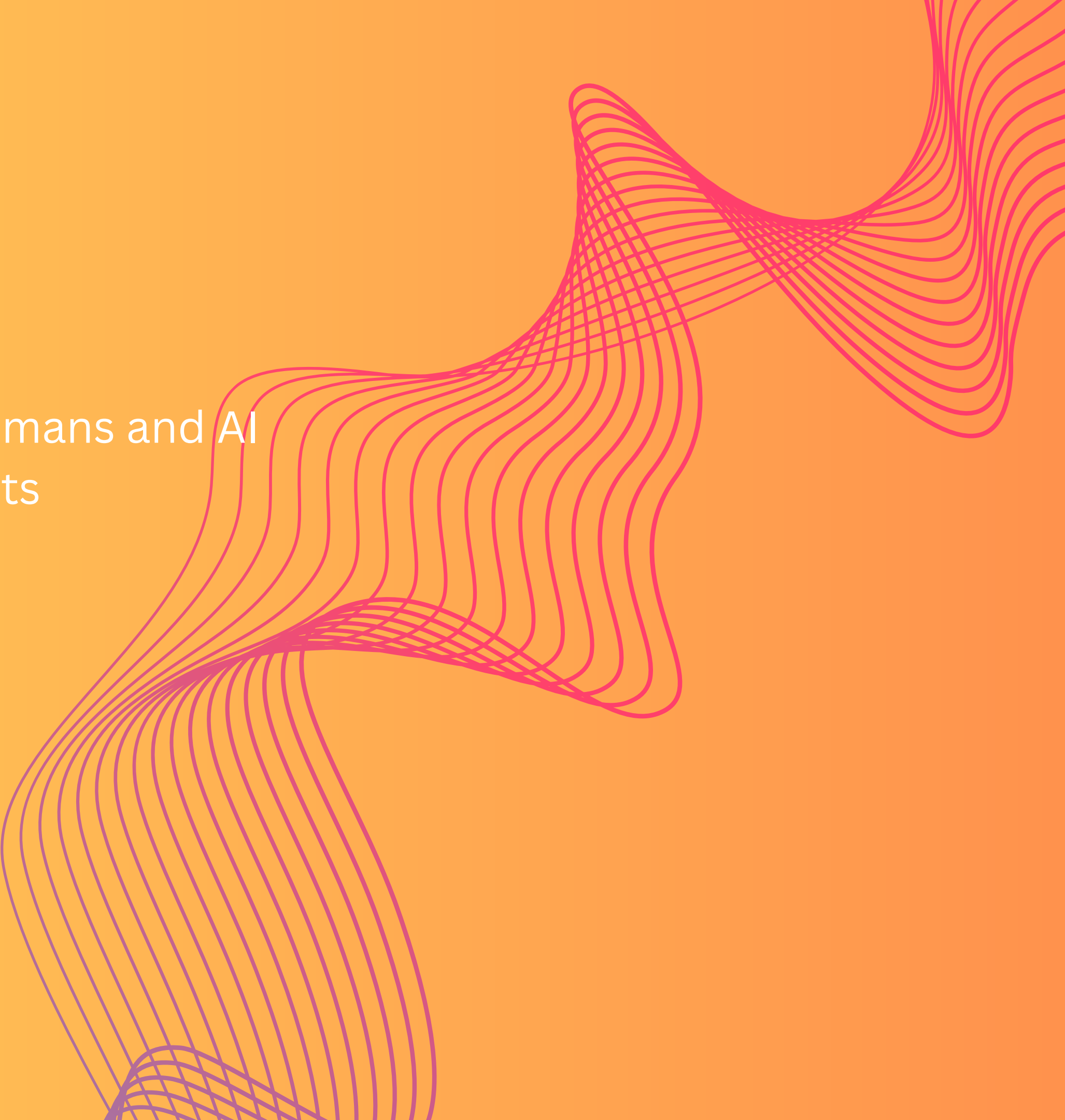
Prompt Engineering



Prompt Engineering:-

Prompt engineering is the process of structuring an instruction that can be interpreted and understood by a generative AI model.

Prompt Engineering:-

- write, refine and optimize prompts
 - perfect the interaction between humans and AI
 - continuously monitor those prompts
 - maintain an up-to-date prompt
- 
- The background of the slide features a series of overlapping, wavy lines in shades of pink and purple. These lines originate from the right side and flow towards the left, creating a sense of movement and depth. The lines are thin and closely spaced, forming a complex, organic pattern that complements the warm orange background.

Why Prompt Engineering ?

Lets see one demo:

Prompt-1:

Correct my paragraph: “Today was great day in the world for me I went Disneyland with my mom. It coukd have been better if it wasnt raining.”

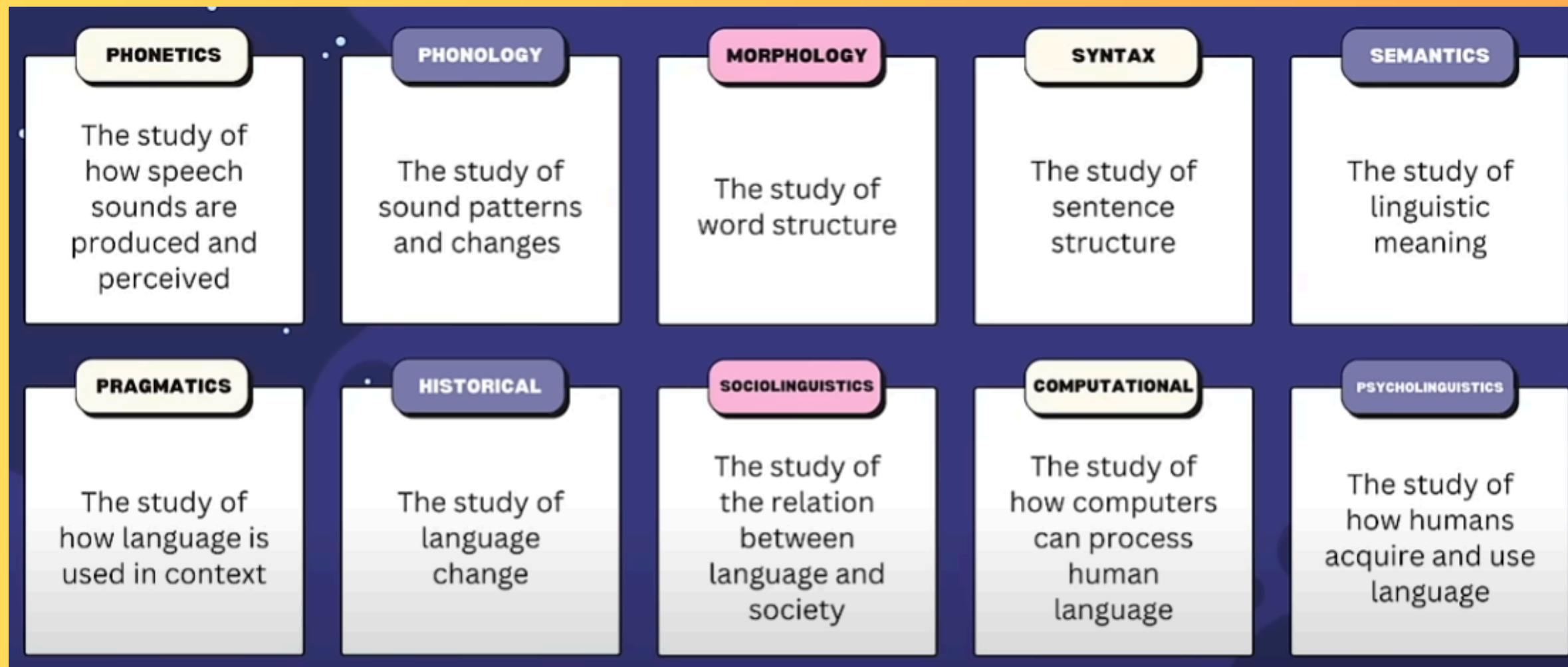
Why Prompt Engineering ?

Prompt-2:

"I want you to act as a spoken English teacher. I will speak to you in English and you will reply to me in English to practice my spoken English. I want you to keep your reply neat, limiting the reply to 100 words. I want you to strictly correct my grammar mistakes and typos. I want you to ask me a question in your reply. Now let's start practicing, you could ask me a question first. Remember, I want you to strictly correct my grammar mistakes, typos and factual errors"

Linguistics:-

Linguistics is the study of language



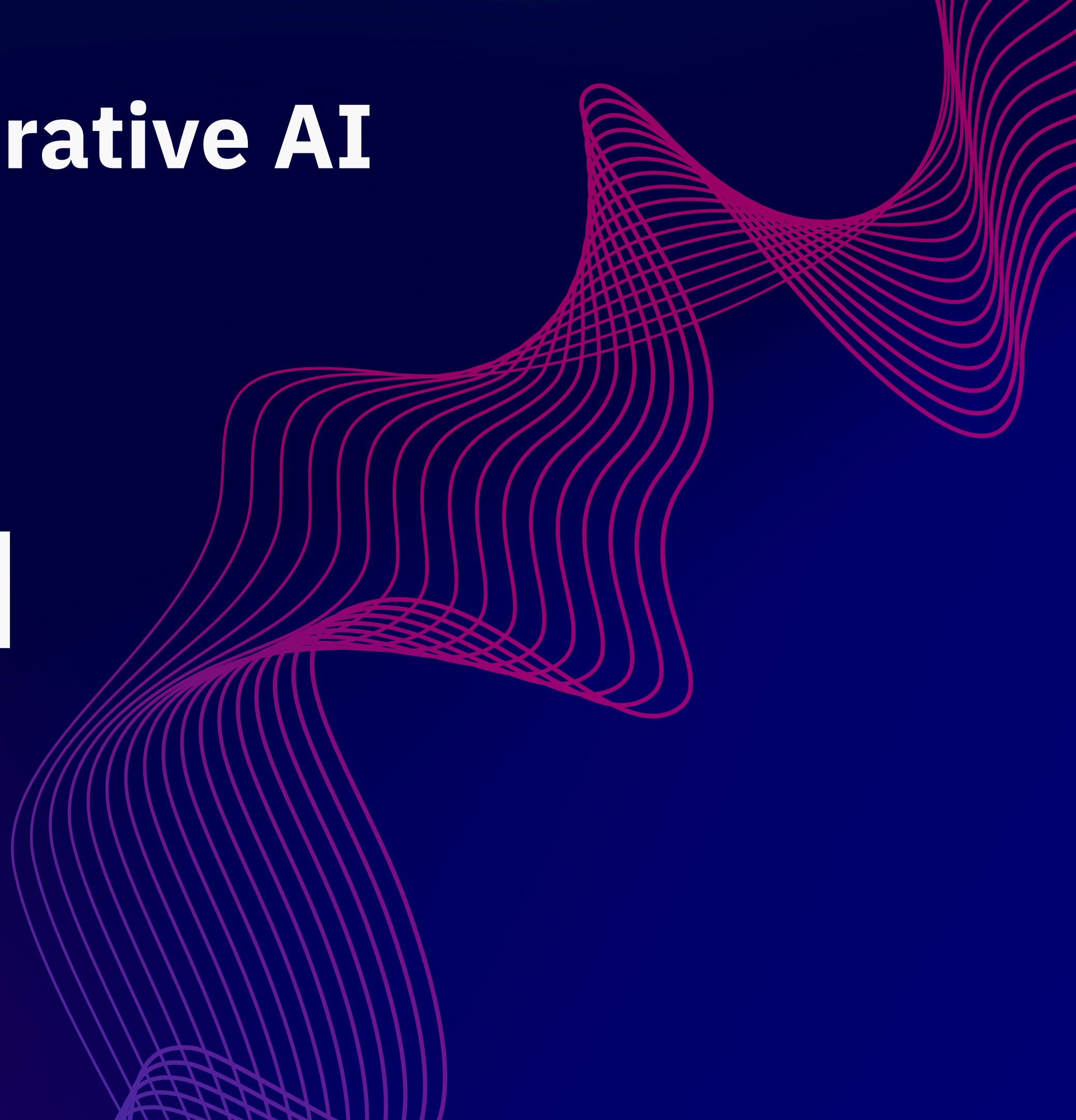
Generative AI Model (LLMs)

Model	Vendor/Creator
GPT(GPT-3, GPT-4, etc.)	OpenAI
BERT	Google
LaMDA	Google
PaLM	Google
BLOOM	Hugging Face
LLaMA	Meta
Claude	Anthropic
NeMO LLM	Nvidia
Generate	Cohere



Complete Generative AI

Retrieval Augmented Generation (RAG)



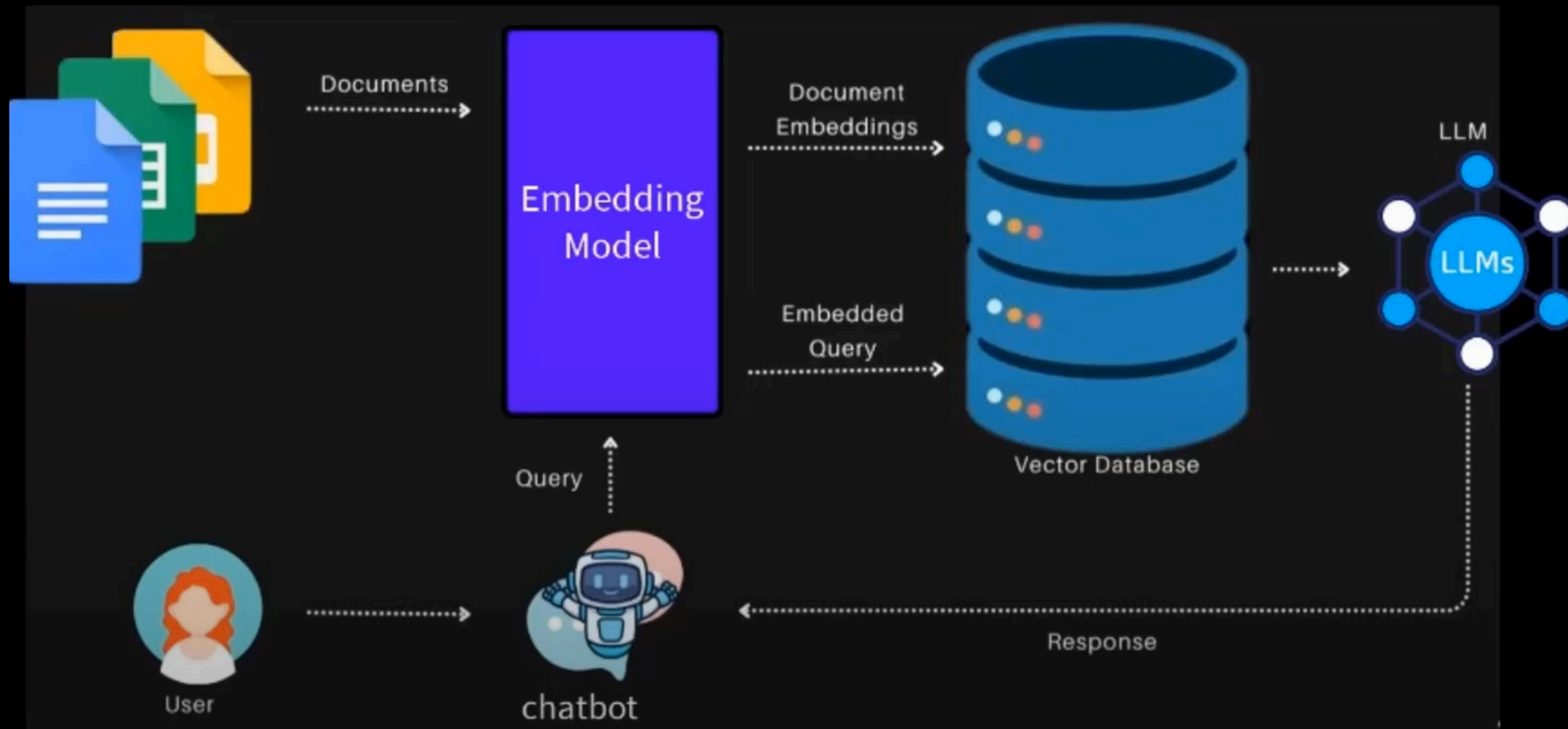
Retrieval Augmented Generation (RAG)

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



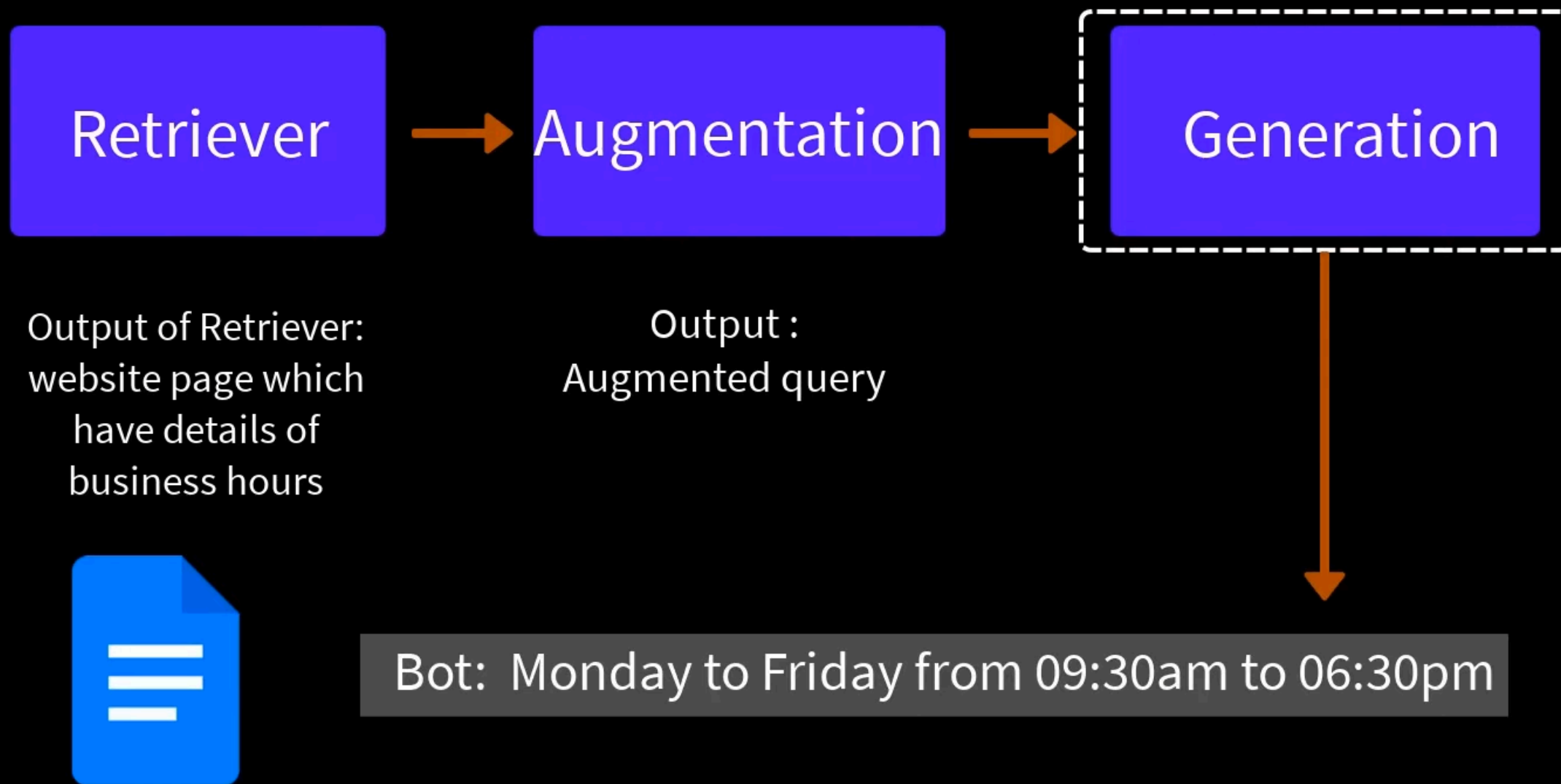
Retrieval Augmented Generation (RAG)

RAG allows LLMs to use external sources for better responses

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

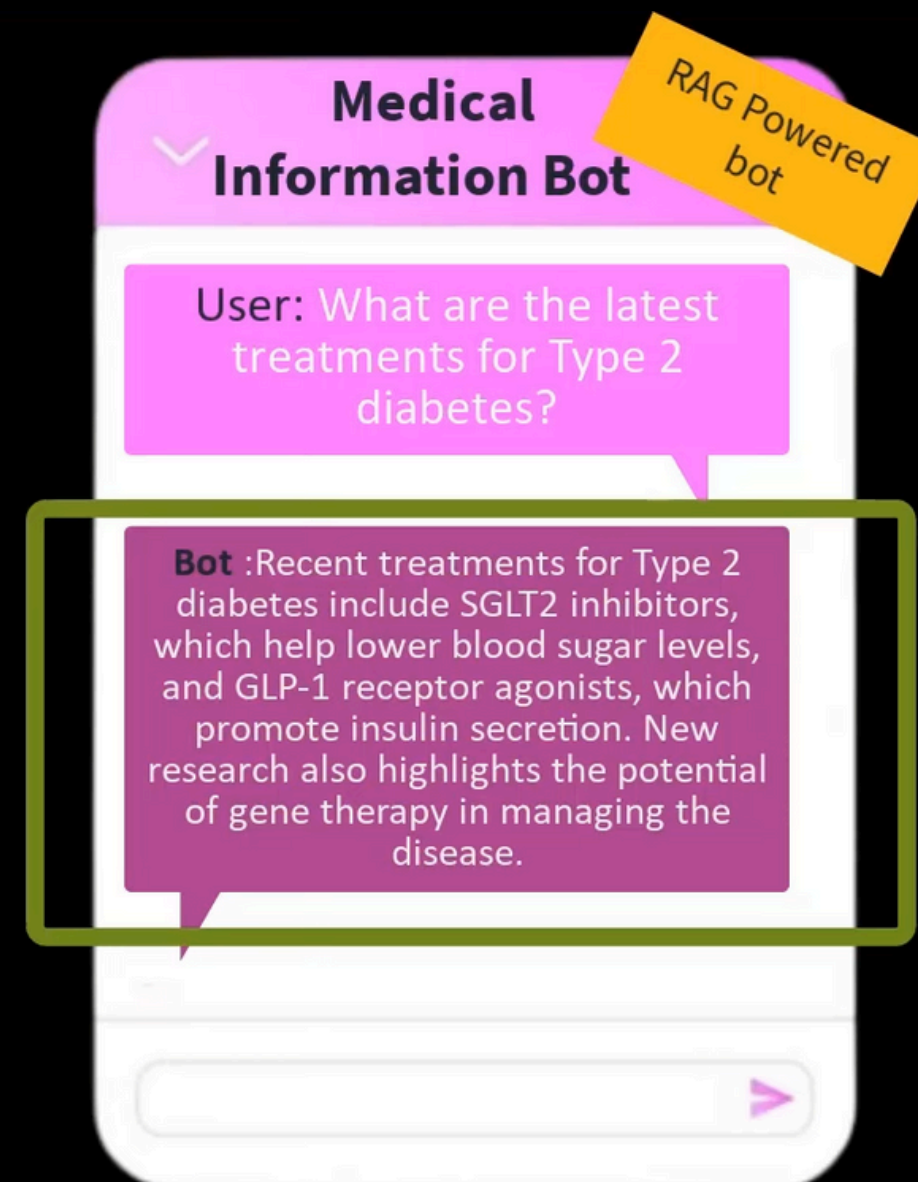
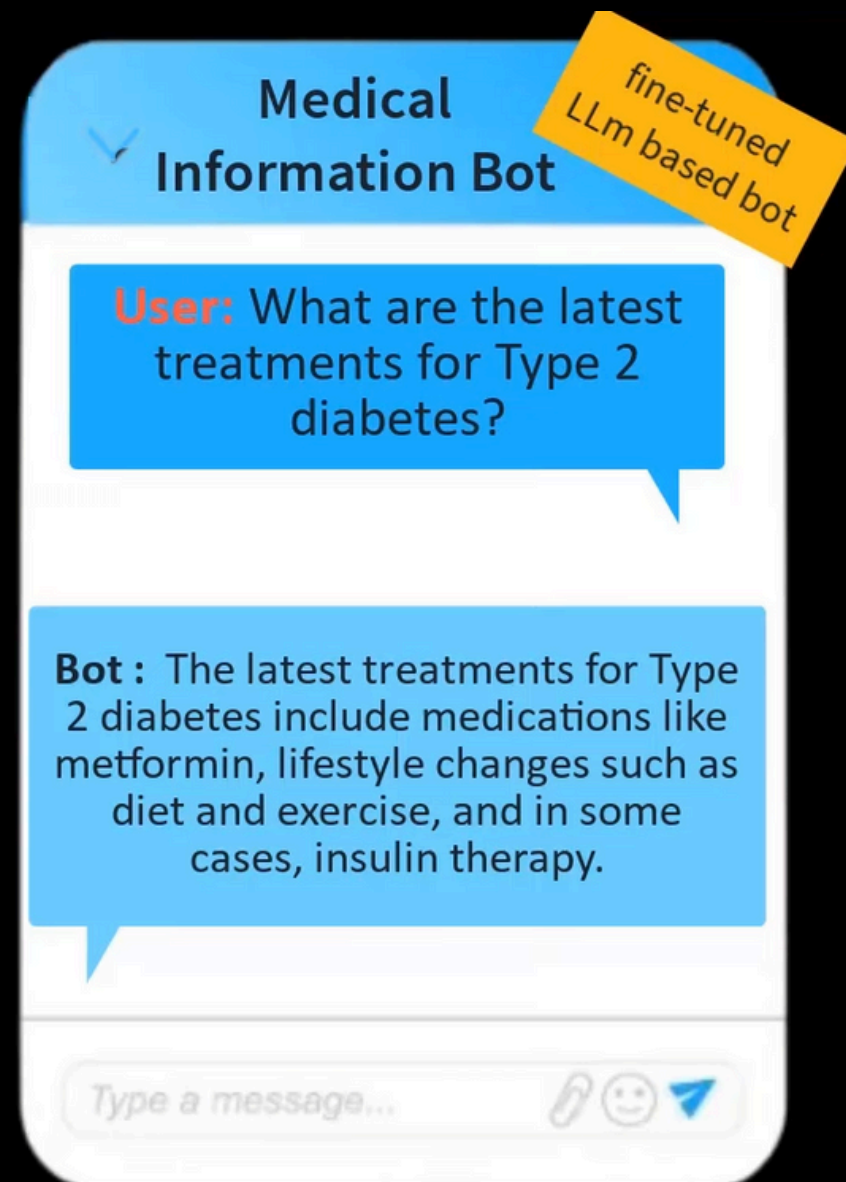
Retrieval Augmented Generation (RAG)

User: What are your business hours?



RAG vs. Fine-tuning

Both extend the abilities of language models



RAG Projects

Build a Q&A App with RAG using Gemini Pro and Langchain

