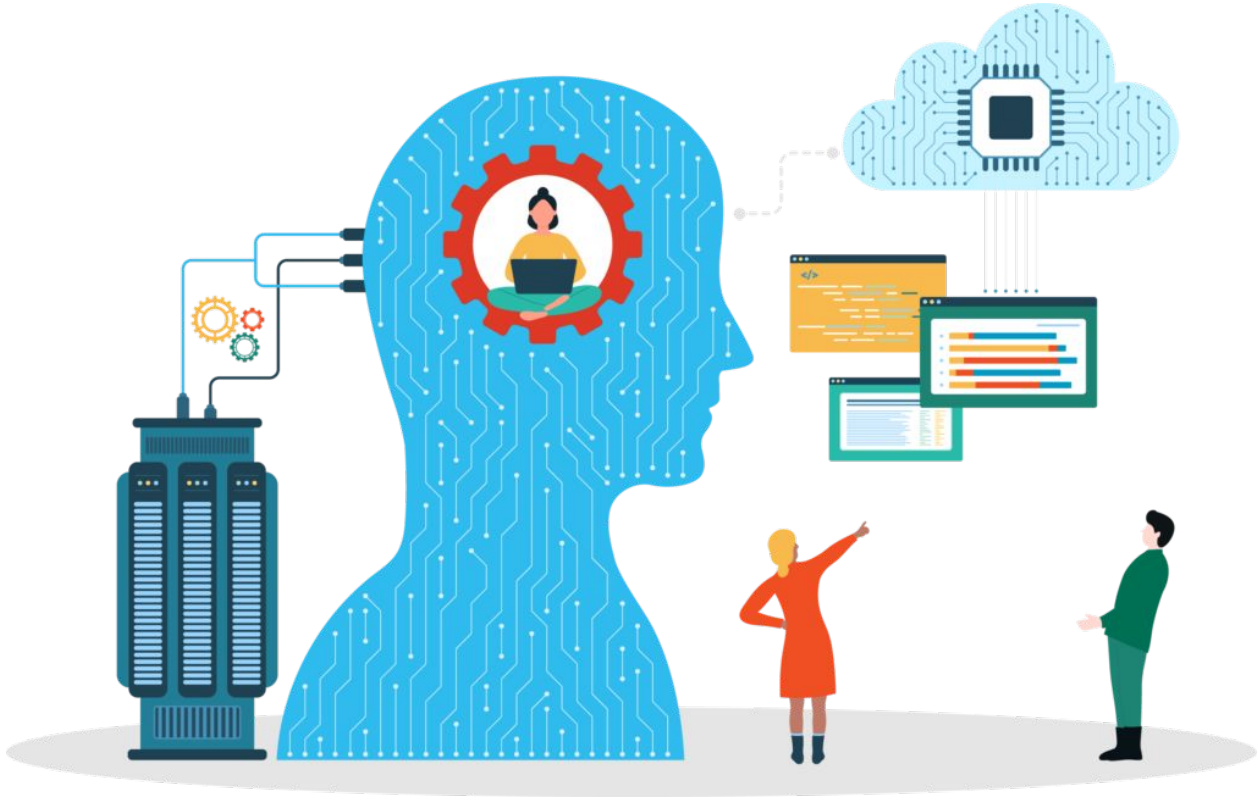
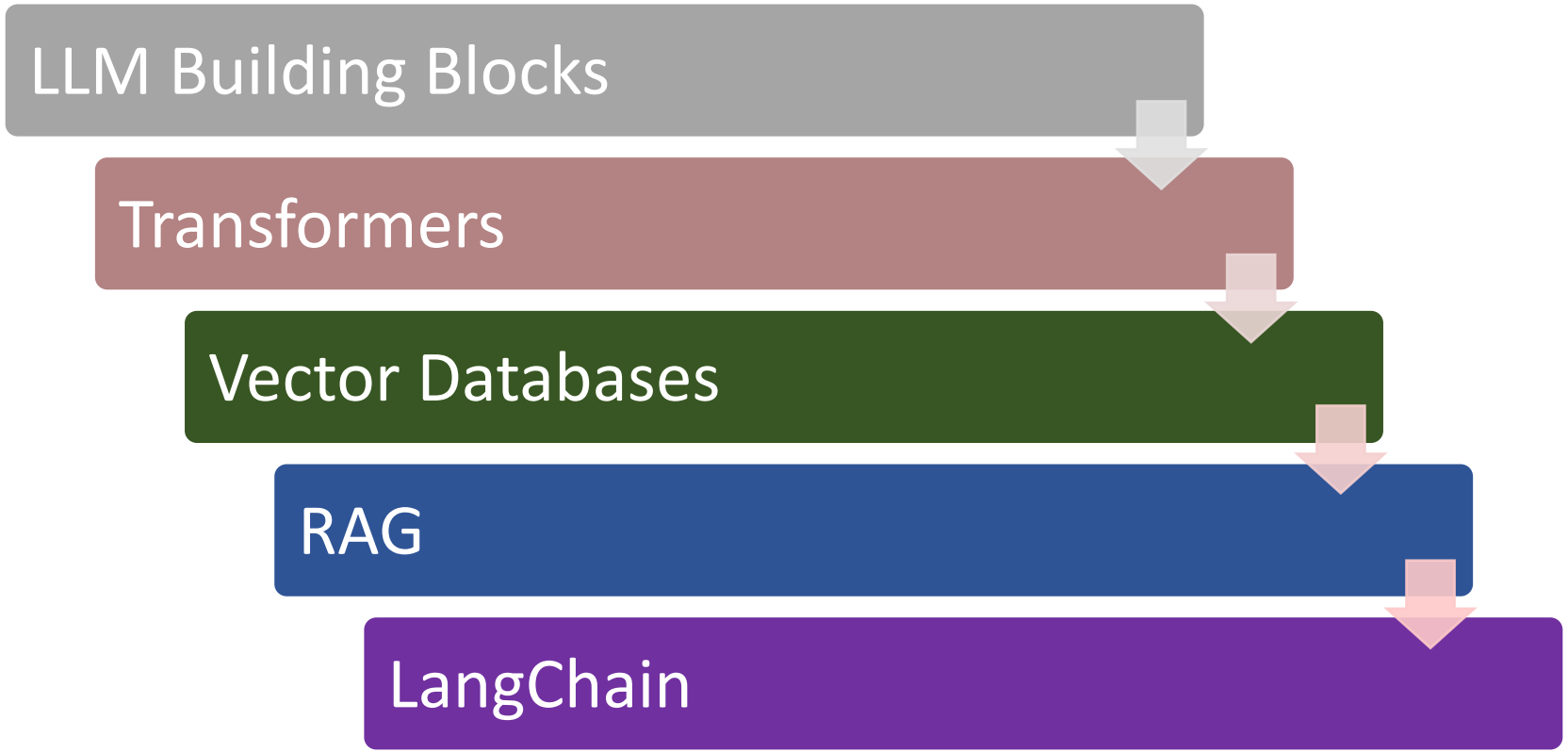
Agenda

- ✔ Generative AI Core Concepts
  - LLM Building Blocks
  - Transformers
  - RAG
  - Vector Databases
  - LangChain
- ✔ Workshop Overview
- ✔ Let us build something live!
- ✔ Follow ups

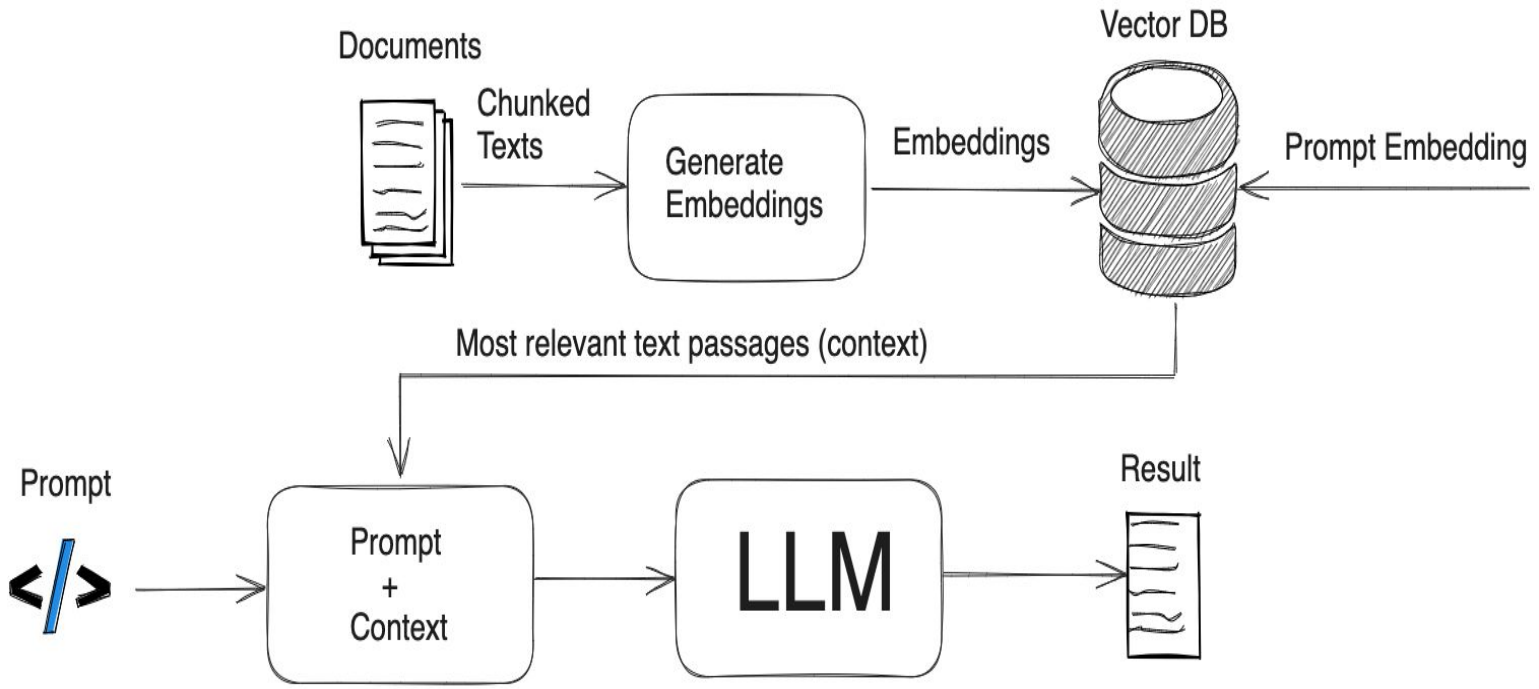


Generative AI - Core Concepts



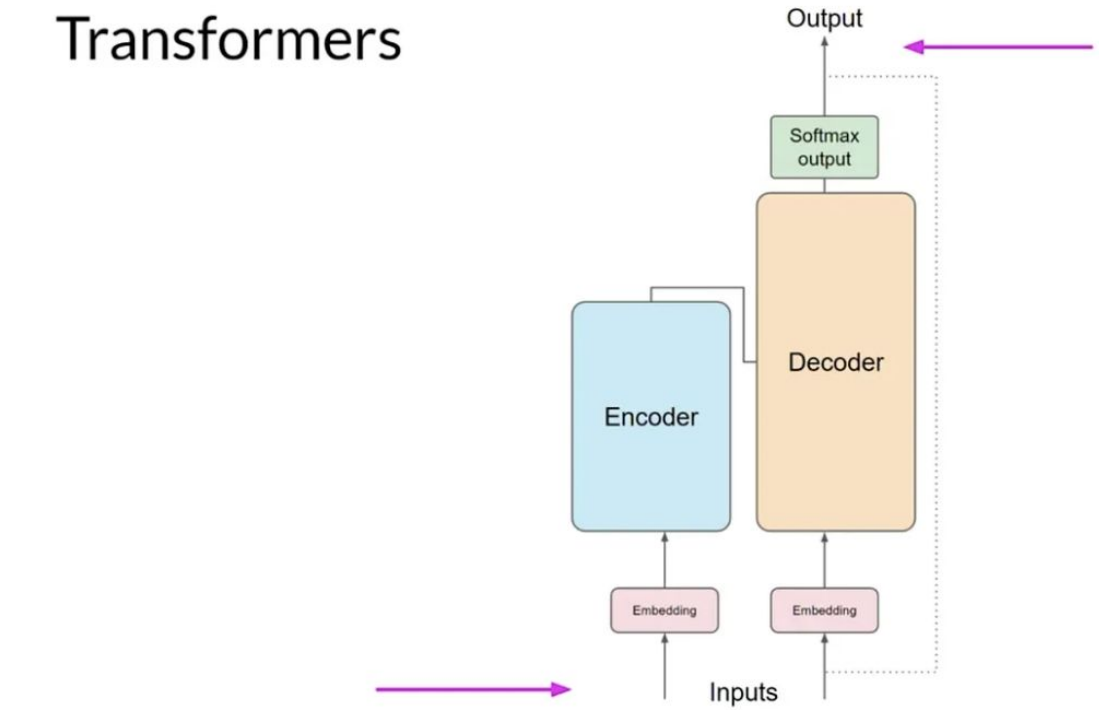
LLM Building Blocks

- ✔ Tokens
- ✔ Vectors
- ✔ Embeddings



Transformers

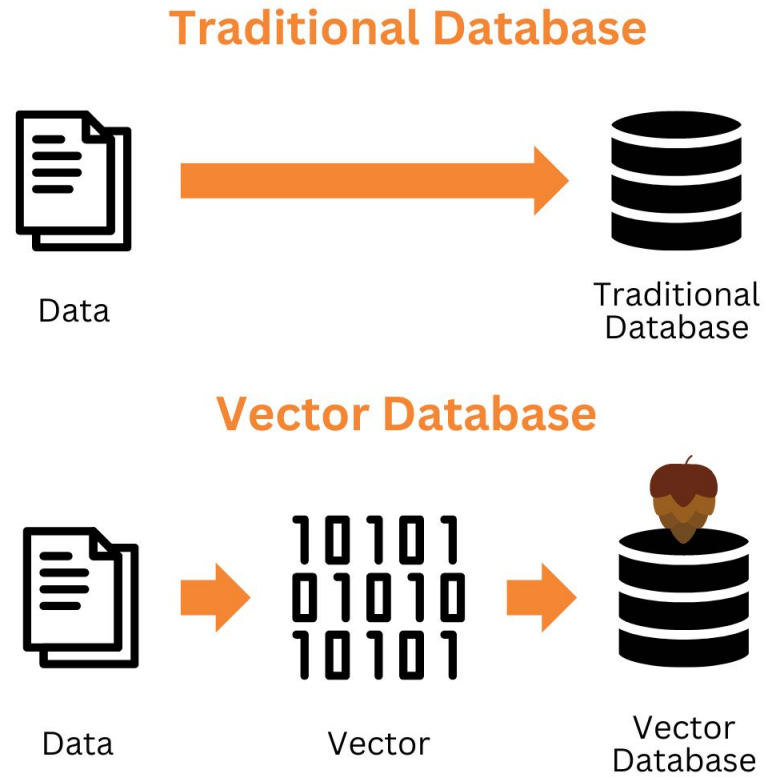
- ✔ Positional Encodings
- ✔ Attention
- ✔ Self-Attention



* Image fromhttps://python.plainenglish.io/a-basic-high-level-view-of-transformer-architecture-llms-from-35000-feet-c036dd2f7c25

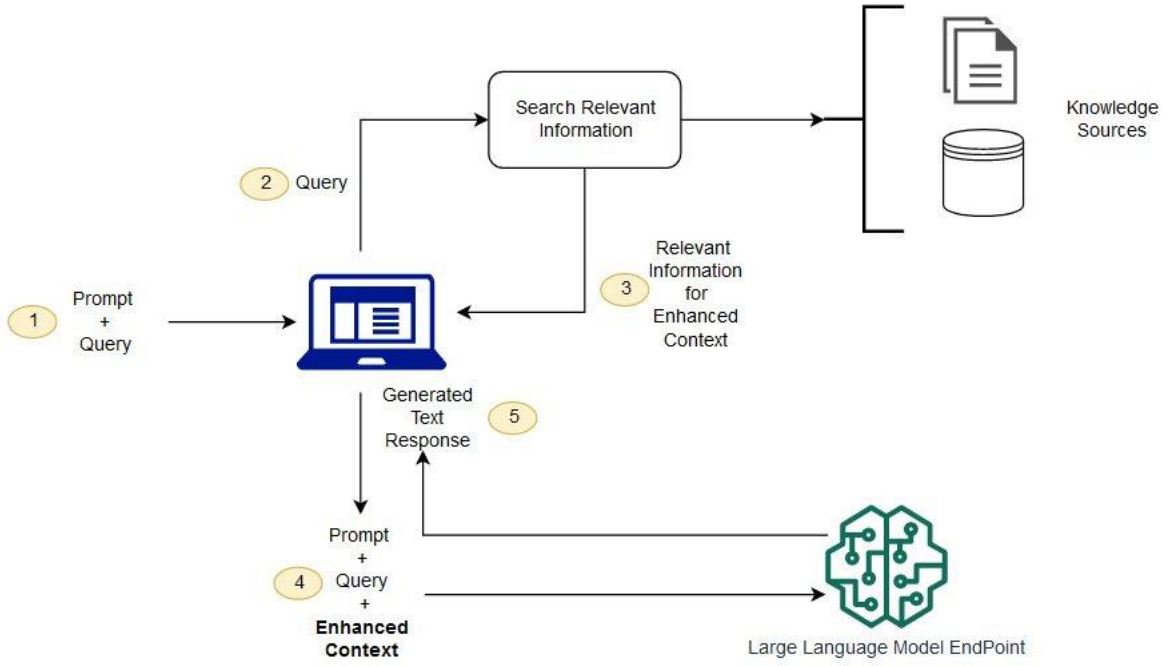
Vector Database

- ✔ What are Embeddings
- ✔ How to store
- ✔ Why use Embeddings?



Retrieval Augmented Generation AKA RAG

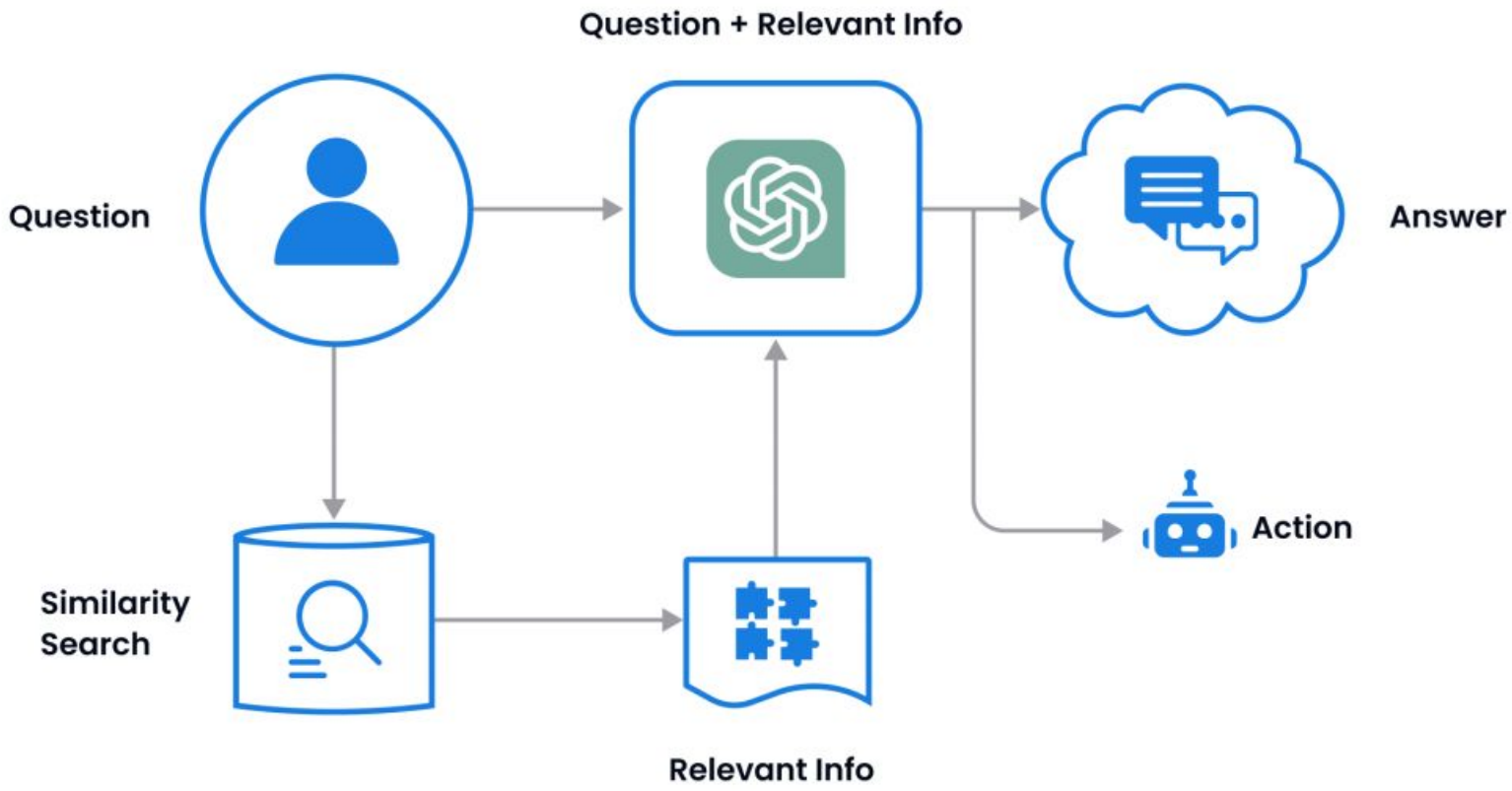
- ✔ Custom Knowledge Base
- ✔ Vector Databases
- ✔ Retrieval Component
- ✔ Generation Component
- ✔ RAG-LLM Combination
- ✔ RAG-LLM Benefits



* Image from https://docs.aws.amazon.com/sagemaker/latest/dg/jumpstart-foundation-models-customize-rag.html

LangChain

- ✔ Templates
- ✔ Agents
- ✔ The Chains



Workshop Overview

- ✔ Overview: We will build a Q and A application using LLM, RAG and Langchain
- ✔ Git Repo : https://github.com/coder-lgtm/langchain-demo OR git@github.com:coder-lgtm/langchain-demo.git
- ✔ Software and tools used:
  - Python - Python 3.8
    - Open AI API - https://openai.com/api
  - Streamlit Python Library - https://streamlit.io/
  - LangChain Python Library - https://python.langchain.com/docs/get_started/introduction/

- ✔ What do you need?
  1. If you want to build this app live then a Laptop with internet connection. ***This is optional***. You can always try this at a later time per your convenience.
  2. Please follow Setup Instructions here https://github.com/coder-lgtm/langchain-demo beforehand.

Let's Build Something LIVE!!

THANK YOU!

Questions??