♯ Agendas:

✓① Comprehensive tool Stack for Gen AI Application

✓② Different Orchestration Frameworks

✓③ RAG High Level concepts

✓④ How to choose best Orchestration Framework
   for proffessional level LLM App Dev.

① External API ⟶ Tavily, SerpAPI, Weather API

② LLMs ⟶ GPT, Mistral, Llama, Falcon, Groq API

③ Cloud Storage ⟶ S3, GCP Bucket, Azure Blob

④ Orchestration Framework ⟶ Langchain, LLamaIndex, openAI API

⑤ Vector DataBase ⟶ Chroma DB, FAISS, Pinecone, weaviate
   Redis, MongoDB Atlas

⑥ Backend Server ⟶ Python, Langchain, Flask, FastAPI

⑦ Frontend Server ⟶ Flask, Fast API, Django, Next.js, React.

⑧ Cloud ⟶ AWS, GCP, Azure

   ✓⟶ AWS Bedrock

   ✓⟶ GCP vertexAI

   ⟶ Azure openAI ✗

2023 — October

2024 ⟶

What is Feature

Questions ———→ [LLM] ———→ Response
what is pattern   GPT

→ KB

= === =
= ———
— —.

[RAG] ← [90%]

Features ←
[batch]
↙ 10 ᵐ
① Langchain
② LlamaIndex ← (PA) → ↗
③ OpenAI — SDK → GPT

Tena iPs

universal OF

LLMs → GPT ———→ OPEN AI API ←
Llama ——→ llama cPP ←
Gemini → Google Gen A2 ←
(Falcon —→ ←
mistal —→ ←

Langchain —— LLM

→ Data loader
→ Load LLM
→ Vector DB
→ 3rd API
→ Data privacy
→ eg t
— —

→ Langsmith → monitoring (Lo
→ Langserve → model serving
→ Langgraph → Agents / multi