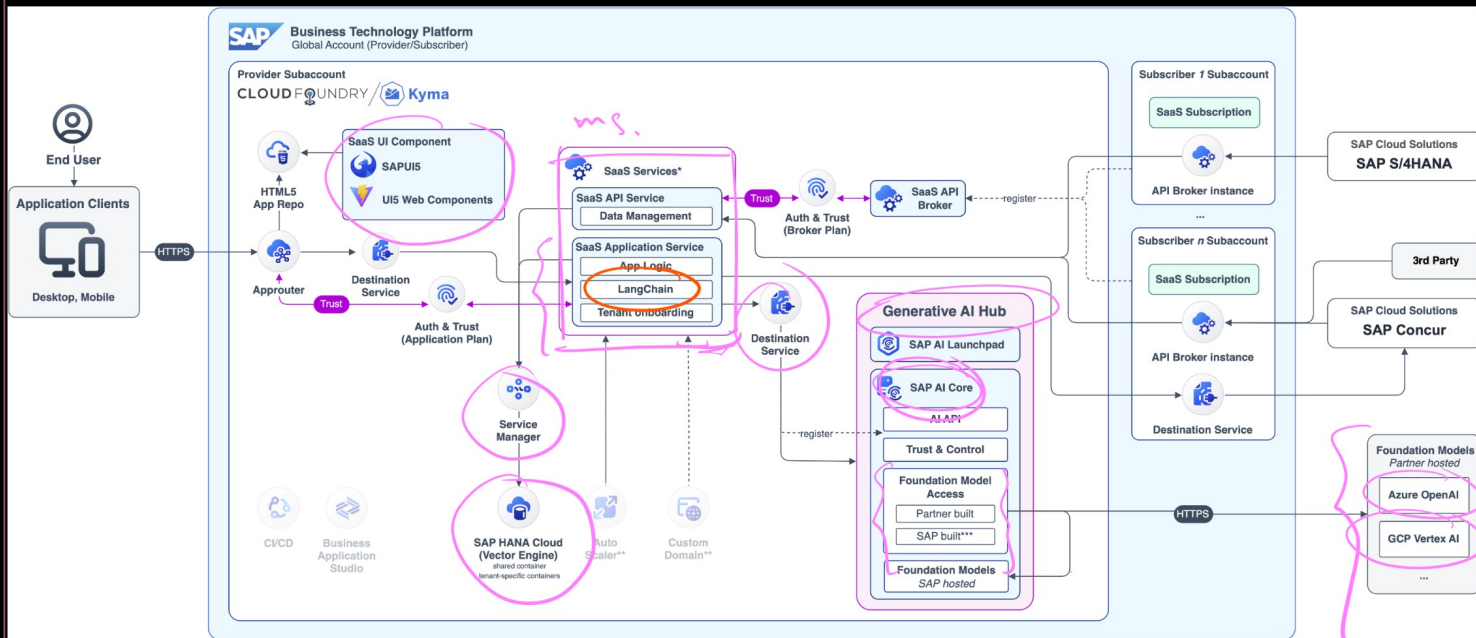
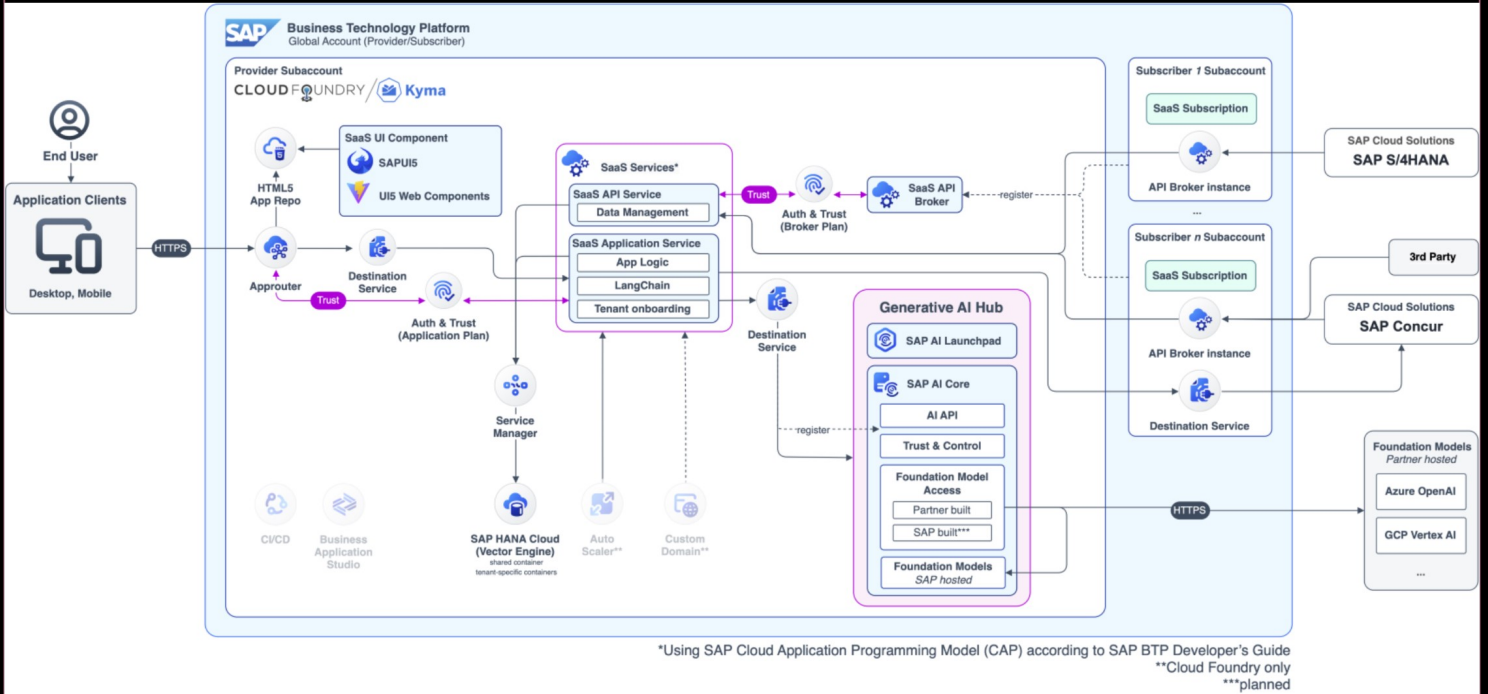


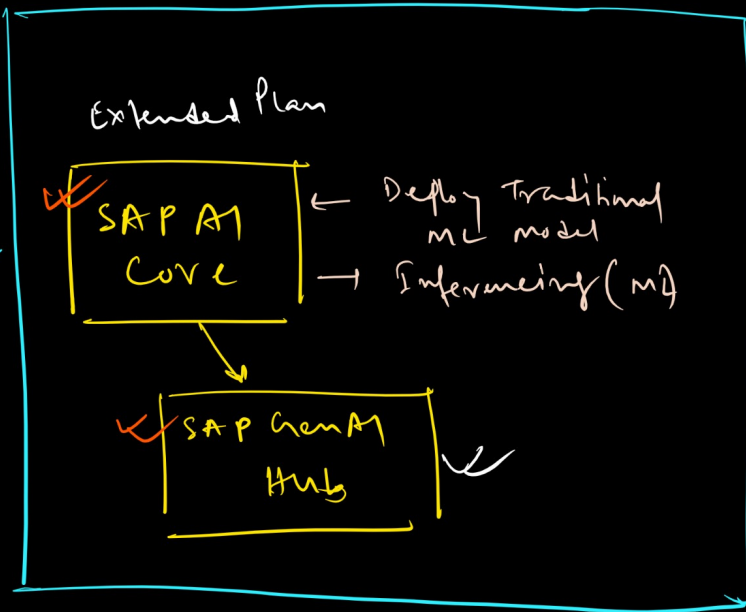
- ⊗ SAP GenAI Reference Architecture
- ⊗ SAP BTP GenAI Hub (Overview)
- ⊗ Build Chat application
- ⊗ Build SQL Agent application
- ⊗ RAG (Retrieval-Augmented Generation)
- ⊗ Build RAG Application
using SAP Gen AI Hub + HANA



*Using SAP Cloud Application Programming Model (CAP) according to SAP BTP Developer's Guide
 **Cloud Foundry only
 ***planned



✓ AI Launchpad



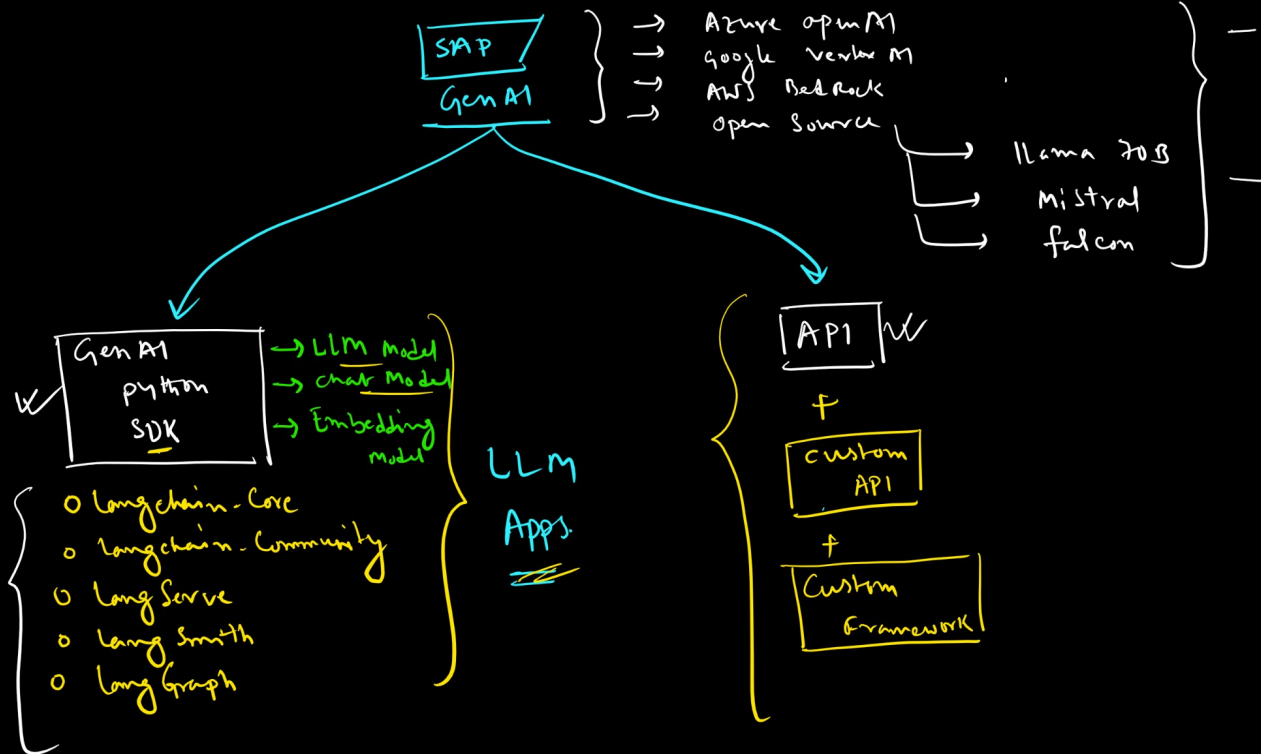
→ open source
 → open source
 → GCP (Gemini)
 → AWS Bedrock
 Instance
 (use)
 LLM Apps.

```
from gen_ai_hub.proxy.native.openai import
from gen_ai_hub.proxy.native.openai
from gen_ai_hub.proxy.langchain.init_models
from gen_ai_hub.proxy.langchain.openai
```

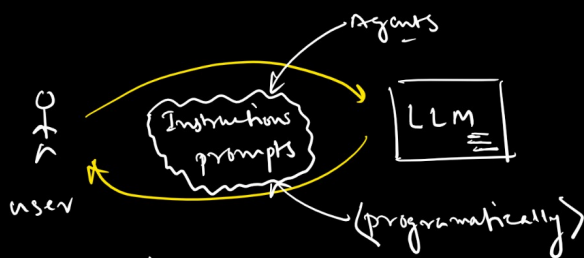
Original LLM models

```
from langchain.chains import LLMChain
from langchain.prompts import PromptTemplate
```

work/low - Agent - chain - [Langchain]



<https://help.sap.com/docs/sap-ai-core/sap-ai-core-service-guide/consume-generative-ai-models-using-sap-ai-core>

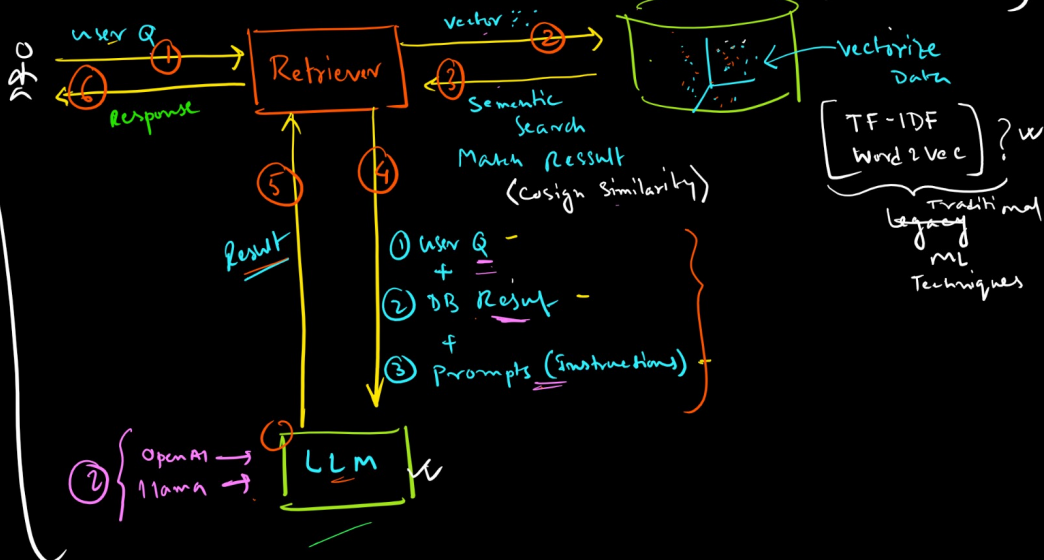


RAG → Retrieval Augmented Generation

③ langchain

① HANA Vector Engine
pine cone
Pg Vec Embedding

PDF
text
html
word
company policy



- Limitations →
- Knowledge cutoff
 - Accuracy
 - Enterprise knowledge
 - Hallucination
 - Scalability
↳ Retrain expensive

<https://www.pinecone.io/learn/vector-database/>

① → HANA cloud (BTP)
→ client ID / secret (gen M Hwls) ⇒

Chat Application (PDF - Text)

