

LLM Engineering

MASTER AI & LARGE LANGUAGE MODELS





PROGRESS REPORT

Time for LangChain

What you can now do

- Generate text and code with Frontier Models including AI Assistants with Tools, and with open-source models with HuggingFace transformers
 - Confidently choose the right LLM for your project, backed by metrics
 - Explain how RAG uses vector embeddings and vector datastores to add context to prompts
-

By end of today you'll be able to

- Describe the LangChain framework, with benefits and limitations
- Use LangChain to read in a Knowledge Base of documents
- Use LangChain to divide up documents into overlapping chunks

INTRODUCING LangChain

LangChain

- OpenSource framework launched in October 2022
- Provides a common framework for interfacing with many LLMs
- Includes its own declarative language: LangChain Expression Language (LCEL)

Pros & Cons

- Greatly simplifies the creation of applications using LLMs (eg AI assistants, RAG, summarization) - fast time to market
- Wrapper code around LLMs makes it easy to swap models
- As APIs for LLMs have matured, converged and simplified, the need for a unifying framework like LangChain has decreased

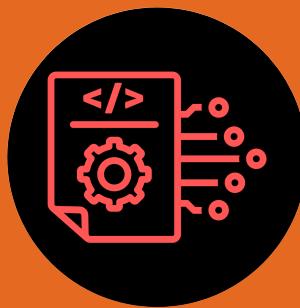


LangChain

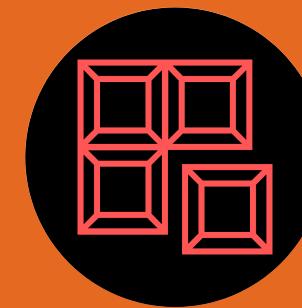
We will now use LangChain to load our Knowledge Base



Read in the documents in all folders



Add meta-data to the documents



Break down the contents into
overlapping chunks



PROGRESS REPORT

55% Towards Mastering LLMs

What you can now do

- Generate text and code with Frontier Models including AI Assistants with Tools, and with open-source models with HuggingFace transformers
- Confidently choose the right LLM for your project, backed by metrics
- Explain how RAG uses vector embeddings and vector datastores to add context to prompts, define LangChain and read / split Documents

By end of the next session you'll be able to

- Convert our chunks of text into Vectors using OpenAIEmbeddings
- Store the Vectors in Chroma, one of the most popular open-source Vector datastores
- Visualize and explore Vectors in a Chroma Vector Datastore in 2D and 3D