RAG from scratch: Overview
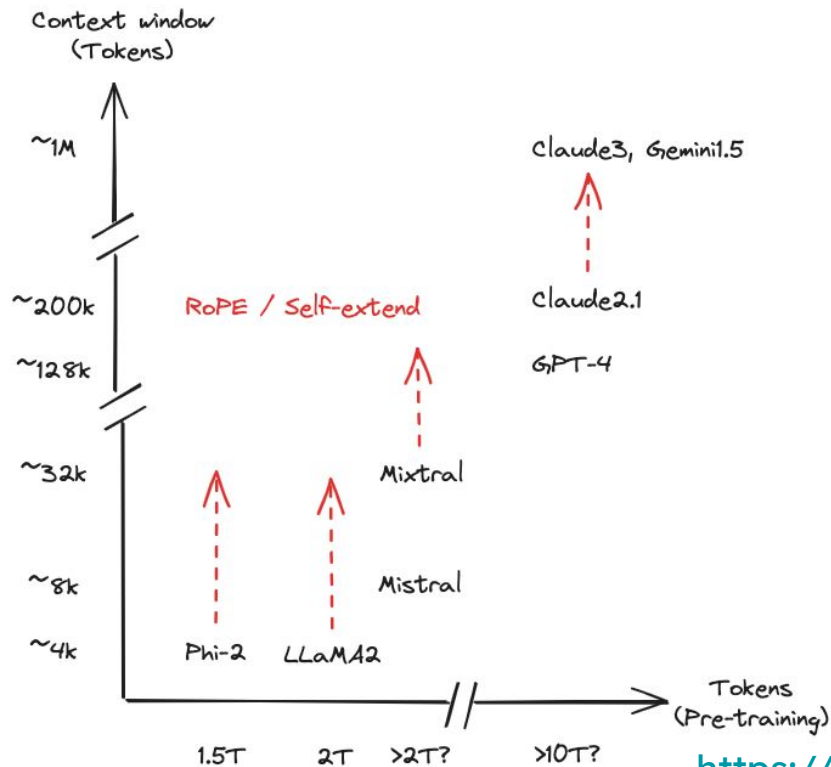
Lance Martin
Software Engineer, LangChain
@RLanceMartin

> 95% of the world's data is "private", but we can "feed it" to LLMs

https://huggingface.co/blog/mixtral
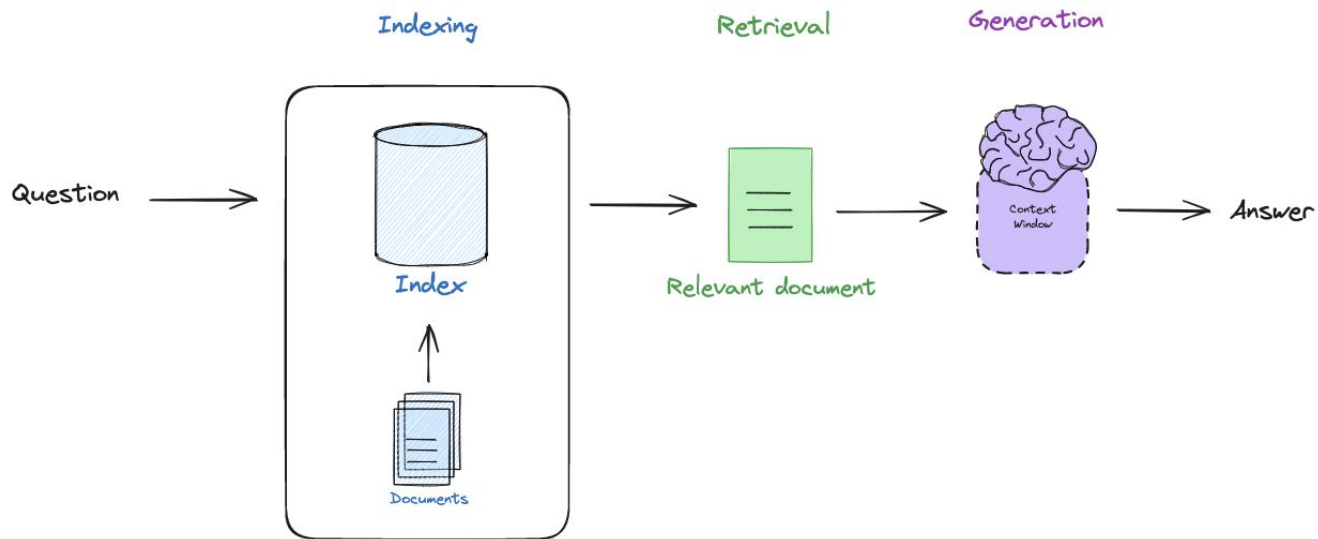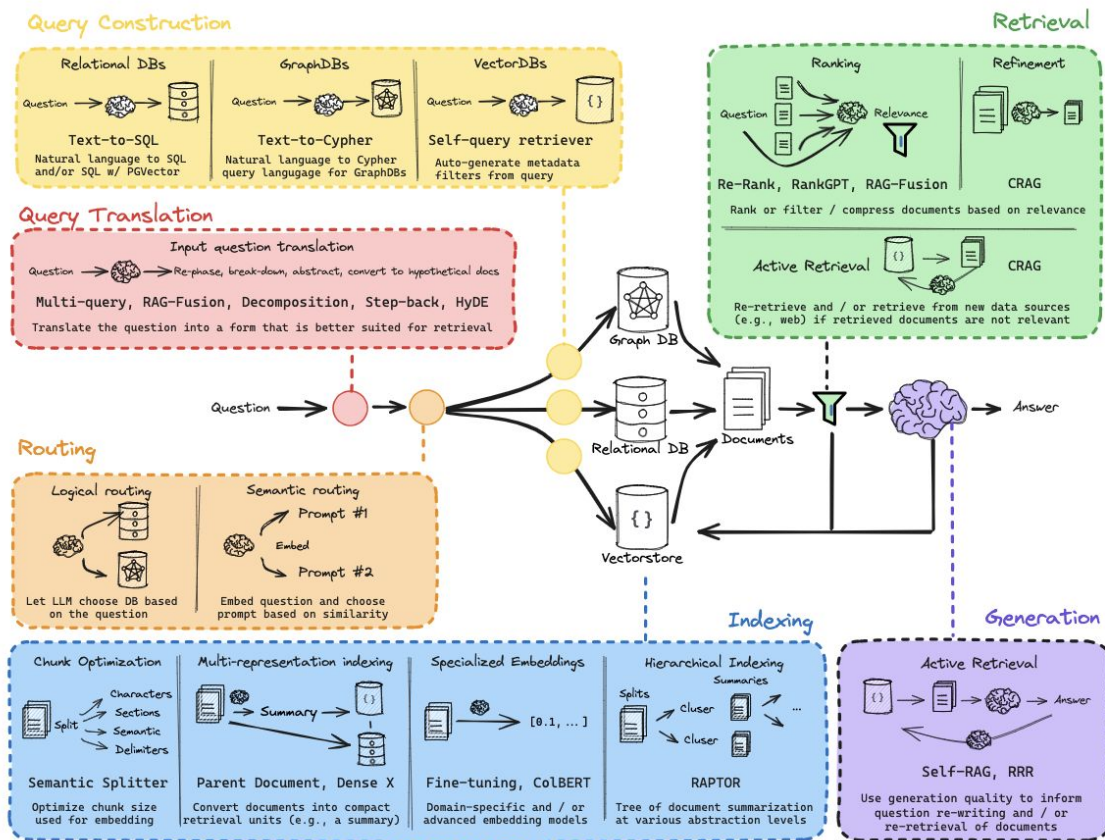https://x.com/RihardJarc/status/1778082161595208124

Connecting LLMs to external data is a central need

🦜🔗 LangChain
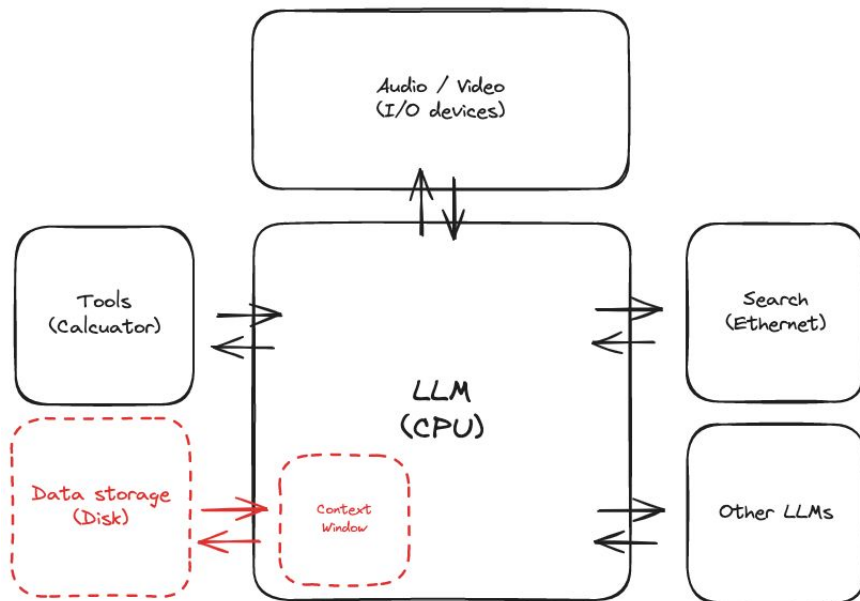
Retrieval Augmented Generation

We'll start from scratch, and build up to state of the art in RAG

Outline

Basics
- Indexing
- Retrieval
- Generation

Advanced
- Query transformations
- Routing
- Query construction
- Indexing
- Retrieval
- Generation

[Code walk-through](#)