LangChain Framework Guide

LangChain is a framework for developing applications powered by language models. It provides modular components for building complex LLM pipelines.

Core Components:

- Models: Interfaces to LLMs and embedding models

- Prompts: Template management and few-shot examples

- Chains: Sequences of calls to models and utilities

- Agents: LLMs that can use tools and take actions

- Memory: Persistence of state across interactions

LangChain Expression Language (LCEL) enables composable, streaming-first chain construction with built-in observability.

Building RAG with LangChain

LangChain simplifies RAG implementation with dedicated components:

Document Loaders:

Load data from PDFs, web pages, databases, and more.

Text Splitters:

Chunk documents with configurable overlap for context preservation.

Vector Stores:

Native integrations with ChromaDB, Pinecone, Weaviate, and others.

Retrievers:

Customize retrieval with MMR, self-query, or ensemble strategies.

Chains:

RetrievalQA, ConversationalRetrievalChain for different use cases.