Welcome to LangChain

Quickstart Guide

Concepts

Tutorials

Models

Getting Started

LLMs

Generic Functionality

How to use the async API for LLMs

How to write a custom LLM wrapper

How (and why) to use the fake LLM

How (and why) to use the human input LLM

How to cache LLM calls

How to serialize LLM classes

How to stream LLM and Chat Model responses

How to track token usage

Integrations

AI21

Aleph Alpha

Anyscale

Azure OpenAI

Banana

Beam integration for langchain

CerebriumAI

Cohere

C Transformers

Databricks

DeepInfra

ForefrontAI

Google Cloud Platform Vertex AI PaLM

GooseAI

GPT4All

Hugging Face Hub

Hugging Face Local Pipelines

Huggingface TextGen Inference

Structured Decoding with JSONFormer

Llama-cpp

Manifest

Modal

MosaicML

NLP Cloud

OpenAI

OpenLM

Petals

PipelineAI

PredictionGuard

PromptLayer OpenAI

Structured Decoding with RELLM

Replicate

Runhouse

SageMakerEndpoint

StochasticAI

Writer

Reference

Chat Models

How-To Guides

How to use few shot examples

How to stream responses

Anthropic

Azure

PromptLayer ChatOpenAI

Text Embedding Models

AzureOpenAI

Fake Embeddings

InstructEmbeddings

Jina

MiniMax

ModelScope

MosaicML embeddings

SageMaker Endpoint Embeddings

Self Hosted Embeddings

Sentence Transformers Embeddings

TensorflowHub

Prompts

Prompt Templates

Connecting to a Feature Store

How to create a custom prompt template

How to create a prompt template that uses few shot examples

How to work with partial Prompt Templates

How to serialize prompts

PromptTemplates

Example Selector

Output Parsers

Chat Prompt Template

Example Selectors

How to create a custom example selector

LengthBased ExampleSelector

Maximal Marginal Relevance ExampleSelector

NGram Overlap ExampleSelector

Similarity ExampleSelector

CommaSeparatedListOutputParser

Enum Output Parser

OutputFixingParser

PydanticOutputParser

RetryOutputParser

Structured Output Parser

Memory

ConversationBufferMemory

ConversationBufferWindowMemory

Entity Memory

Conversation Knowledge Graph Memory

ConversationSummaryMemory

ConversationSummaryBufferMemory

ConversationTokenBufferMemory

VectorStore-Backed Memory

How to add Memory to an LLMChain

How to add memory to a Multi-Input Chain

How to add Memory to an Agent

Adding Message Memory backed by a database to an Agent

Cassandra Chat Message History

How to customize conversational memory

How to create a custom Memory class

Dynamodb Chat Message History

Momento

Mongodb Chat Message History

Motörhead Memory

How to use multiple memory classes in the same chain

Postgres Chat Message History

Redis Chat Message History

Zep Memory

Indexes

Document Loaders

CoNLL-U

Copy Paste

CSV

Email

EPub

EverNote

Facebook Chat

File Directory

HTML

Images

Jupyter Notebook

JSON

Markdown

Microsoft PowerPoint

Microsoft Word

Open Document Format (ODT)

Pandas DataFrame

PDF

Sitemap

Subtitle

Telegram

TOML

Unstructured File

URL

WebBaseLoader

Weather

WhatsApp Chat

Arxiv

AZLyrics

BiliBili

College Confidential

Gutenberg

Hacker News

HuggingFace dataset

iFixit

IMSDb

MediaWikiDump

Wikipedia

YouTube transcripts

Airbyte JSON

Apify Dataset

AWS S3 Directory

AWS S3 File

Azure Blob Storage Container

Azure Blob Storage File

Blackboard

Blockchain

ChatGPT Data

Confluence

Diffbot

Discord

Docugami

DuckDB

Figma

GitBook

Git

Google BigQuery

Google Cloud Storage Directory

Google Cloud Storage File

Google Drive

Image captions

Iugu

Joplin

Microsoft OneDrive

Modern Treasury

Notion DB 2/2

Notion DB 1/2

Obsidian

Psychic

ReadTheDocs Documentation

Reddit

Roam

Slack

Spreedly

Stripe

2Markdown

Twitter

Text Splitters

Character

LaTeX

NLTK

Python Code

Recursive Character

spaCy

Tiktoken

Hugging Face tokenizer

tiktoken (OpenAI) tokenizer

Vectorstores

AnalyticDB

Annoy

Atlas

Chroma

Deep Lake

DocArrayHnswSearch

DocArrayInMemorySearch

ElasticSearch

FAISS

LanceDB

Milvus

MyScale

OpenSearch

PGVector

Pinecone

Qdrant

Redis

SKLearnVectorStore

Supabase (Postgres)

Tair

Typesense

Vectara

Weaviate

Zilliz

Retrievers

Azure Cognitive Search Retriever

ChatGPT Plugin

Self-querying with Chroma

Cohere Reranker

Contextual Compression

Databerry

ElasticSearch BM25

kNN

Metal

Pinecone Hybrid Search

Self-querying

SVM

TF-IDF

Time Weighted VectorStore

VectorStore

Vespa

Weaviate Hybrid Search

Self-querying with Weaviate

Chains

Async API for Chain

Creating a custom Chain

Loading from LangChainHub

LLM Chain

Router Chains

Sequential Chains

Serialization

Transformation Chain

Analyze Document

Chat Over Documents with Chat History

Graph QA

Hypothetical Document Embeddings

Question Answering with Sources

Question Answering

Summarization

Retrieval Question/Answering

Retrieval Question Answering with Sources

Vector DB Text Generation

API Chains

Self-Critique Chain with Constitutional AI

FLARE

GraphCypherQAChain

BashChain

LLMCheckerChain

LLM Math

LLMRequestsChain

LLMSummarizationCheckerChain

Moderation

Router Chains: Selecting from multiple prompts with MultiPromptChain

Router Chains: Selecting from multiple prompts with MultiRetrievalQAChain

OpenAPI Chain

PAL

SQL Chain example

Agents

Tools

Defining Custom Tools

Multi-Input Tools

Tool Input Schema

Apify

ArXiv API Tool

AWS Lambda API

Shell Tool

Bing Search

ChatGPT Plugins

DuckDuckGo Search

File System Tools

Google Places

Google Search

Google Serper API

Gradio Tools

GraphQL tool

HuggingFace Tools

Human as a tool

IFTTT WebHooks

Metaphor Search

OpenWeatherMap API

Python REPL

Requests

SceneXplain

Search Tools

SearxNG Search API

SerpAPI

Twilio

Wolfram Alpha

YouTubeSearchTool

Zapier Natural Language Actions API

Agent Types

Custom Agent

Custom LLM Agent

Custom LLM Agent (with a ChatModel)

Custom MRKL Agent

Custom MultiAction Agent

Custom Agent with Tool Retrieval

Conversation Agent (for Chat Models)

Conversation Agent

MRKL

MRKL Chat

ReAct

Self Ask With Search

Structured Tool Chat Agent

Toolkits

Azure Cognitive Services Toolkit

CSV Agent

Gmail Toolkit

Jira

JSON Agent

OpenAPI agents

Natural Language APIs

Pandas Dataframe Agent

PlayWright Browser Toolkit

PowerBI Dataset Agent

Python Agent

Spark Dataframe Agent

Spark SQL Agent

SQL Database Agent

Vectorstore Agent

Agent Executors

How to combine agents and vectorstores

How to use the async API for Agents

How to create ChatGPT Clone

Handle Parsing Errors

How to access intermediate steps

How to cap the max number of iterations

How to use a timeout for the agent

How to add SharedMemory to an Agent and its Tools

Plan and Execute

Callbacks

Autonomous Agents

Agent Simulations

Question Answering over Docs

Chatbots

Querying Tabular Data

Code Understanding

Interacting with APIs

Extraction

Evaluation

Agent Benchmarking: Search + Calculator

Agent VectorDB Question Answering Benchmarking

Benchmarking Template

Data Augmented Question Answering

Generic Agent Evaluation

Using Hugging Face Datasets

Evaluating an OpenAPI Chain

Question Answering Benchmarking: Paul Graham Essay

Question Answering Benchmarking: State of the Union Address

QA Generation

SQL Question Answering Benchmarking: Chinook

Installation

API References

Embeddings

Docstore

Text Splitter

Vector Stores

Document Compressors

Document Transformers

Agent Toolkits

Utilities

Experimental Modules

AI21 Labs

Aim

AtlasDB

Beam

ClearML Integration

Comet

Google Serper

Graphsignal

Hazy Research

Helicone

Hugging Face

Llama.cpp

MLflow

NLPCloud

Prediction Guard

PromptLayer

Rebuff: Prompt Injection Detection with LangChain

RWKV-4

scikit-learn

Unstructured

Weights & Biases

WhyLabs Integration

Wolfram Alpha Wrapper

Yeager.ai

Dependents

Deployments