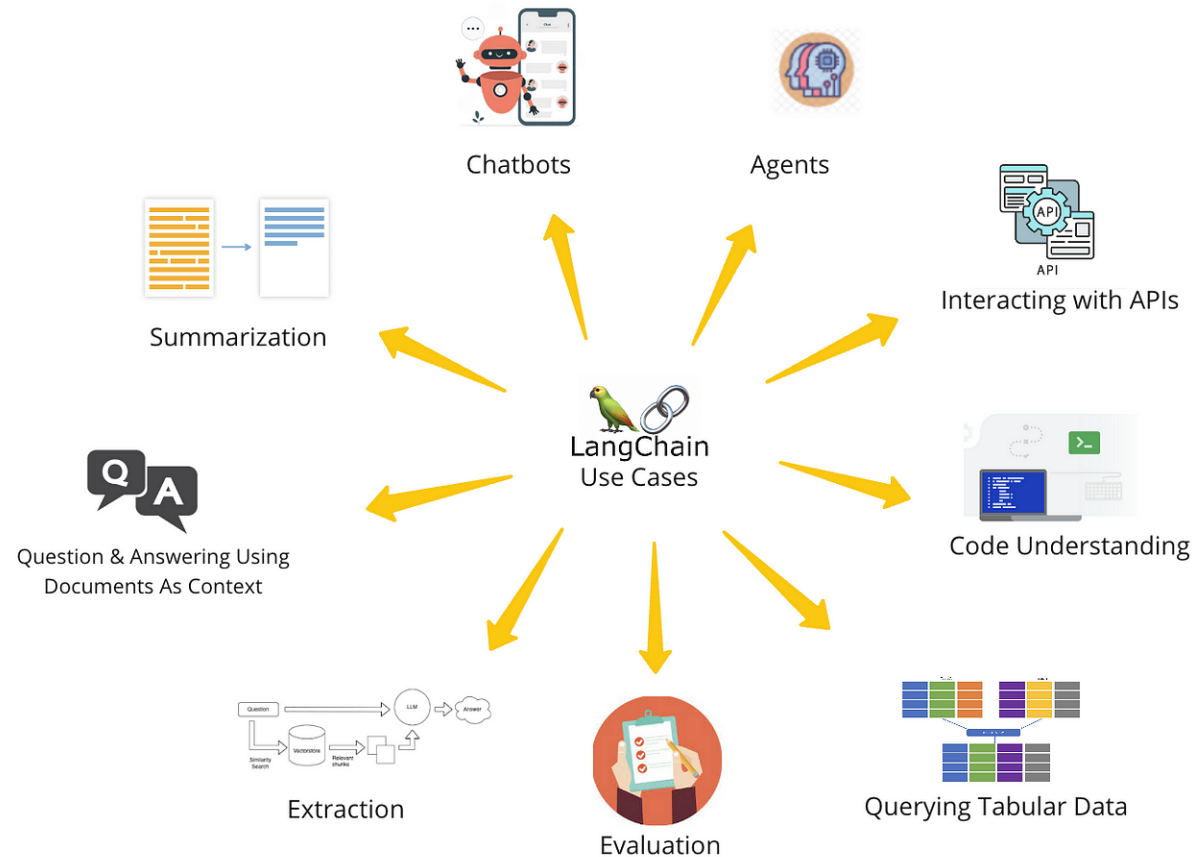


8. LangChain 활용사례



Q&A with RAG

RAG는 추가 데이터로 LLM의 지식을 보강하는 기술입니다.

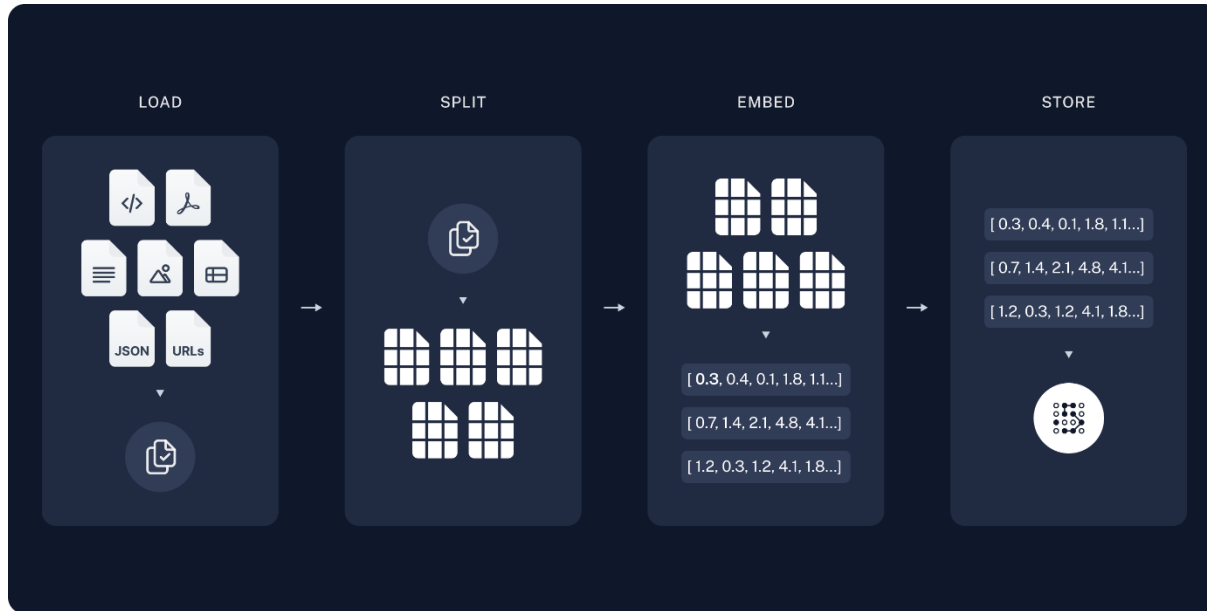
LLM은 광범위한 주제에 대해 추론할 수 있지만, 그 지식은 학습된 특정 시점까지의 공개 데이터로 제한됩니다.

비공개 데이터나 모델의 마감일 이후에 도입된 데이터에 대해 추론할 수 있는 AI 애플리케이션을 구축하려면

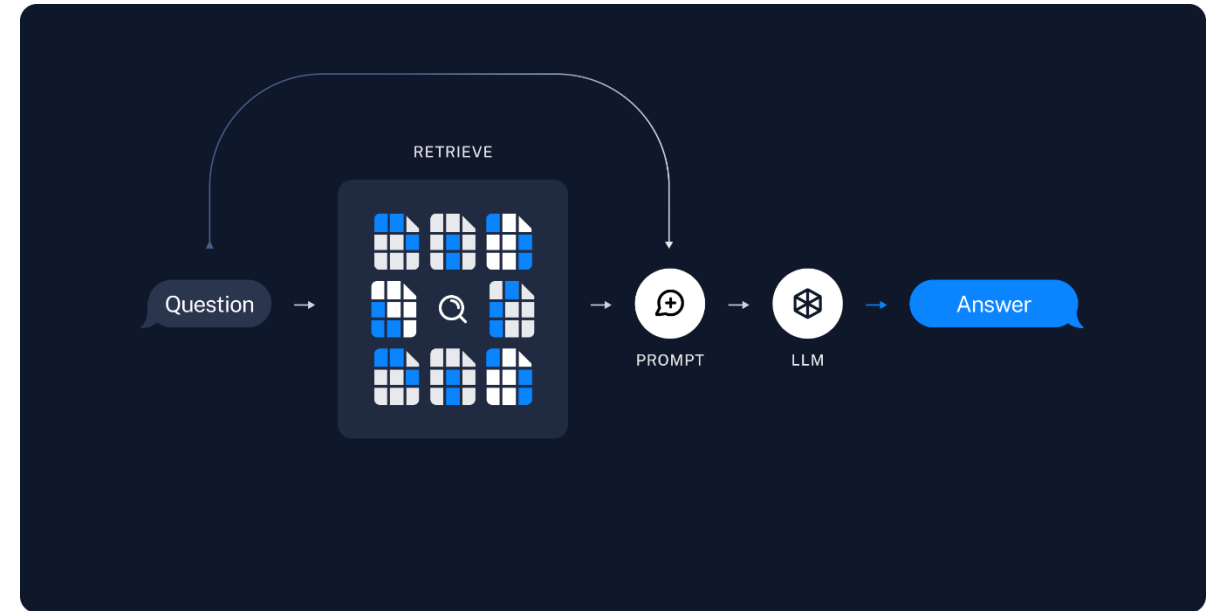
모델에 필요한 특정 정보로 모델의 지식을 보강해야 합니다.

적절한 정보를 가져와 모델 프롬프트에 삽입하는 프로세스를 검색 증강 생성(RAG)이라고 합니다.

Indexing



Retrieval and generation



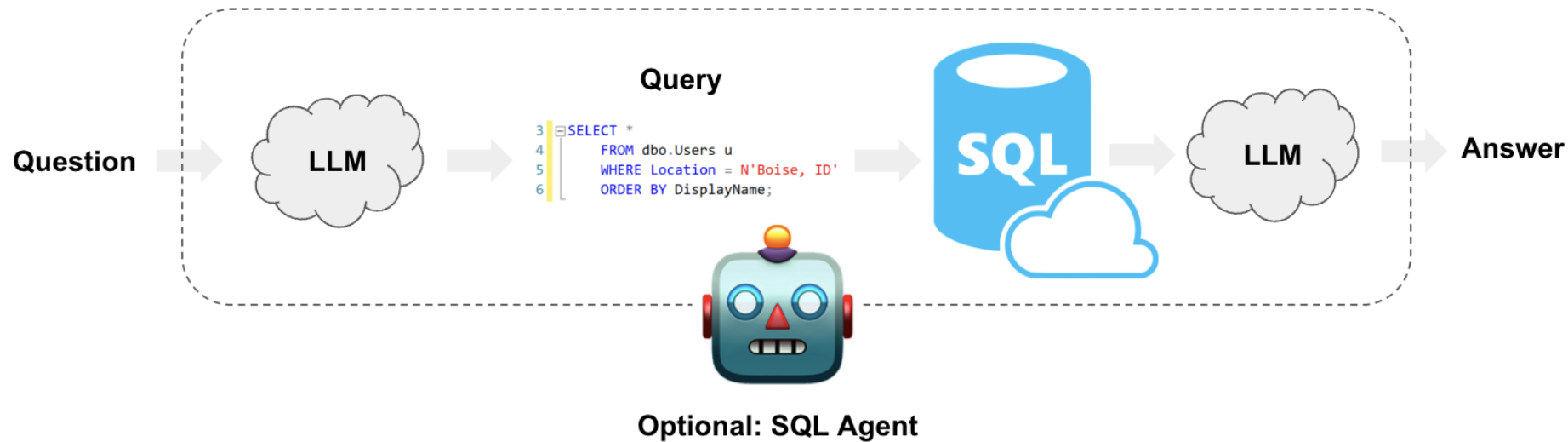
question_answering

SQL

Q&A 시스템을 구축할 수 있는 가장 일반적인 유형의 데이터베이스 중 하나는 SQL 데이터베이스입니다. LangChain에는 SQLAlchemy에서 지원하는 모든 SQL 언어(예: MySQL, PostgreSQL, Oracle SQL, Databricks, SQLite)와 호환되는 수많은 기본 제공 체인 및 에이전트가 포함되어 있습니다.

다음과 같은 사용 사례를 지원합니다.

- 자연어 질문을 기반으로 실행될 쿼리 생성
- 데이터베이스 데이터를 기반으로 질문에 답변할 수 있는 챗봇 만들기
- 사용자가 분석하고자 하는 인사이트를 기반으로 사용자 지정 대시보드 구축

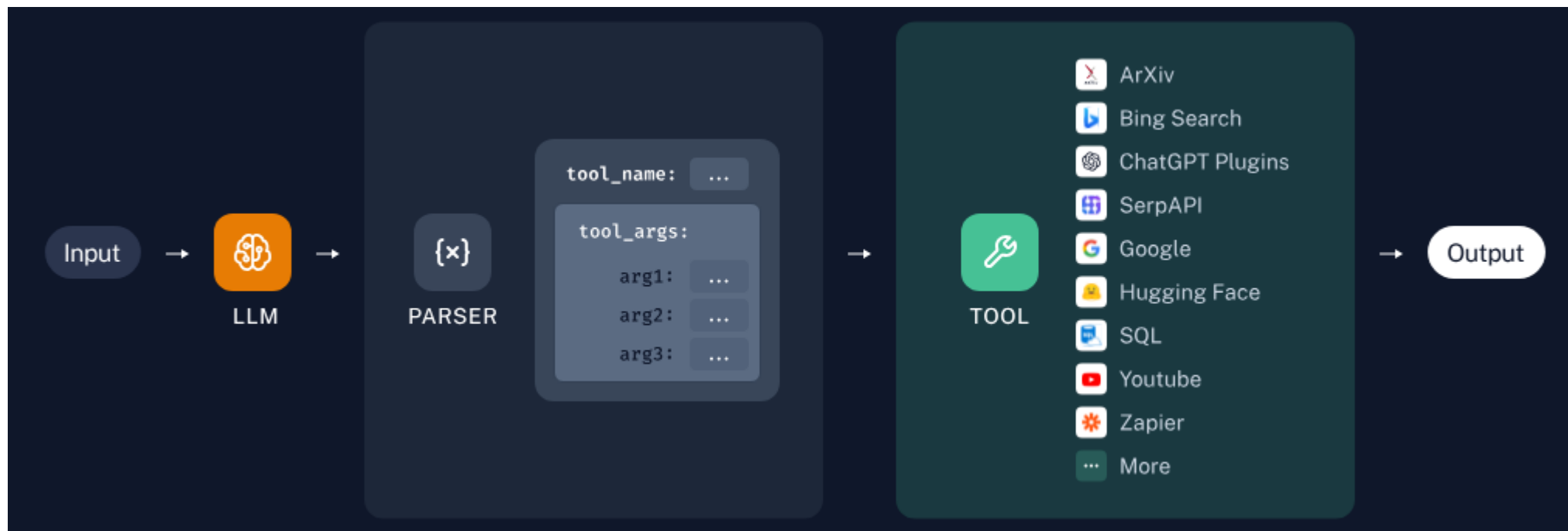


sql

Tool use

LLM의 흥미로운 사용 사례는 API, 함수, 데이터베이스 등 다른 '도구'를 위한 자연어 인터페이스를 구축하는 것입니다. LangChain은 자연어 인터페이스를 구축하는 데 매우 유용합니다

- 우수한 모델 출력 구문 분석 기능으로 모델 출력에서 JSON, XML, OpenAI 함수 호출 등을 쉽게 추출할 수 있습니다.
- 다양한 기본 제공 도구 모음
- 도구를 호출하는 방식에 있어 많은 유연성을 제공합니다.



tool

CSV

LLM은 다양한 유형의 데이터 소스에 대한 질문 답변 시스템을 구축하는 데 유용합니다.

이를 수행하는 두 가지 주요 방법은 다음과 같습니다.

- (권장) CSV를 SQL 데이터베이스에 로드하고 SQL 사용 사례 문서에 설명된 접근 방식을 사용합니다.
- LLM이 데이터와 상호 작용하기 위해 Pandas와 같은 라이브러리를 사용할 수 있는 Python 환경에 대한 액세스 권한을 부여합니다.

SQL

임의의 Python을 사용할 때보다 권한을 제한하고 쿼리를 처리하기가 더 쉽기 때문에 SQL을 사용하여 CSV 데이터와 상호 작용하는 것이 권장되는 방식입니다.

대부분의 SQL 데이터베이스는 CSV 파일을 테이블로 쉽게 로드할 수 있습니다(DuckDB, SQLite 등).

Pandas

SQL 대신 Pandas와 같은 데이터 분석 라이브러리와 LLM의 코드 생성 기능을 사용하여 CSV 데이터와 상호 작용할 수도 있습니다.

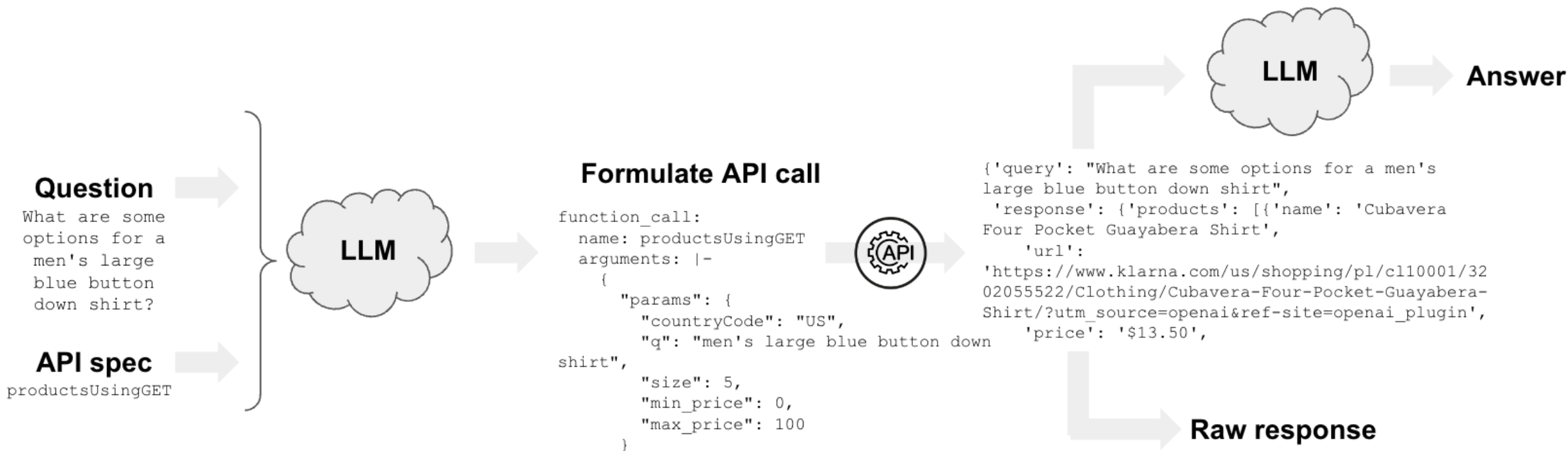


CSV

Interacting with API

LLM을 외부 API와 인터페이스하는 방법에는 크게 두 가지가 있습니다.

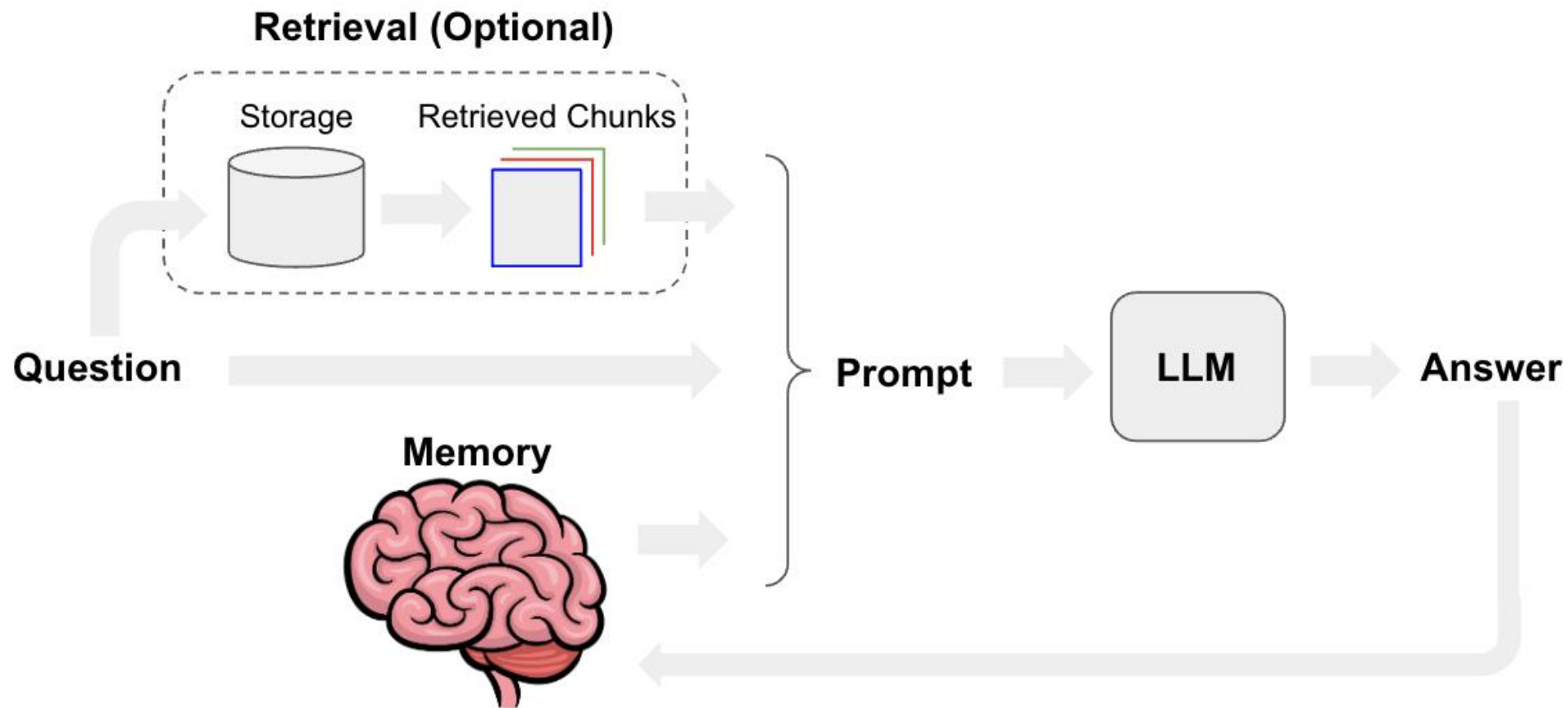
- Function : 예를 들어 OpenAI 함수는 널리 사용되는 방법 중 하나입니다.
- LLM 생성 인터페이스 : API 문서에 액세스할 수 있는 LLM을 사용하여 인터페이스를 생성합니다.



API

Chatbot

챗봇은 LLM의 가장 인기 있는 사용 사례 중 하나입니다. 챗봇의 핵심 기능은 장기간의 상태 저장 대화를 할 수 있고 관련 정보를 사용하여 사용자의 질문에 답변할 수 있다는 것입니다.

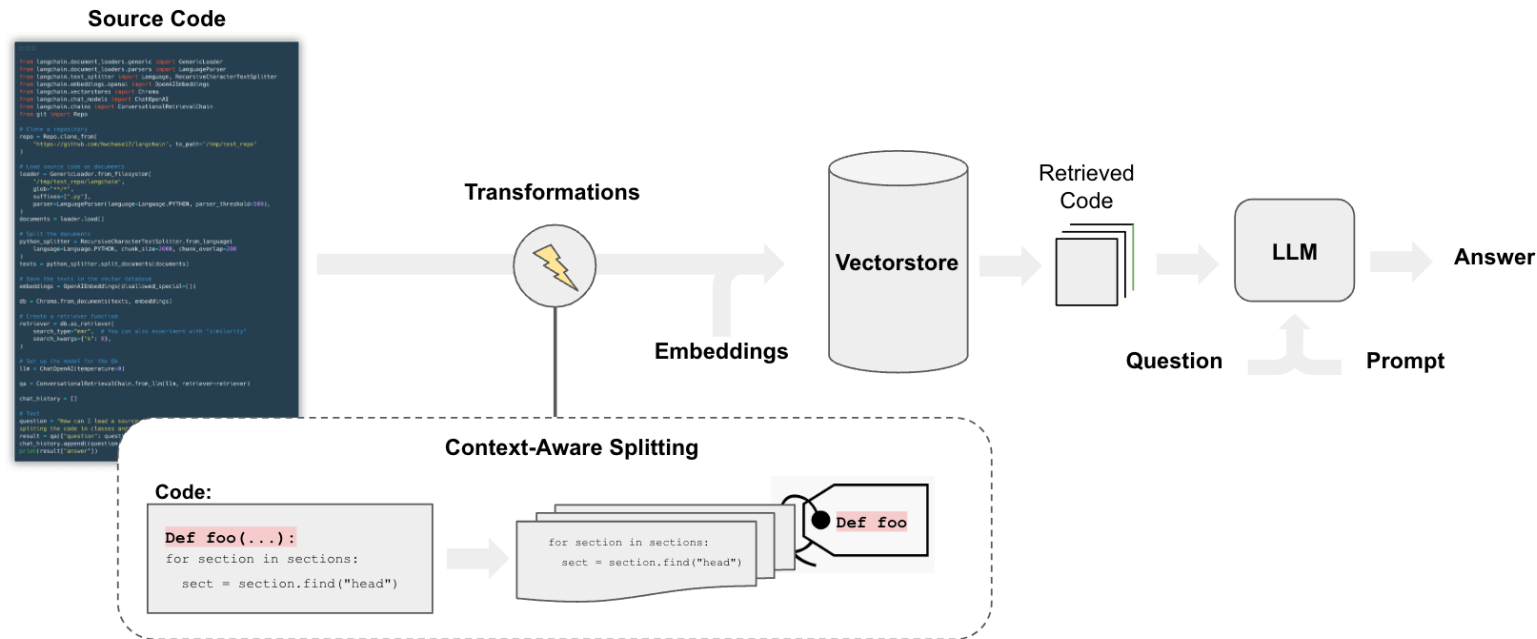


Chatbot

Code understanding

소스 코드 분석은 다음과 같은 사용 사례에 가장 많이 사용되는 LLM 애플리케이션(예: GitHub Copilot, 코드 인터프리터, Codium 및 Codeium) 중 하나입니다,

- Code base에 대한 Q&A를 통해 작동 방식 이해
- 리팩터링 또는 개선 사항 제안을 위한 LLM 사용 LLM을 사용하여 코드 문서화하기



code understanding

Synthetic data generation

합성 데이터는 실제 이벤트에서 수집된 데이터가 아닌 인위적으로 생성된 데이터입니다.
개인 정보를 침해하거나 현실적인 제약에 부딪히지 않고 실제 데이터를 시뮬레이션하는 데 사용됩니다.

합성 데이터의 이점

개인정보 보호 및 보안: 실제 개인 데이터가 유출될 위험이 없습니다.

데이터 증강: 머신러닝을 위한 데이터 세트 확장.

유연성: 특정 또는 희귀 시나리오를 생성할 수 있습니다.

비용 효율적: 실제 데이터 수집보다 저렴한 경우가 많습니다.

규제 준수: 엄격한 데이터 보호법을 준수하는 데 도움이 됩니다.

모델 견고성: AI 모델을 더 잘 일반화할 수 있습니다.

신속한 프로토타이핑: 실제 데이터 없이도 빠르게 테스트할 수 있습니다.

제어된 실험: 특정 조건을 시뮬레이션할 수 있습니다.

데이터 액세스: 실제 데이터를 사용할 수 없을 때 대안으로 사용할 수 있습니다.



data generation

Extraction

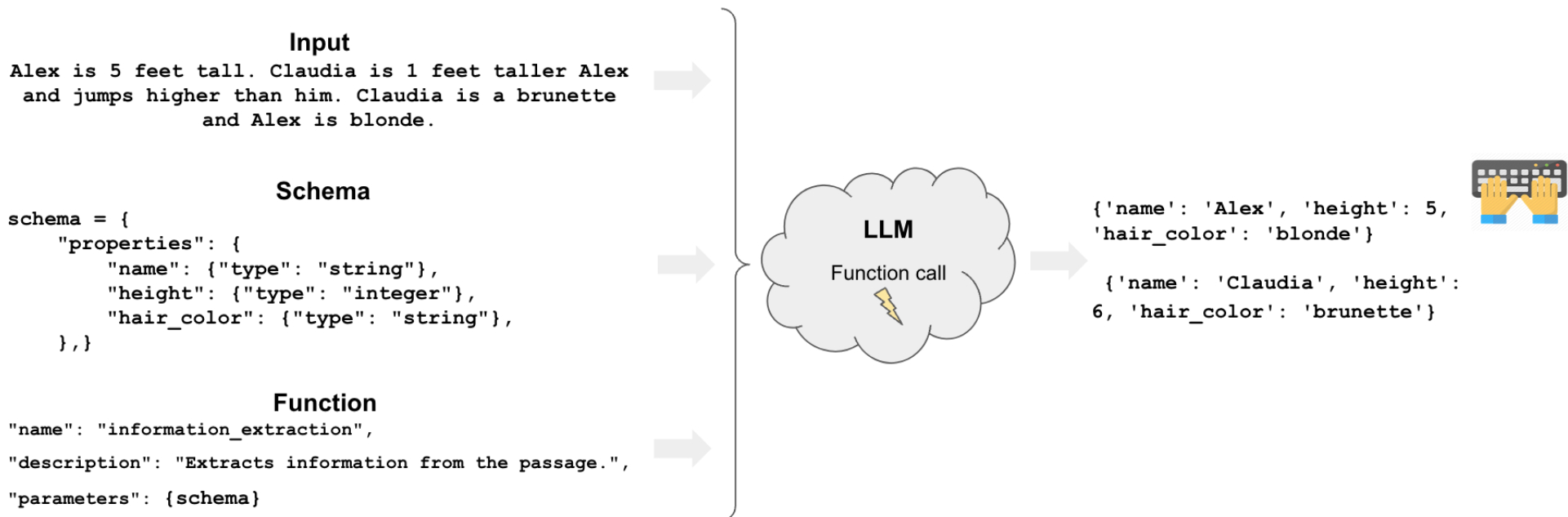
원시 LLM 생성에서 구조화된 출력을 얻는 것은 어렵습니다.

예를 들어 특정 스키마로 포맷된 모델 출력이 필요하다고 가정해 보겠습니다.

- 데이터베이스에 삽입할 구조화된 행을 추출하는 경우
- API 매개변수 추출
- 사용자 쿼리의 다른 부분 추출(예: 시맨틱 검색과 키워드 검색)

출력을 구조화 하는 두 가지 기본 접근 방식이 있습니다.

- Function : 일부 LLM은 함수를 호출하여 LLM 응답에서 임의의 엔티티를 추출할 수 있습니다.
- Parsing : Output parsers는 LLM 응답을 구조화 하는 클래스입니다.



Graph querying

그래프 데이터베이스는 실제 관계를 표현하고 쿼리할 수 있는 강력한 방법을 제공합니다.
다양한 그래프 DB와 상호 작용하기 위해 LLM을 쉽게 사용할 수 있는 여러 가지 체인이 있습니다.

Diffbot Graph Transformer

Open In Colab

ArangoDB QA chain

Open In Colab

Memgraph QA chain

This notebook shows how to use LLMs to provide a natural langu...

NebulaGraphQAChain

This notebook shows how to use LLMs to provide a natural langu...

Neo4j DB QA chain

This notebook shows how to use LLMs to provide a natural langu...

FalkorDBQAChain

This notebook shows how to use LLMs to provide a natural langu...

NetworkX

NetworkX is a Python package for the

Ontotext GraphDB QA Chain

This notebook shows how to use LLMs to provide natural langua...

HugeGraph QA Chain

This notebook shows how to use LLMs to provide a natural langu...

KuzuQAChain

This notebook shows how to use LLMs to provide a natural langu...

GraphSparqlQAChain

Graph databases are an excellent choice for applications based on

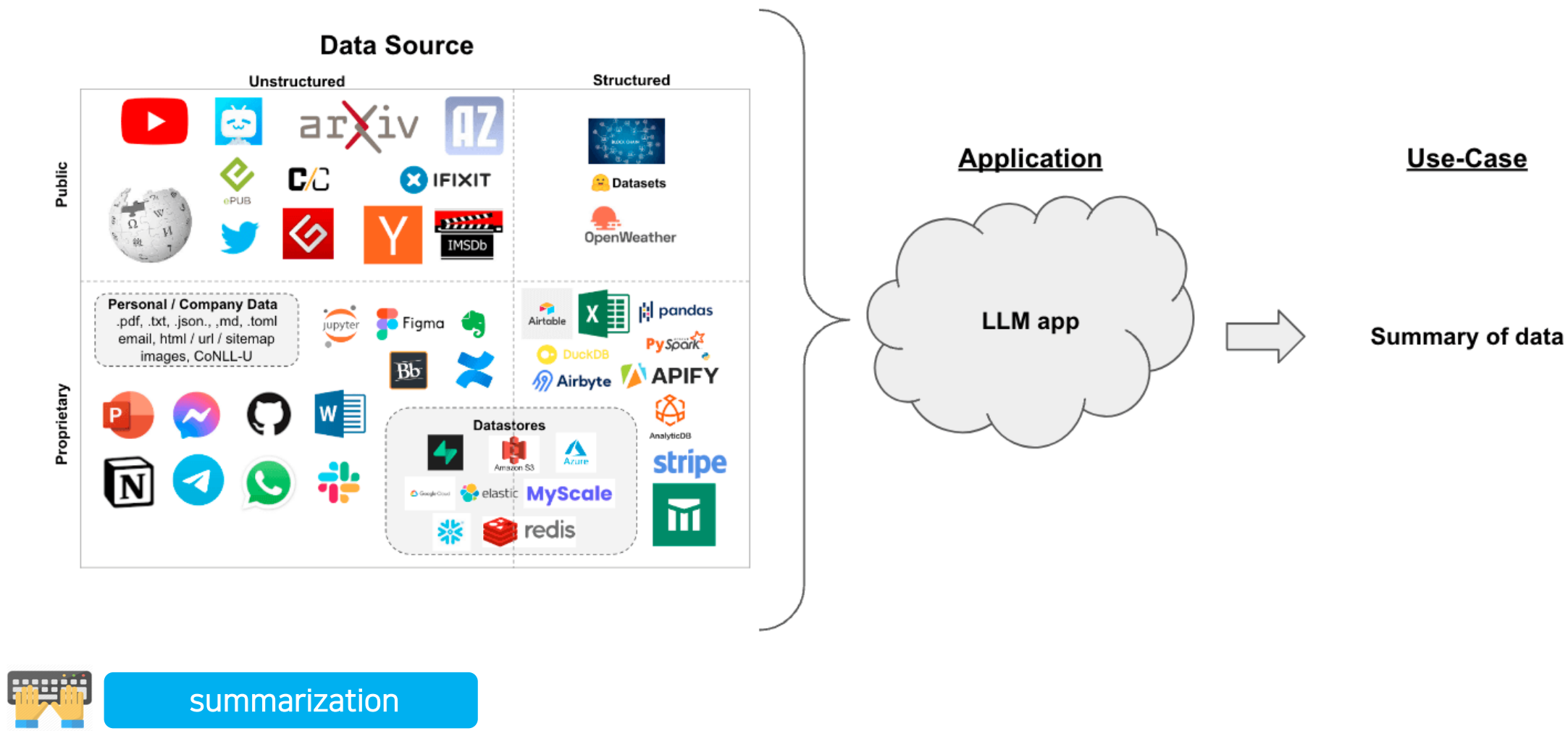
Neptune Open Cypher QA Chain

This QA chain queries Neptune graph database using openCyphe...

https://python.langchain.com/docs/use_cases/graph/

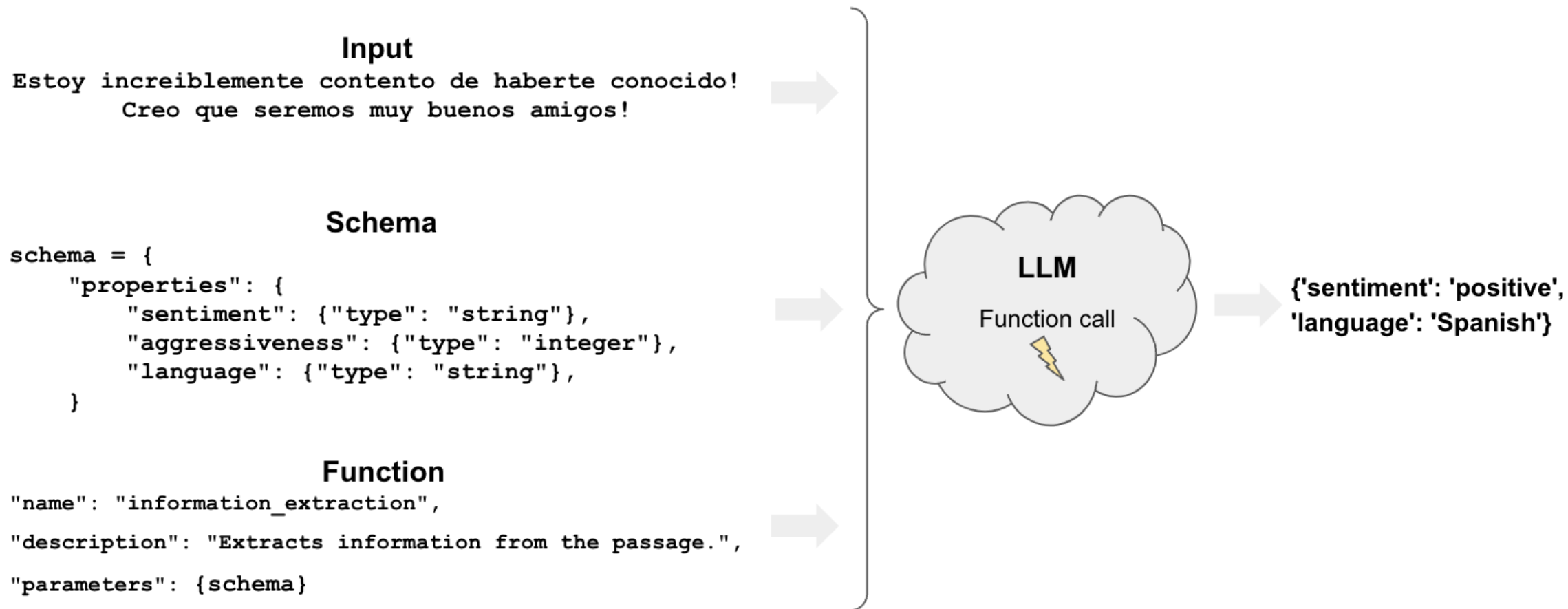
Summarization

텍스트를 이해하고 생성하는 LLM은 문서(PDF, Notion 페이지, 고객 질문 등) 내용을 요약하는데 훌륭한 도구입니다.



Tagging

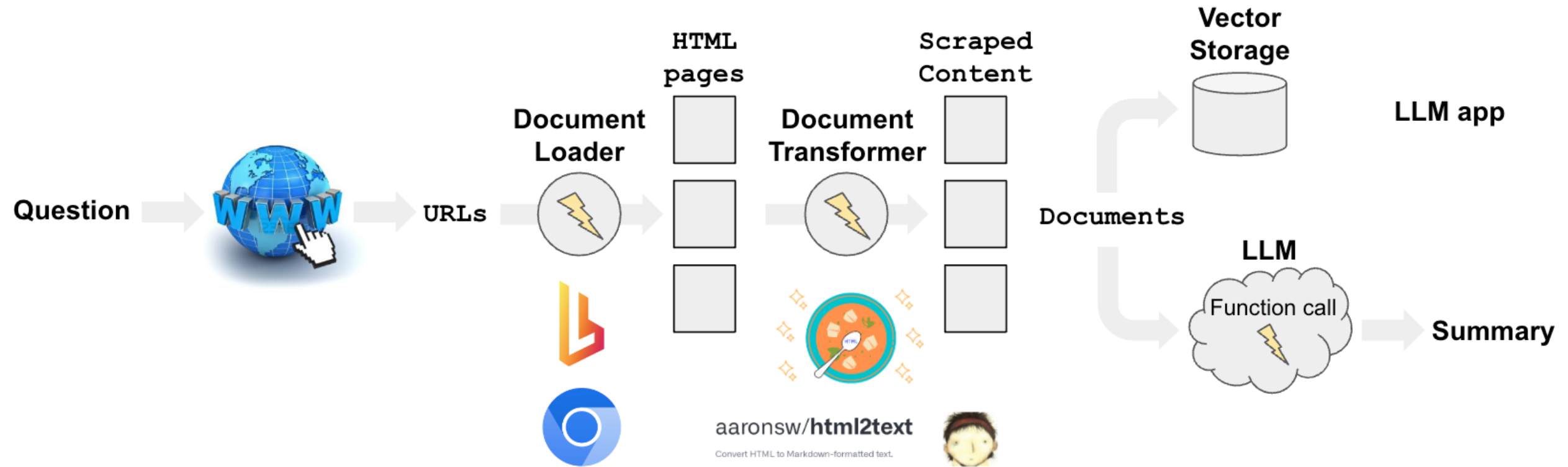
태그 지정은 문서에 감성, 언어, 스타일, 토픽, 정치적 성향 같은 클래스로 레이블을 지정하는 것입니다.



tagging

Web scraping

웹 리서치는 킬러 LLM 애플리케이션 중 하나입니다.



web scraping





수고하셨습니다. 감사합니다.