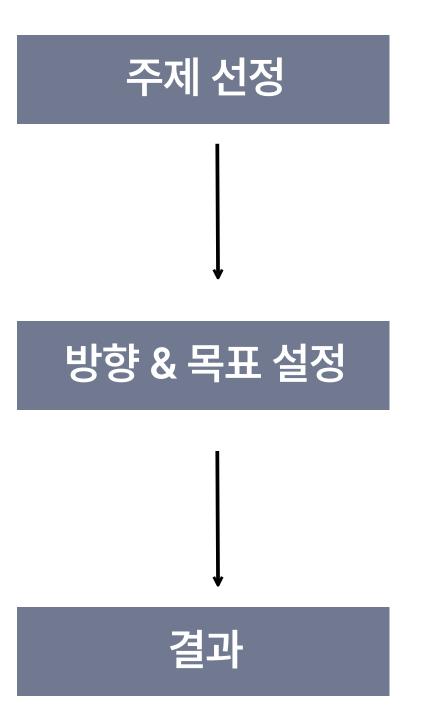
# 부동산 전월세 실거래가 분석 및 질의응답 챗봇 서비스

발표 일시 : 2025.05.20(화)

발표자: 권선옥

#### 프로젝트 해결 과정



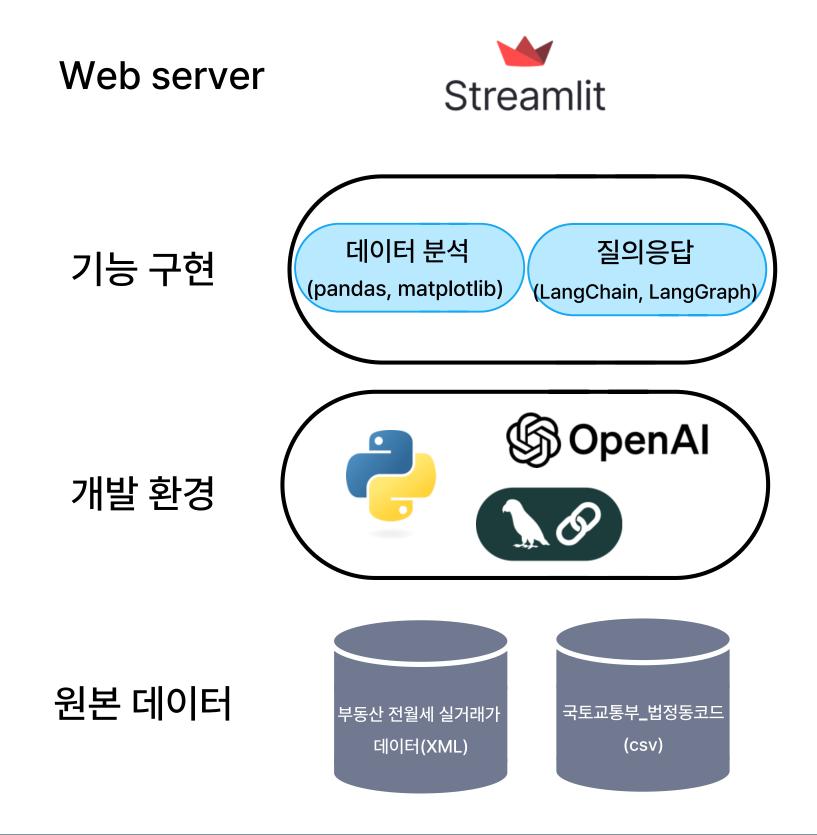
- Target : 최신 부동산 실거래 데이터를 파악하고 싶은 사람
- Pain Point:
  - 방대한 부동산 데이터 분석의 어려움(시간, 비용)
  - 특정 지역 등 조건별 데이터를 찾고 요약하기 어려움
- 방향: 부동산 전월세 추이와 주요 결과를 쉽게 파악할 수 있게 하자!
- Solution : 부동산 전월세 실거래가 분석 및 대화형 챗봇 기획

● 전월세 실거래가 추이 시각화 분석 및 질의응답 챗봇 앱 완성

### 기술 스택

분류	기술 / 라이브러리	역할 / 설명
웹 프레임워크	Streamlit	웹 앱 UI 구성 및 사용자 입력 처리
데이터 수집	requests, httpx, xmltodict	공공데이터 API 비동기 호출 및 XML 파싱
비동기 처리	asyncio	여러 개월 데이터를 병렬로 수집
LLM 기반 QA	LangChain (ChatOpenAl, FAISS, TextSplitter, PromptTemplate, Document, create_retrieval_chain)	벡터 검색 기반 RAG QA 시스템 구성
메모리	ConversationBufferMemory	LangGraph 챗봇 대화 히스토리 저장
챗봇 아키텍처	LangGraph (StateGraph, RunnableLambda, END)	LangChain 기반 챗봇의 상태 전이 그래프 구성

**Project Architecture** 



- Vector Store : FAISS

- Prompt Template

1) 질문 : {input}

Retrieval

2 Augmented

- Embedding model: OpenAlEmbeddings

2) 참고 문서(vectorDB) : {context}

3) 대화 기록 : {chat\_history}

#### Service Diagram

컬럼명 항목명
sggCd 지역코드
umdNm 법정동
mhouseNm 연립다세대명

- LangGraph **3** Generation  $Start([Start]) \rightarrow user\_input$  $\rightarrow$  generate  $\rightarrow$  End([End]) 부동산 전월세 실거래가 데이터 분석 데이터(XML) 조건 설정 필터링된 질의응답 '지역코드'와 질문 입력 (pandas, matplotlib) 맵핑 (지역, 유형, 조회 기간 등) (LangChain, LangGraph) 데이터

보증금 & 전/월세 추이

시각화

	법정동코드	법정동명	폐지여부
	1100000000	서울특별시	존재
	1111000000	서울특별시 종로구	존재

국토교통부\_법정동코드

(csv)

모델 : gpt-4o

필터링된 데이터 기반 응답

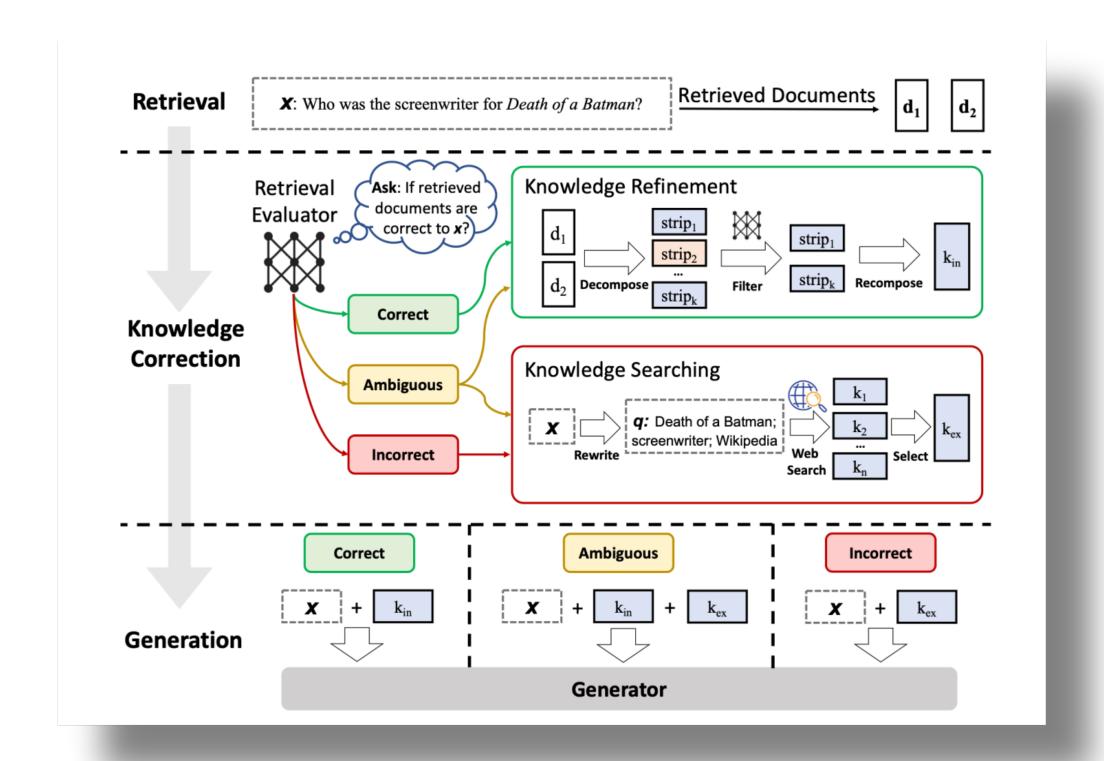
시연

## 코드 및 앱 시연

#### **PSEUDO LAB 10TH ACADEMY**

#### 후기

- 배운점
  - 공공 데이터 API 호출 방법
  - LangChain, LangGraph 연동 방법
  - 프롬프트 설정의 중요성
- 이후 개선 방향
  - CRAG(CorrectiveRAG) 적용



출처 : https://arxiv.org/pdf/2401.15884