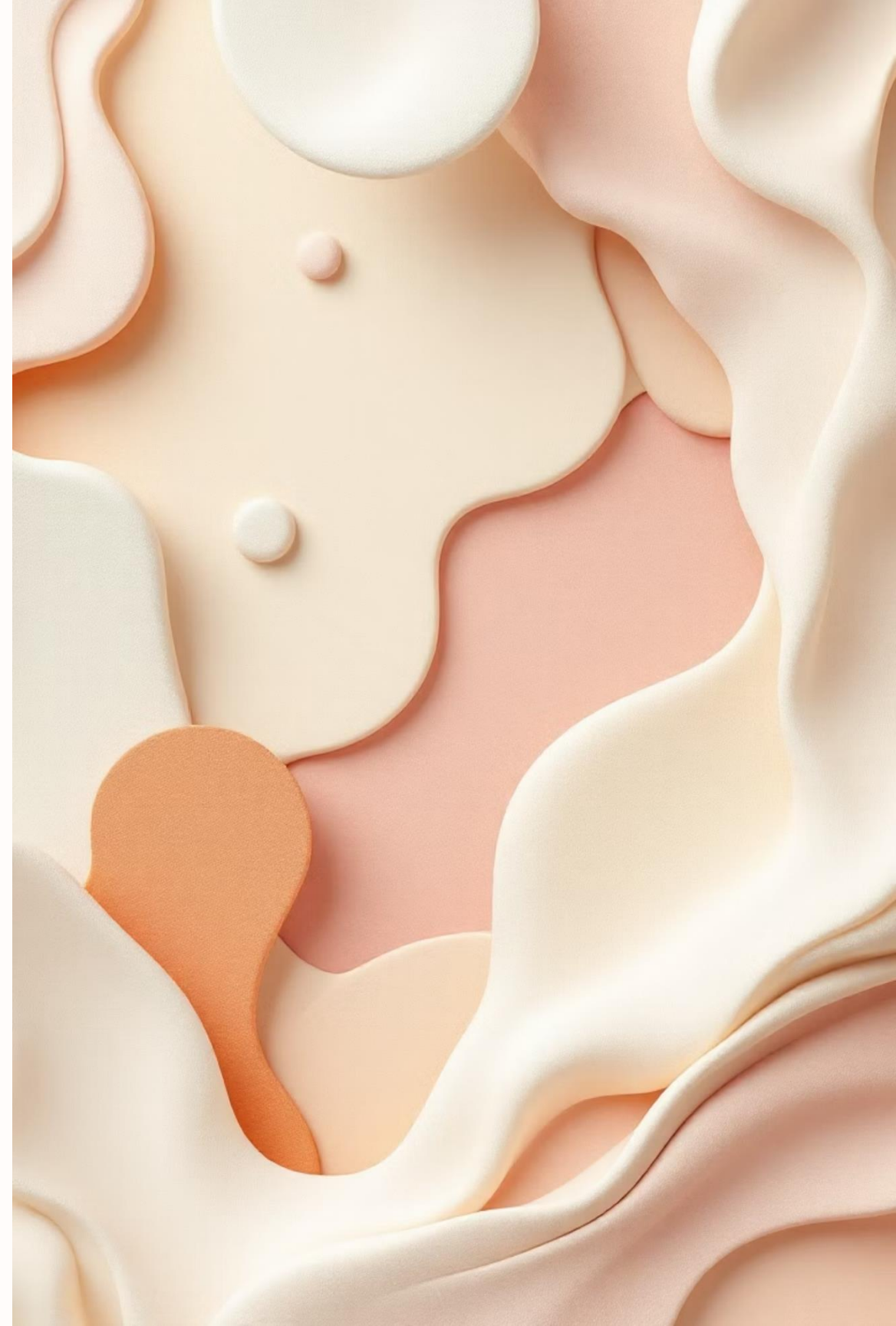


Vision Data 지능형 검색 시스템

CCTV, 바디캠, 모바일 영상 데이터의 혁신적 활용을 위한 AI 솔루션





해결하고자 하는 문제

데이터 검색의 어려움

CCTV, 바디캠, 모바일에서 수집되는 대용량 비디오와 이미지 데이터에서 원하는 콘텐츠를 찾기 어려운 현실

비구조화된 데이터

Vision Data의 특성상 텍스트 기반 검색으로는 한계가 있어 효율적인 활용이 불가능

개발 생산성 저하

AI 개발자와 빅데이터 엔지니어가 필요한 비디오를 찾는데 많은 시간 소요

타겟 사용자

AI 개발자

- 컴퓨터 비전 모델 학습용 데이터 수집
- 특정 시나리오별 영상 데이터 필요
- 빠른 데이터 접근과 품질 검증 요구

빅데이터 엔지니어

- 대용량 Vision Data 파이프라인 구축
- 데이터 전처리 및 품질 관리
- 실시간 스트리밍 데이터 처리



솔루션 개요

01

데이터 수집 및 저장

서비스 시나리오에 맞게 데이터웨어하우스에 유입되는 비디오는, 비디오 전체 요약 / 오디오 대화록 추출 / 1 fps로 각 프레임 장면에 대한 묘사 정보를 체계적으로 저장

02

지능형 인덱싱

AI 기반으로 비디오 콘텐츠를 분석하여 검색 가능한 메타데이터와 시맨틱 정보를 자동 생성

03

챗봇 인터페이스

자연어 쿼리를 통해 원하는 비디오 데이터를 직관적으로 검색하고 즉시 확인할 수 있는 사용자 친화적 인터페이스

기술 아키텍처

LangChain/LangGraph 기반 시스템 구조

1

데이터 수집 레이어

CCTV, 바디캠, 모바일 등으로 수집한 비디오에 대한 분석 및 저장

2

처리 및 분석 레이어

LangChain을 활용하여 비디오 분석 정보를 VectorDB에 저장하고,
RAG 시스템 구축

3

검색 및 응답 레이어

LangGraph 기반 에이전트가 사용자 쿼리를 처리하고 최적화된
결과 제공

핵심 기술 구성요소



지능형 RAG 시스템 구축

수집 데이터 및 ChromaDB를 활용하여 RAG 시스템 구축



Adaptive RAG & Sub-graph

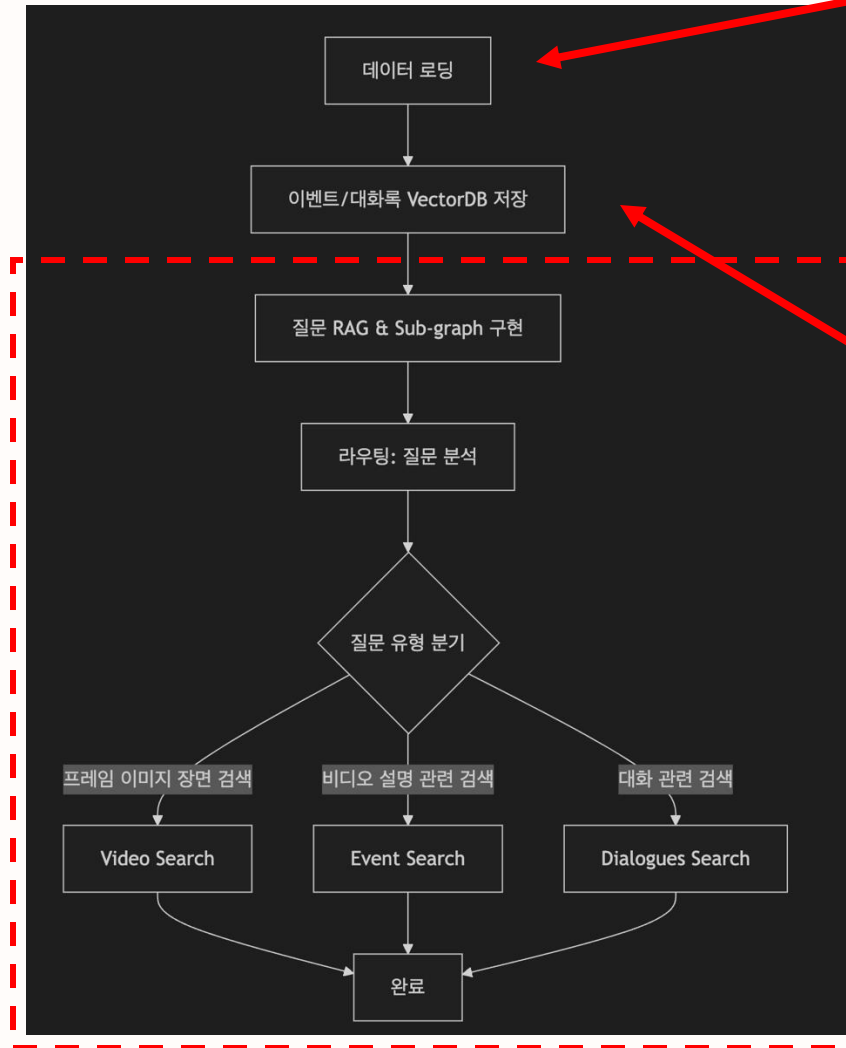
검색 시나리오별 RAG 시스템을 라우팅하고, 기능을 모듈화함



프롬프트 엔지니어링

검색 최적화를 위한 도구(tool) 개발 및 이를 활용한 프롬프트 구현

데이터 흐름도(1/2)



< 워크플로우 디자인 >

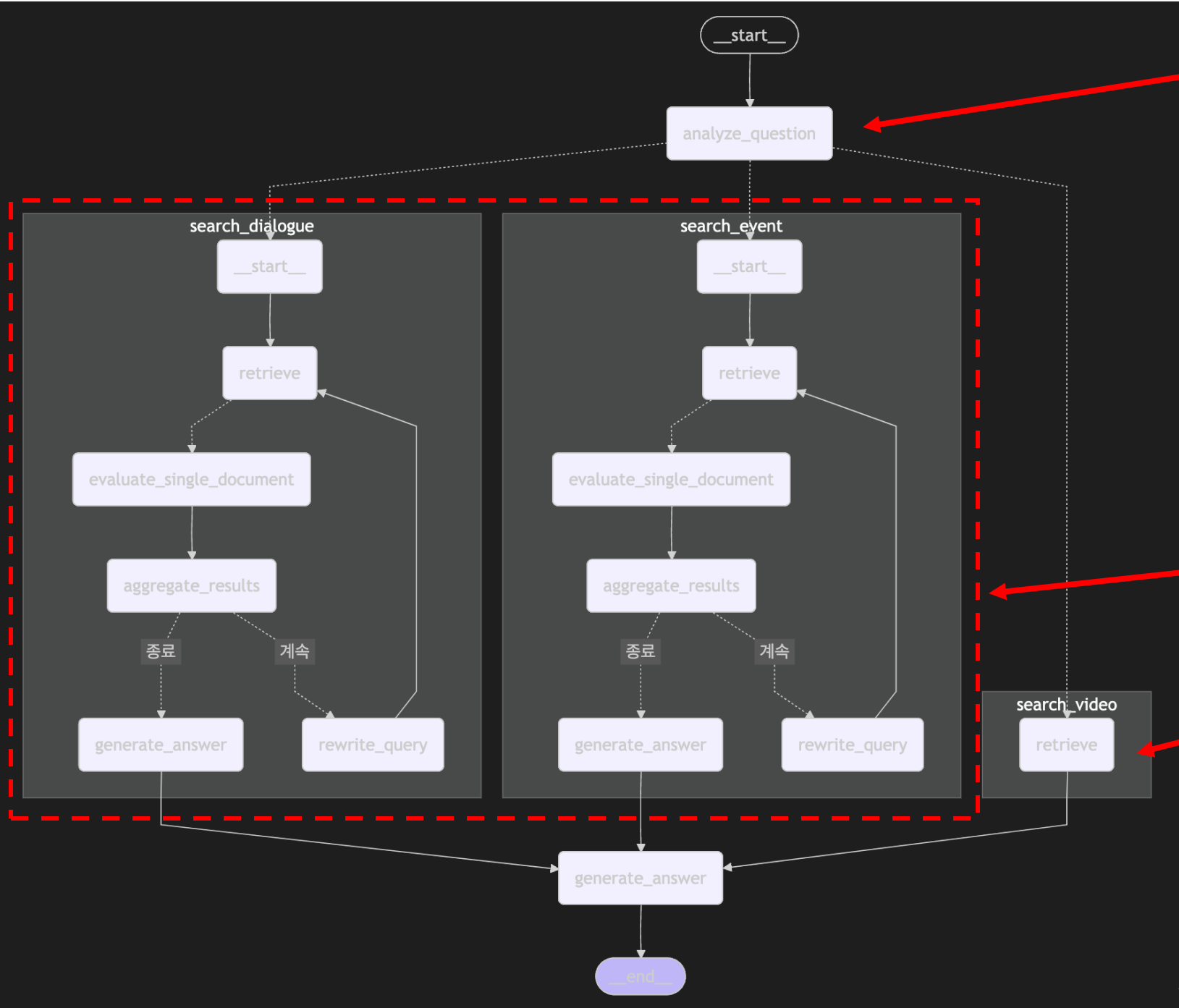
- 챗봇에 사용할 데이터 준비

- data/synopsis_events.txt: 비디오에 대한 전체 요약 정보 저장(15초씩 동영상을 clipping하여 Qwen2.5 VL을 이용하여 클립 영상에 대한 설명 추출)
- data/synopsis_dialogues.txt: 비디오에 포함된 대화 기록 저장(mp3 추출 후 WhisperX 모델을 이용하여 시간별 대화기록 추출)
- Semantic Search APIs: 비디오 프레임 이미지(1fps)에 대한 Image Embedding(OpenClip)을 저장하고 Text-to-Image Similarity Search 지원

- ChromaDB를 사용하여 지능형 RAG 시스템 구축

- 사용자의 질문을 분석하여 적절한 에이전트를 선택 (Adaptive RAG 적용)
- LangGraph & Sub-graph를 사용하여 시나리오 구현

데이터 흐름도(2/2)



- 검색 라우팅 프롬프트

```
...
## 비디오 검색 도구
- *semantic search*: 비디오 속 한 장면에 대한 묘사로 질문
- *event synopsis*: 비디오 전체 상황에 대한 설명으로 질문
- *dialogues*: 비디오 속에서 말하거나, 대화 또는 이야기를 하고 있는 내용에 대한 질문
...
```

- 검색 시나리오별 Agent RAG 구축

- 별도 검색 API(Semantic Search)를 사용하는 도구 추가

< LangGraph 기반 챗봇 구현 >

라이브 데모

실제 시스템 작동 시연

1 특정 프레임에 대한 묘사

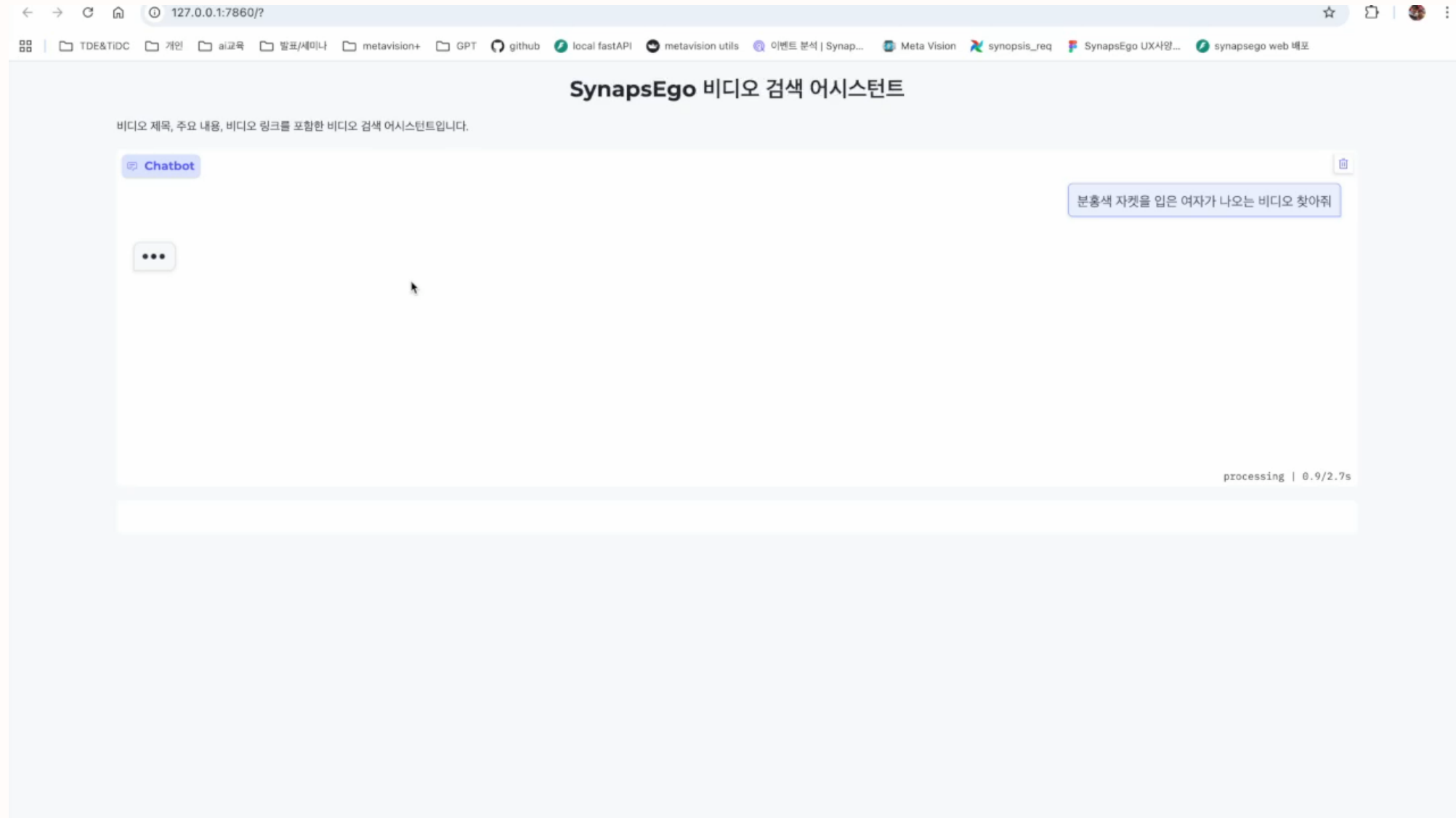
"분홍색 자켓을 입은 여자가 나오는 비디오 검색"

2 비디오 요약 정보

"경찰체포와 관련된 내용이 있는 비디오 검색"

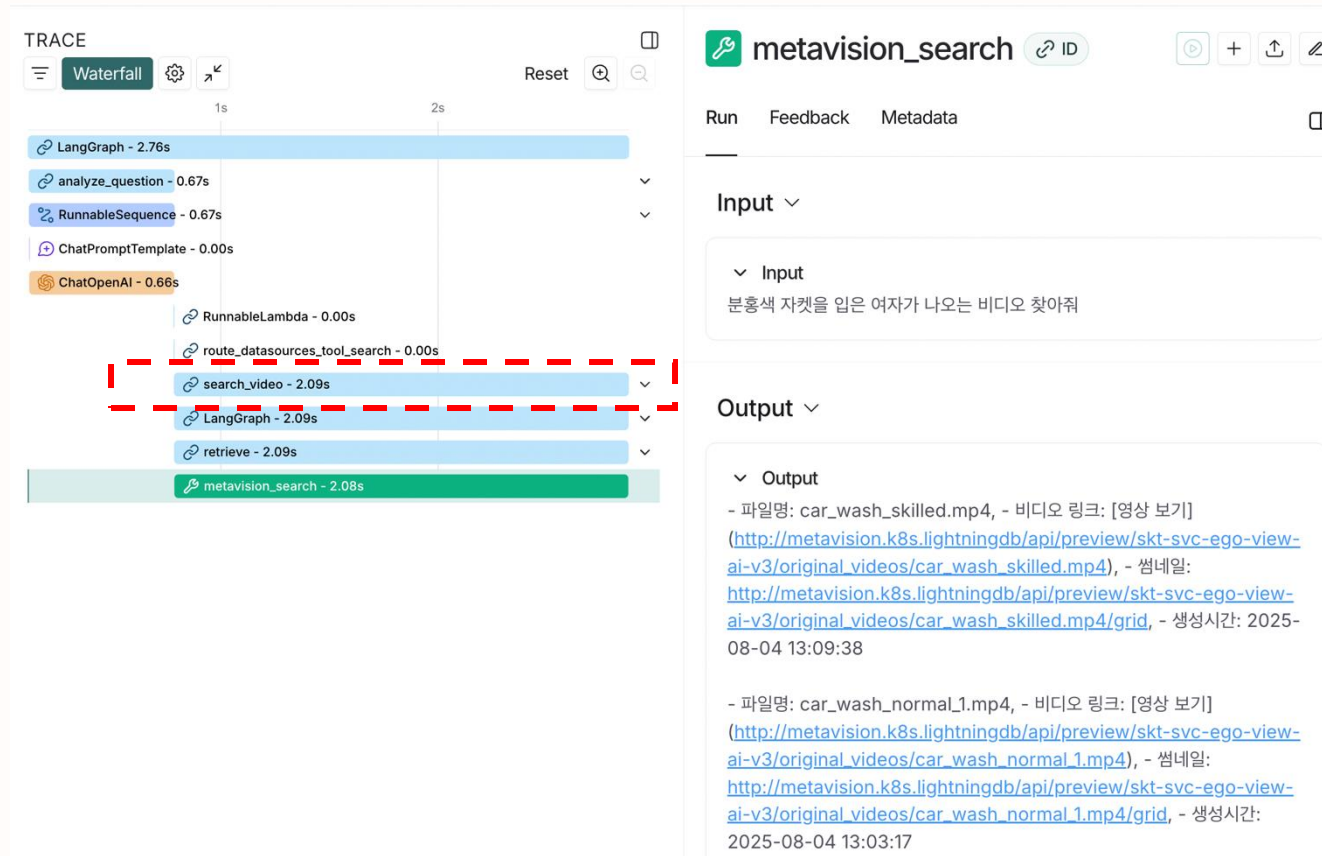
3 대화기록 검색

"GPU에 대해 말하고 있는 비디오 검색"

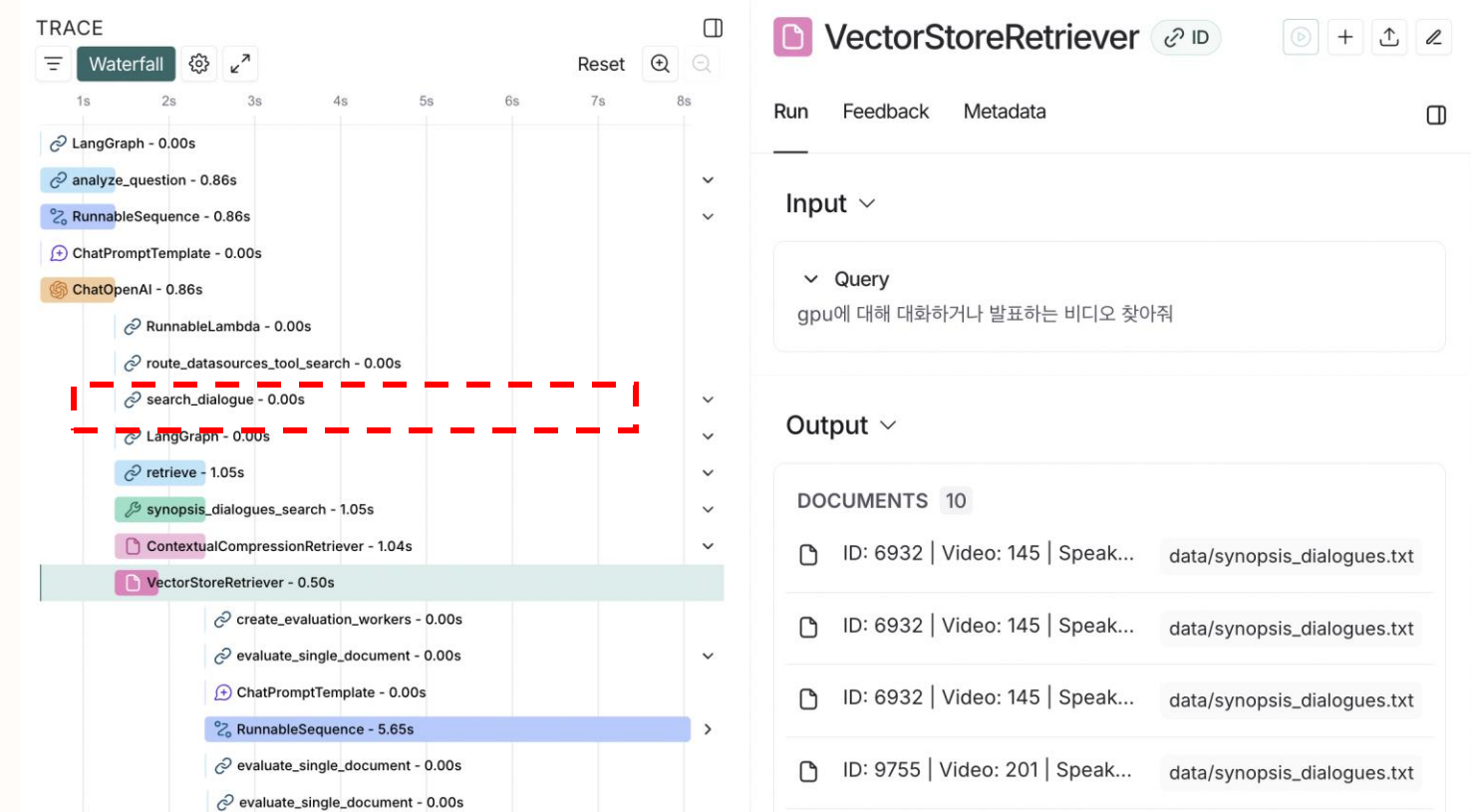


모니터링

LangSmith를 활용하여 모니터링



< 'Search Video' 사용 시나리오 >



< 'Search Dialogue' 사용 시나리오 >

향후 확장 계획

기술적 개선

- 실시간 스트리밍 데이터 처리 최적화
- 다국어 지원 및 음성 검색 기능
- 엣지 컴퓨팅 환경 지원

비즈니스 확장

- 공공안전 분야 특화 솔루션
- 소매업계 고객 행동 분석
- 스마트시티 통합 플랫폼
- 클라우드 SaaS 서비스 제공

"Vision Data의 새로운 패러다임을 제시하며, AI 기술로 더 스마트한 세상을 만들어갑니다"

