

사용자 이미지 기반 AI 챗봇 문제정의서

9조 허정우 김보배

목차

Table of Contents

연구 개발의 필요성

- 주제 설명
- 사용자 관점에서의 문제점

연구 개발의 목표 및 내용

- 핵심 문제
- 아이디어

설문 인사이트

- 설문 내용
- 도출된 인사이트

기대 효과 및 향후 확장 가능성

소개할 내용을 입력해 주세요
포함되는 내용을 나열해 보세요.

연구 개발의 추진전략

- 추진 일정
- 구성 및 역할

주제

사용자 이미지 기반 AI 챗봇

- Graph RAG를 멀티모달 데이터에 적용하기 위하여 특정 도메인을 설정함.
- 사용자 이미지(갤러리)의 사진에 대한 정보를 그래프 형태로 저장, 사용자의 자연어 질문에 대한 응답을 제공하는 챗봇

사용자 입장에서의 문제

- 기존 갤러리 검색은 단순히 시간/위치/라벨링 기반 필터링만 제공한다.
- 사용자는 “작년 생일 때 뭐 먹었더라?” 와 같은 의도 기반 질문을 할 수 없음.
- 사용자는 기억을 떠올리기 위해 수동으로 사진을 훑어야 하는 귀찮음이 존재함.
- 현재 까지 존재하는 대부분 챗봇은 일반 지식 기반 데이터만 제공한다.
- 사용자의 개인의 삶에서 발생한 이벤트를 기억해주는 기능이 필요하다.

연구 개발의 목표 및 내용

- 궁극적인 목표 (TO-BE)

사용자가 자연어로 질문하면, 개인 이미지 갤러리에서 의미 있는 장면을 구조적으로 분석하고
그래프 기반 AI + 멀티모달 검색 기술을 통해 정답에 가까운 응답과 이미지를 제공하는 서비스 개발

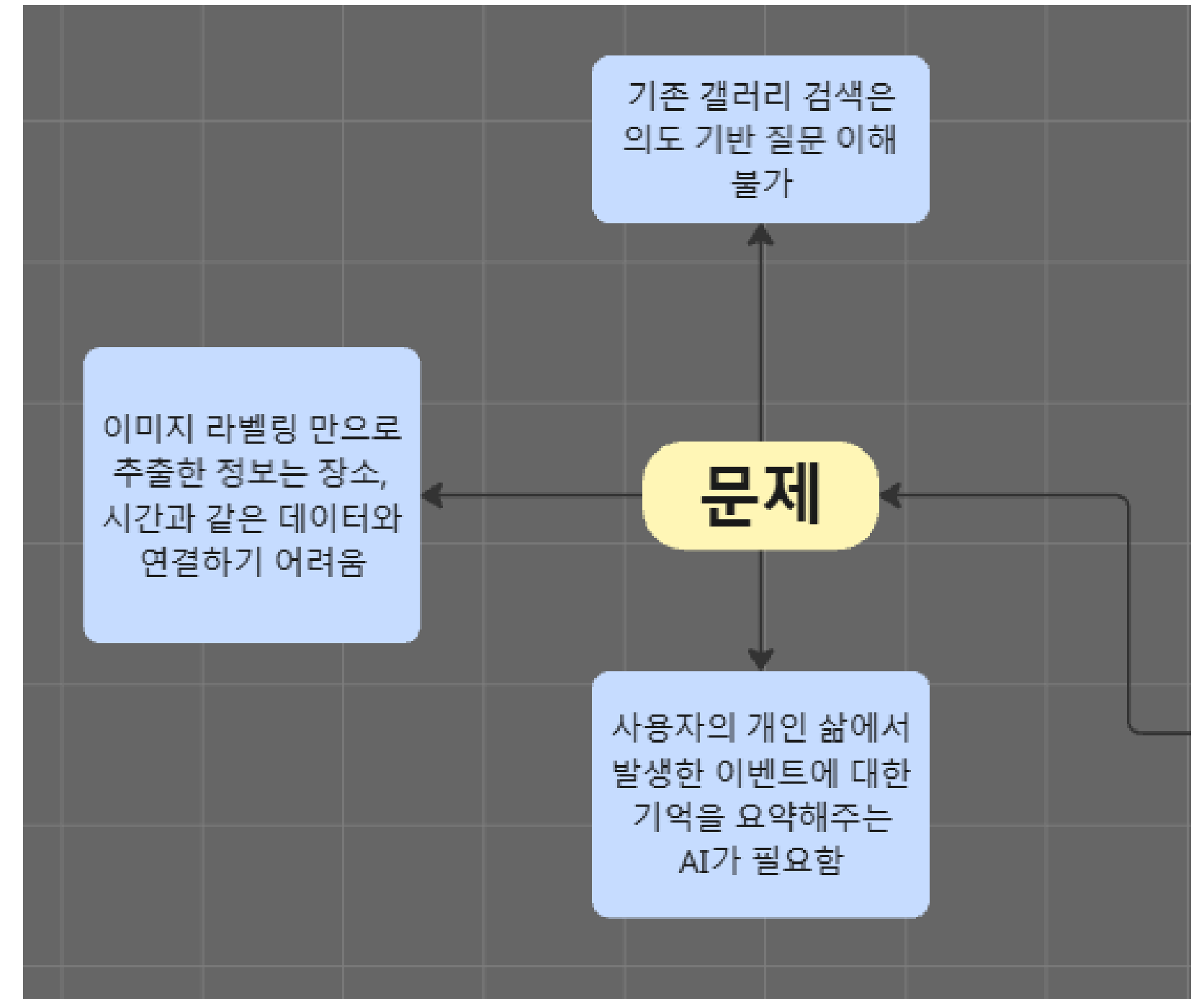
- 멀티모달 GraphRAG의 실제 적용

텍스트 + 이미지 + 시간 + 장소 -> 다양한 modality 연결
멀티모달 노드를 동적으로 생성하고 탐색하는 GraphRAG 구조를 실제 사례에 적용



핵심 문제

- 기존 갤러리 검색은 의도 기반 질문에 대한 대답을 제공할 수 없음. 또 자연어 질문 처리 불가
- 이미지 라벨링으로 추출한 정보 만으로 장소, 시간과 같은 데이터와 연결하기 어려움.
- 사용자의 개인 삶에서 발생한 이벤트에 대한 기억을 요약해주는 도우미가 필요함.



아이디어

- 기존 갤러리 검색은 의도 기반 질문에 대한 답을 제공할 수 없음. 또 자연어 질문 처리 불가

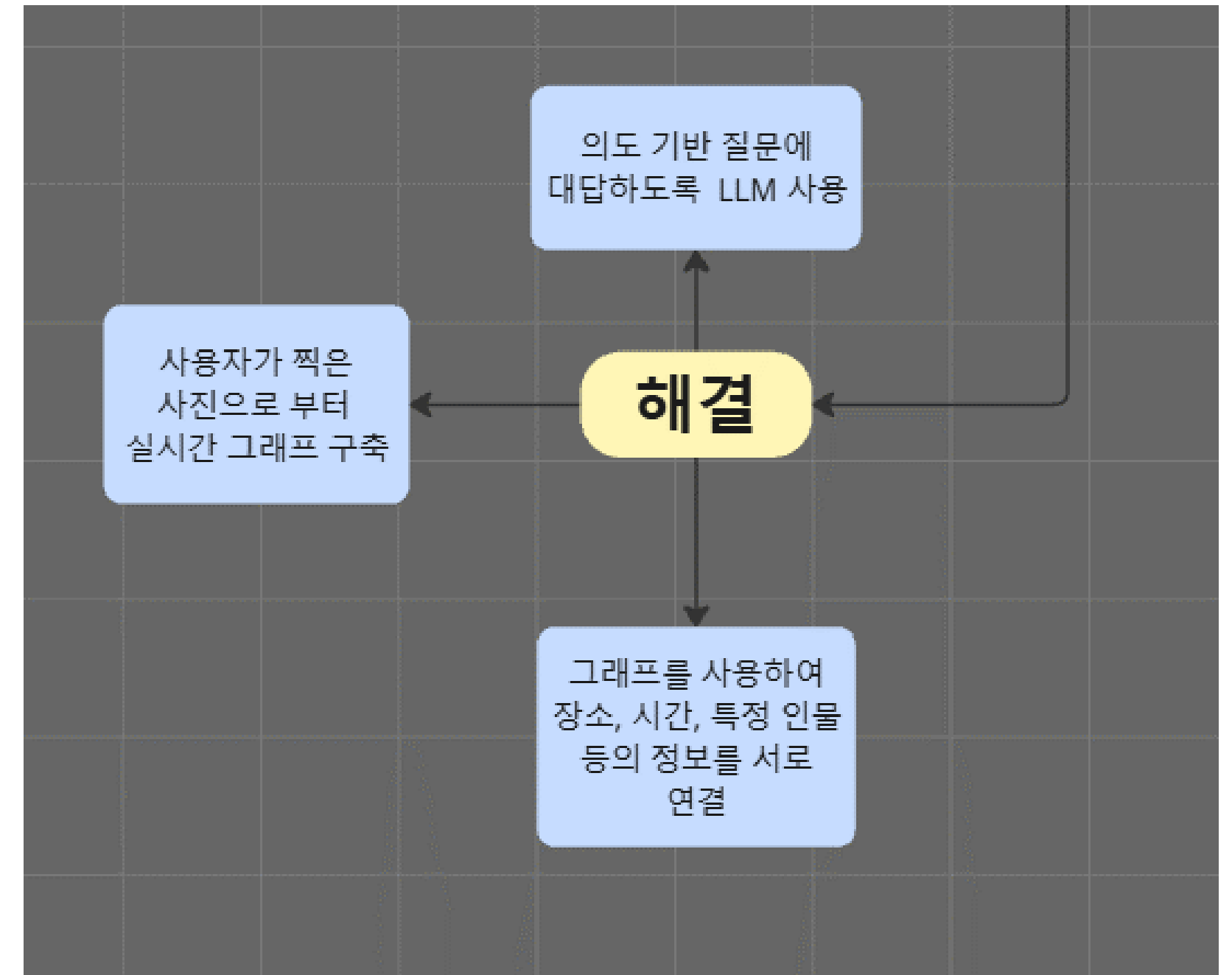
=> 의도 기반 자연어 질문 처리와 답변 생성을 위해 LLM 사용

- 이미지 라벨링으로 추출한 정보 만으로 장소, 시간과 같은 데이터와 연결하기 어려움.

=> 이미지에서 추출한 정보들을 그래프로 구축하여 연결함

- 사용자의 개인 삶에서 발생한 이벤트에 대한 기억을 요약해주는 도우미가 필요함.

=> RAG를 사용하여 사용자 개인에 맞는 데이터 기반으로 답변 생성



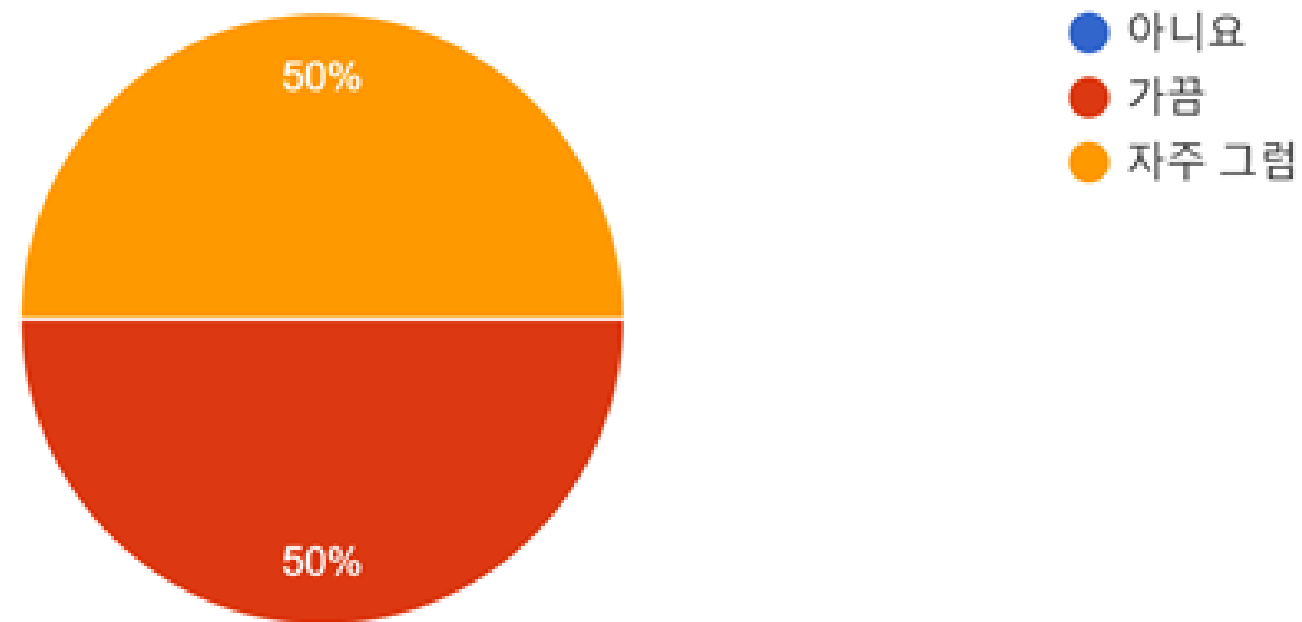
설문 조사

항목	세부내용
기간	4/3 ~ 4/4
목표	<ul style="list-style-type: none">• 사용자가 느끼는 불편함 조사• 현재 제공되는 서비스에 대한 불편함 조사• 새로운 서비스에 대한 의견 도출
조사 도구	Google form 설문
참여 인원 수	14명

주요 질문 및 응답

사진 갤러리에서 과거의 사진을 찾는데 어려움을 느낀 적이 있나요?

응답 14개

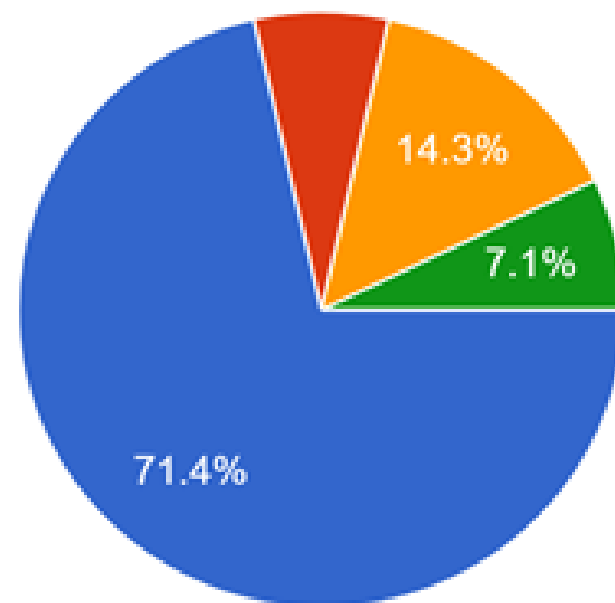


- 응답자 전부가 갤러리에서 사진을 찾는데 어려움을 느낀다고 응답

주요 질문 및 응답

기존 갤러리의 검색 기능을 사용해보셨나요?

응답 14개



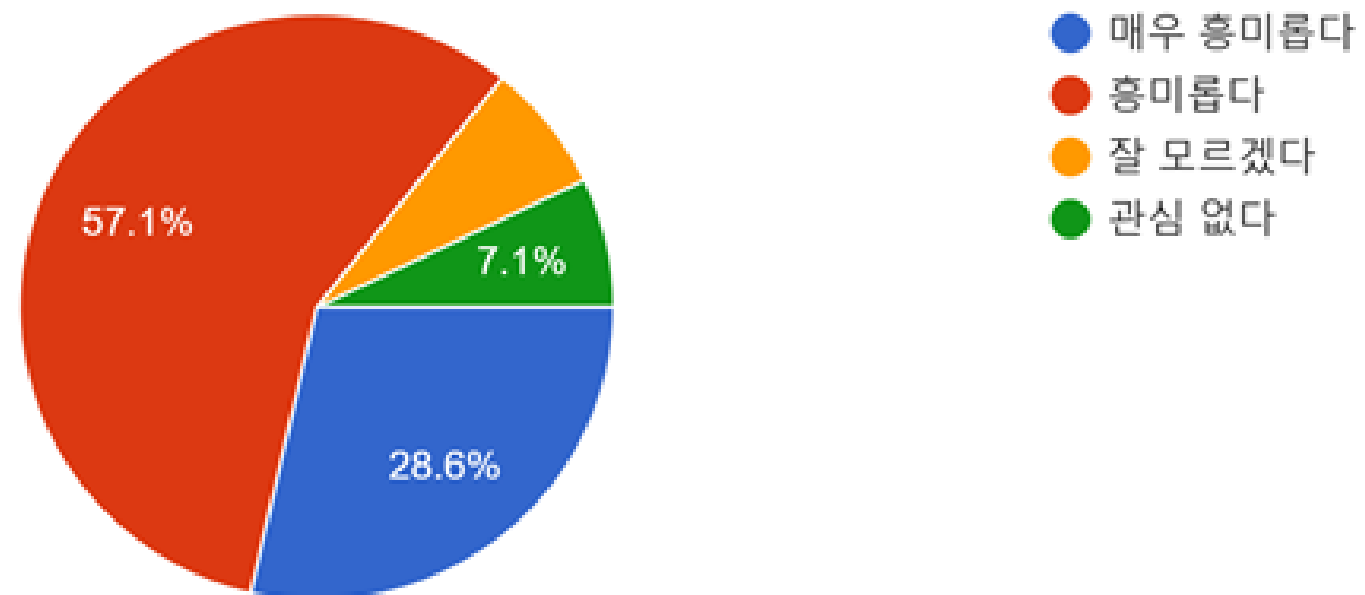
- 있는지도 모름
- 아니요
- 써봤는데 별로임
- 좋아서 자주 사용하고 있다

- 응답자 대부분이 갤러리 검색 기능의 존재를 모른다고 대답.

주요 질문 및 응답

AI가 내 앨범 사진을 기반으로 질문에 대답해준다면 어떨 것 같나요?

응답 14개



- 응답자 대부분이 AI 챗봇에 대해 흥미롭다는 반응이 나타남

기타 의견

- 개인 정보 관련 문제에 대한 지적이 있었다.

기대 효과 및 향후 확장 가능성

- 사용자 관점 기대효과
 - 개인 갤러리에서의 사진 검색 시간 절감으로 기억 회상, 일상 보조 도구로의 활용
 - 개인 사용자의 정서적 만족도 상승 -> 1차 테스트 기준 사용자 만족도 4.0 목표

기대 효과 및 향후 확장 가능성

- 확장 가능성

- 유사 장면 추천 기능

”이런 분위기의 캠핑 사진 또 있어?” 와 같은 질문에 개해 시각적 + 개념적 유사성 기반으로 특정 사진과 비슷한 장면을 자동 추천

- OCR + 텍스트 기반 검색 기능

사진 속 슬씨 (간판, 문서, 포스트잇 등)에 대한 검색 기능으로 “쿠폰 유효기간 끝난거 뭐 있었지?” 와 같은 질문에 대해 사진 내 텍스트를 검색한 후 이에 대한 구체적인 응답을 제공

- 멀티모달 확장

영상, 음성 등의 여러 형태의 데이터를 통합하여 갤러리 이외의 다른 앱과의 연결

- 경량화 LLM + 프라이버시 RAG

온디바이스 동작이 가능한 성능 좋은 LLM 모델이 출시된다면 각 사용자 별 독립 그래프 구성 + 로컬 추론이 가능해져 프라이버시 보호 가능

추진 일정

1. 조사 및 기획 (4월)

- 사용자 요구사항, 필요기능 분석
- 기술 구성 요소 정의

2. 데이터 수집 및 모델 설계 (5월)

- 이미지 기반 객체 인식 및 관계 추론
- Scene Graph 모델 선정 및 학습
- 그래프 저장 구조 데이터베이스 설계 (Neo4j)

3. GraphRAG 구성 및 검색 파이프라인 구축 (5월)

- 질의 처리 파이프라인 개발
- Scene Graph 내 의미기반 탐색 알고리즘 구축
- 초기 자연어 응답 생성

4. 정보 처리 방법 (5월)

- 이미지 정보 추출 방법 탐색 및 논의
- 이미지 정보 추출 및 Graph 구성 방법 탐색
- 그래프 탐색 방법 설정

5. 통합 프로토타입 개발 (6월)

- 전체 시스템 통합
- 유저 질의부터 이미지 응답까지 완성형 흐름 구성
- 테스트 가능한 사용자 UI 구성

6. 테스트 및 평가 (6월)

- 사용자 테스트 진행 및 피드백 수집
- 정확도 응답시간, 만족도 분석
- 오류 및 개선 방안 탐색

역할

- 허정우

사용자 UI/UX 프로토타입 제작
멀티모달 검색 결과 시각화 구현
Scene Graph 생성 알고리즘

- 협업 방식

깃허브 공동 개발 환경 : 코드/데이터 버전 관리
노션 : 일정 및 자료 공유, 회의록 관리
매주 1회 회의 : 진행상황 점검 및 논의

- 김보배

Graph Database 설계 및 구축
GraphRAG 기반 검색 파이프라인 설계 및 구현
질의 분석 및 LLM 응답 생성

감사합니다