

문제정의서(연구개발계획서)

Project Name	사용자 이미지 기반 AI 챗봇(GraphRAG)
-----------------	----------------------------

9 조

201902770 허정우

202002470 김보배

지도교수: 임성수 교수님 (서명)

Document Revision History

REV#	DATE	AFFECTED SECTION	AUTHOR
1	2023/04/04	1,2	허정우
2	2023/04/05	3	허정우
3	2023/04/05	4,5	김보배

Table of Contents

1. 연구 개발의 필요성	5
2. 연구 개발의 목표 및 내용	6
3. 이해당사자 인터뷰/ 설문 인사이드	7
4. 기대 효과 및 향후 확장 가능성	9
5. 연구 개발의 추진전략 및 방법	10
6. AI 도구 활용 정보	11
7. 참고문헌 (REFERENCE)	11

List of Figure

그림 목차 항목을 찾을 수 없습니다.

1. 연구 개발의 필요성

주제: Graph RAG기반 사용자 이미지 데이터 AI 챗봇

GraphRAG를 멀티모달 데이터에 적용시키기 위해 문제를 축소하여 사용자 이미지에 대하여 정보를 추출하고 사용자의 자연어 질문에 응답을 주는 챗봇

- 사용자 입장에서의 문제

1. 기존 갤러리 검색의 한계

- 대부분의 갤러리 앱은 단순히 시간/위치 기반 필터링만 제공한다.

기존의 검색 방식은 메타데이터나 단순 태그 기반으로 동작하여 사람·장소·이벤트 간의 의미 있는 관계를 탐색하거나 자연어 질의에 대응하기 어렵다.

사용자는 "지난주 주말에 놀러가서 뭐 먹었더라?" 같은 의도 기반 질문을 할 수 없고, 수동으로 사진을 훑어야 한다. 자연어로 질문하고 자동 응답받는 인터페이스는 현재 상용 갤러리 앱에서 구현되어 있지 않다.

2. 개인 맞춤형 정보 요약의 부재

- 현재까지의 대부분 챗봇은 일반 지식 기반이다. 사용자 개인의 삶에서 발생한 이벤트 (예: 지난주 식사, 여행, 활동 등)를 기억하고 요약해주는 기능이 필요하다.

- 관련 서비스

- 현재 Google에서 Ask Photos라는 비슷한 서비스를 제공하고 있으나, 아직 정식 서비스는 아니고 일부 유저들에게만 키를 발급해주는 형태로 제공중이다. 이 서비스는 장소가 미국이어야만 사용 가능하다는 점과 성능이 좋지 않다는 해외 사용자들의 평가가 있다.

2. 연구 개발의 목표 및 내용

■ 궁극적인 목표 (TO-BE):

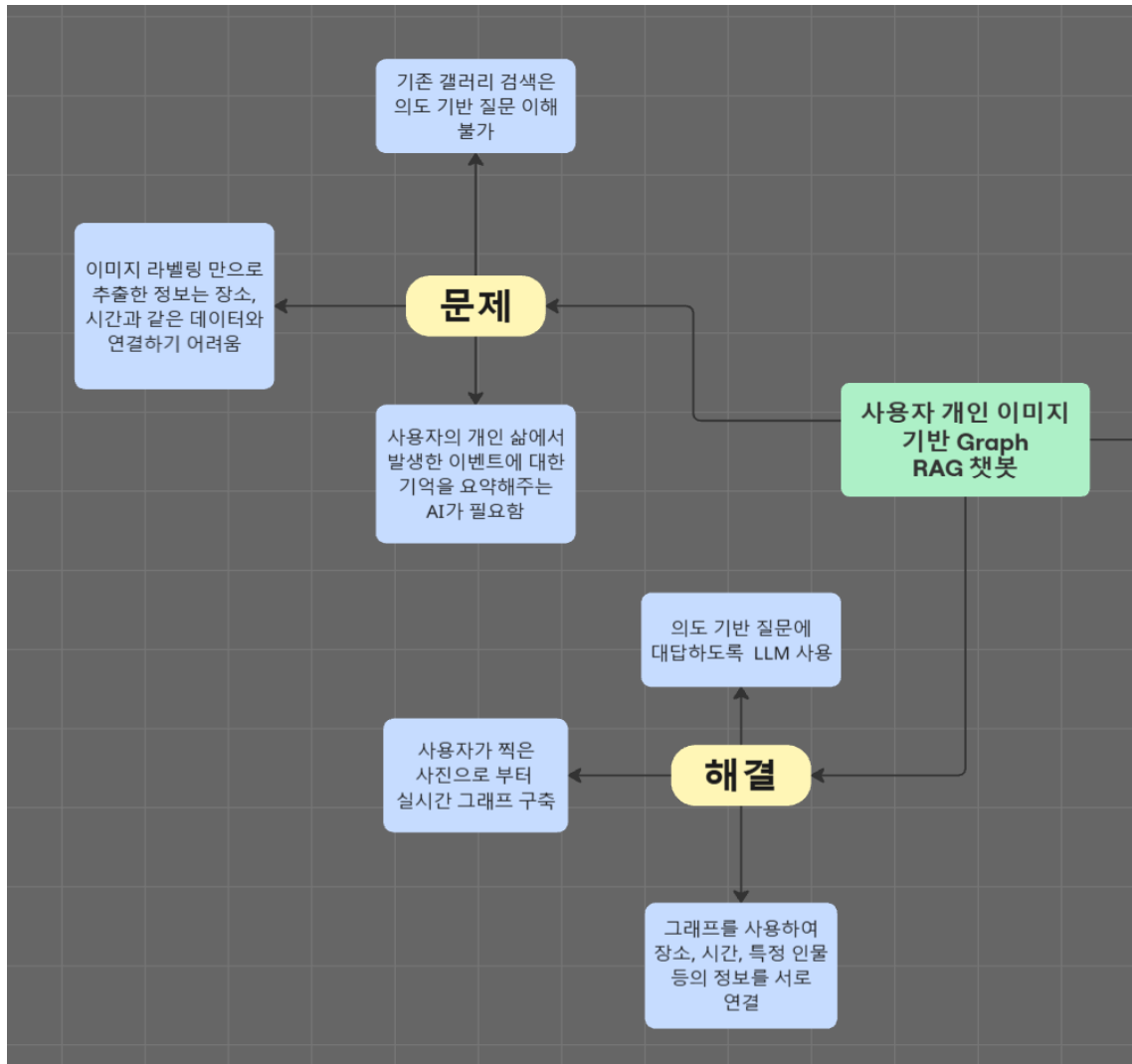
사용자가 자연어로 질문하면, 개인 이미지 갤러리에서 의미 있는 장면을 구조적으로 분석하고 그래프 기반 AI + 멀티모달 검색 기술을 통해 정답에 가까운 응답과 이미지를 제공하는 서비스 개발.

■ 멀티모달 GraphRAG의 실제 적용

텍스트 + 이미지 + 시간 + 장소 → 다양한 modality 연결

멀티모달 노드를 동적으로 생성하고 탐색하는 GraphRAG 구조를 실제 사례에 적용

- 핵심 문제
 1. 의도기반 자연어 질의에 대응 가능한 이미지 검색 구조 설계
 2. 이미지 라벨링 만으론 장소, 시간, 이미지에서 추출한 정보간의 연결이 어렵다.
-> 장면 속 객체, 관계, 맥락을 분석하는 Scene Graph 자동화
 3. 기억 기반 탐색을 통한 사용자 개인화가 필요함(ex. “작년 단풍 구경 간 날”)
- 해결을 위해 적용해 볼 방법
 1. 의도기반 자연어 질문 처리와 답변을 위해 LLM 사용
 2. 그래프를 사용하여 이미지에서 추출한 정보들을 연결하여 정확한 답변 제공
 3. 사용자가 찍은(업로드한) 사진으로부터 실시간 그래프 구축
 4. LLM에 RAG를 사용해 사용자 개인에 맞는 데이터 기반으로 답변 가능.



3. 이해당사자 인터뷰/설문 인사이드

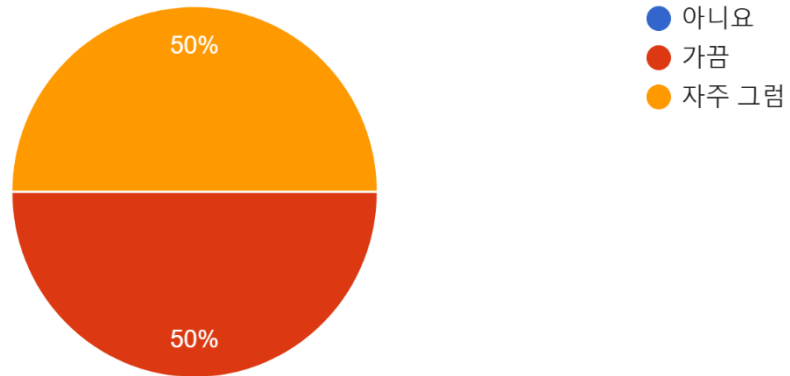
1. 이해당사자 설문

- 방식: 설문 조사 진행
- 기간: 4/3 ~ 4/4
- 목표: 현재 제공되는 서비스에 대한 불편함 조사, 새로운 서비스에 대한 의견 도출
- 조사 도구: google form
- 참여 인원 수: 14

2. 주요 질문 및 응답 요약

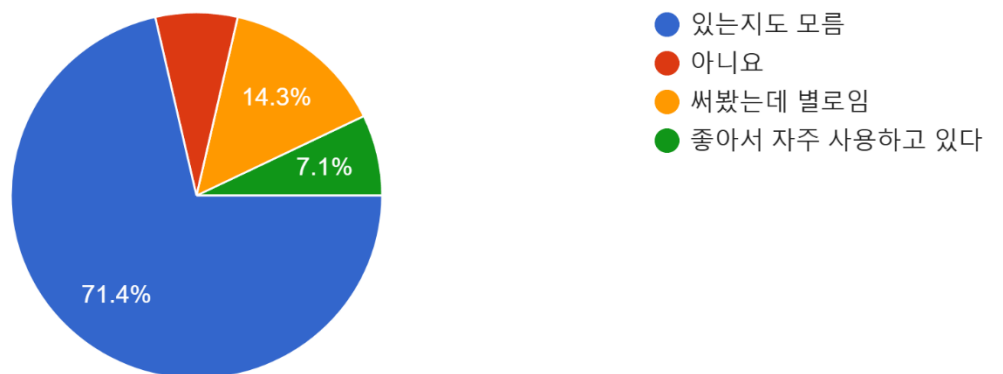
사진 갤러리에서 과거의 사진을 찾는데 어려움을 느낀 적이 있나요?

응답 14개



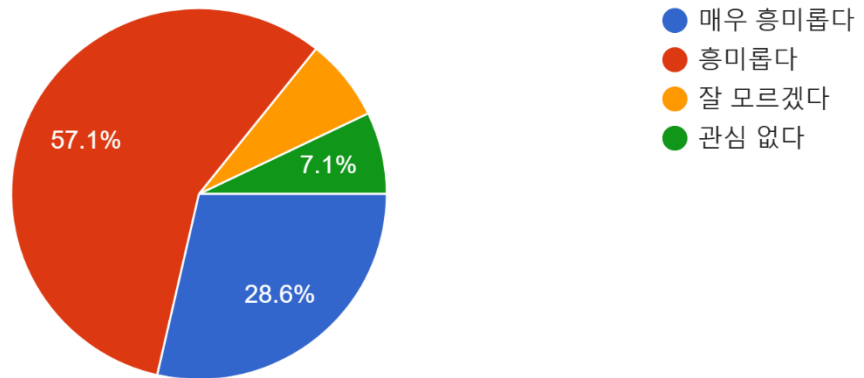
기존 갤러리의 검색 기능을 사용해보셨나요?

응답 14개



AI가 내 앨범 사진을 기반으로 질문에 대답해준다면 어떨 것 같나요?

응답 14개



3. 정리

- 응답자 대부분이 앨범 사진을 활용하는데 어려움이 있으며 기존 갤러리 검색 서비스의 존재를 모르거나 성능에 아쉬움을 느꼈다.
- 새로 제공될 서비스에 대해 흥미가 있다는 응답자가 많았다.
- 우려 사항으로는 개인 정보 관련 문제가 있었다.

4. 기대 효과 및 향후 확장 가능성

■ 사용자 관점 기대 효과

개인 갤러리의 사진 검색 시간 절감으로 기억 회상, 일상 보조 도구로의 활용

개인 사용자의 정서적 만족도 상승

-> 1차 테스트 기준 사용자 만족도 4.0 이상 목표

■ 확장 가능성

1. 유사 장면 추천

- “이런 분위기의 캠핑 장소 더 있어?”와 같은 질문에 대해 시각적 + 개념적 유사성 기반으로 특정 사진과 비슷한 장소를 자동 추천

2. 멀티모달 확장

- 현재는 사진 중심이지만 영상, 음성 등도 통합하여 다른 앱들과 통합 가능

3. OCR + 텍스트 기반 검색

- 사진 속 글씨(간판, 문서, 포스트잇 등)에 대한 검색 기능으로 쿠폰 유효기간 끝난 거 뭐 있었지?” 와 같은 질문에 대해 사진 내 텍스트를 검색한 후 이에 대한 구체적인 응답을 제공하는 서비스

4. 경량형 LLM + 프라이버시 RAG

- 온디바이스 동작이 가능한 성능이 좋은 LLM 모델이 출시 된다면 각 사용자별 독립 그래프 구성 + 로컬 추론이 가능해져 완벽한 프라이버시 보호 가능

5. 연구 개발의 추진전략 및 방법

1. 조사 및 기획(4월)

- 사용자 요구사항, 필요 기능 분석
- 기술 구성 요소 정의

2. 데이터 수집 및 모델 설계(5월)

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

3. 검색 파이프라인 구축(5월)

- 질의 처리 파이프라인 개발
- 초기 자연어 응답 생성

4. 정보 처리 방법(5월)

- 이미지에서 정보를 추출하는 여러 방법 논의
- 이미지에서 정보를 추출하고 Graph에 저장하는 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, 응답 시간, 만족도 분석
- 오류 및 개선 방안 탐색

- 역할
 - 허정우: 사용자 UI, 이미지 데이터 처리
 - 김보배: graphDB 설계 및 구축, 질의 및 응답 생성 파이프라인 설계
- 협업 방식
 - Github: 코드 데이터 버전관리
 - Notion: 일정 및 자료 공유
 - 매주 1회 회의: 진행상황 점검 및 논의

6. AI 도구 활용 정보

사용 도구	GPT-4
사용 목적	문장 다듬기
프롬프트	<ul style="list-style-type: none"> ● 문장을 다음어줘
반영 위치	전반적인 부분
수작업	있음
수정	

7. 참고문헌(Reference)

EGTR: Extracting Graph from Transformer for Scene Graph Generation(Jinbae Im, JeongYeon Nam, Nokyung Park, Hyungmin Lee, Seunghyun Park)