

전북인복지 기반 초개인화 서비스 및 고도화 챗봇 개발

2025학년도 2학기
SW 캡스톤디자인
경진대회

팀 명 새고침
지도교수 김윤경

팀 원 배지혁(컴퓨터공학부, 4), 이윤진(컴퓨터인공지능공학부, 3), 백은지(컴퓨터공학부, 4)
산업체 전북특별자치도 사회서비스원

개발 동기 및 목적

개발 동기

- 기존 챗봇의 한계: 전북인복지 포털의 기존 챗봇 '복댕이'는 Fine-tuning 기반으로 구축되어, 최신 복지 데이터가 반영되지 않고 사용자의 구체적인 질문 의도를 파악하지 못해 Hallucination을 하는 문제가 심각함.

- 복지 정보의 불편화: 전북 14개 시군의 복지 정보가 분산되어 있어, 도민들이 자신에게 맞는 혜택을 찾기 위해 여러 사이트를 직접 탐색해야 하는 어려움 존재.

개발 목적

- 초개인화 상담: 사용자 상황(연령, 거주지, 상태)을 분석하여 맞춤 복지 혜택 필터링.
- 시설 연결: 단순 정보 안내를 넘어, 실제 주변 시설(병원, 복지관) 지도 안내.
- 접근성 & 운영 효율: 음성(STT/TTS) 지원 및 비개발자도 관리 가능한 자동화 파이프라인 구축.

주요 기술

Frontend

- React
- Vite
- Tailwind
- Vanilla JS

Backend

- FastAPI
- LangChain
- APScheduler
- SSE

AI/Embedding model

- gpt-oss-20b(LLM)
- BGE-M3 (Embedding)

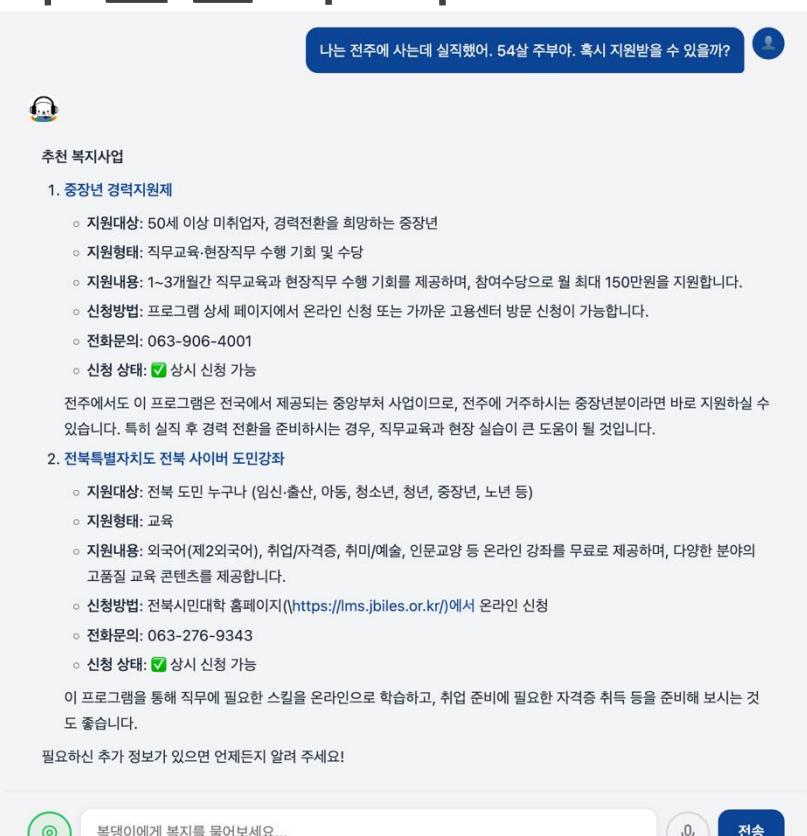
Infra & DevOps

- Environment: Docker Compose (Multi-container)
- Proxy/SSL: Nginx (Reverse Proxy), Let's Encrypt
- Model Serving: OpenRouter API / Ollama

개발 내용

Agentic RAG 기반 지능형 검색 시스템 구축

- Intent Classification: 사용자 질의를 LLM이 실시간으로 분석하여 6가지 의도(복지 검색, 시설 찾기, 통계, 일상 대화 등)로 분류하고, 최적의 검색 전략을 동적으로 실행하는 에이전트 아키텍처 구현.
- 지능형 쿼리 추출 및 하이브리드 검색: LLM이 복잡한 자연어 질의에서 핵심 키워드와 필터 조건(지역, 대상, 연령)을 스스로 추출·보정하며, 이를 벡터 유사도 검색과 키워드 매칭에 동시에 적용하는 하이브리드 검색 엔진 구축.



실시간 데이터 파이프라인

- 자동화된 크롤링 및 임베딩: APScheduler 기반의 데이터 파이프라인을 구축하여 전북인복지 포털의 복지 정보와 시설 데이터를 매일 새벽 자동으로 Crawling, Chunking, 벡터 임베딩 과정을 수행하여 기존 파인튜닝 모델의 데이터 시차 문제 해결.

GIS 기반 시설 안내

- 위치 기반 서비스 (LBS): PostgreSQL의 earthdistance 확장을 활용, 사용자 위치 좌표(GPS) 기준 반경 내 사용자에게 적합한 복지 시설을 거리순으로 정렬하여 안내하는 로직 구현.

결과 및 분석

활용 방안

- 간편한 도입 및 확장: 별도의 서버 구축 없이 스크립트 한 줄 추가만으로 기존 웹사이트에 통합 가능한 형태이며, 향후 도내 타 공공 서비스(일자리, 문화 등)로 범용적 확장 가능.

- 지속 가능한 운영: 자동화된 데이터 파이프라인을 통해 복지 정보를 실시간으로 업데이트하는 방식으로 운영 관리가 용이해짐.

기대 효과

- 도민 편의성 극대화: 복잡한 행정 용어를 몰라도 자연어 대화만으로 본인에게 맞는 복지 혜택과 인근 시설 정보 획득 가능.
- 행정 업무 효율성 증대: 24시간 자동 응답 시스템이 단순 반복 민원을 처리함으로써, 담당 공무원 및 직원의 업무 부담을 경감하고 행정 처리 속도 개선.
- 기술적 혁신: 공공 복지 분야에 최신 AI 기술(Agentic RAG)을 적용하여, 기존 파인튜닝 방식의 한계(최신성 부족, 환각 현상)를 극복한 성공적인 선례 제시.