

건설 SLM(Small Language Model) 개발을 통한 건설현장 품질관리 시스템 및 표준 상용화모델 구축

https://github.com/Yoo-SH/twinsolution_co



2025학년도 2학기 SW 캡스톤디자인 경진대회

팀명 트리플메이저
지도교수 고광신

팀원 산업체

우상수(컴퓨터공학부, 4), 안태윤(컴퓨터공학부, 4), 유승현(컴퓨터공학부, 4)
조진아(문헌정보학과, 4), 황도훈(국제이공학부, 4)
트윈솔루션

개발 동기 및 목적

개발 동기

- 건설 현장의 복잡하고 방대한 법규 및 매뉴얼 정보(1,000개 이상)로 인해 발생하는 휴면 에러와 업무 병목 현상(정보 검색 시간 평균 60분)을 해결하고자 함

목적

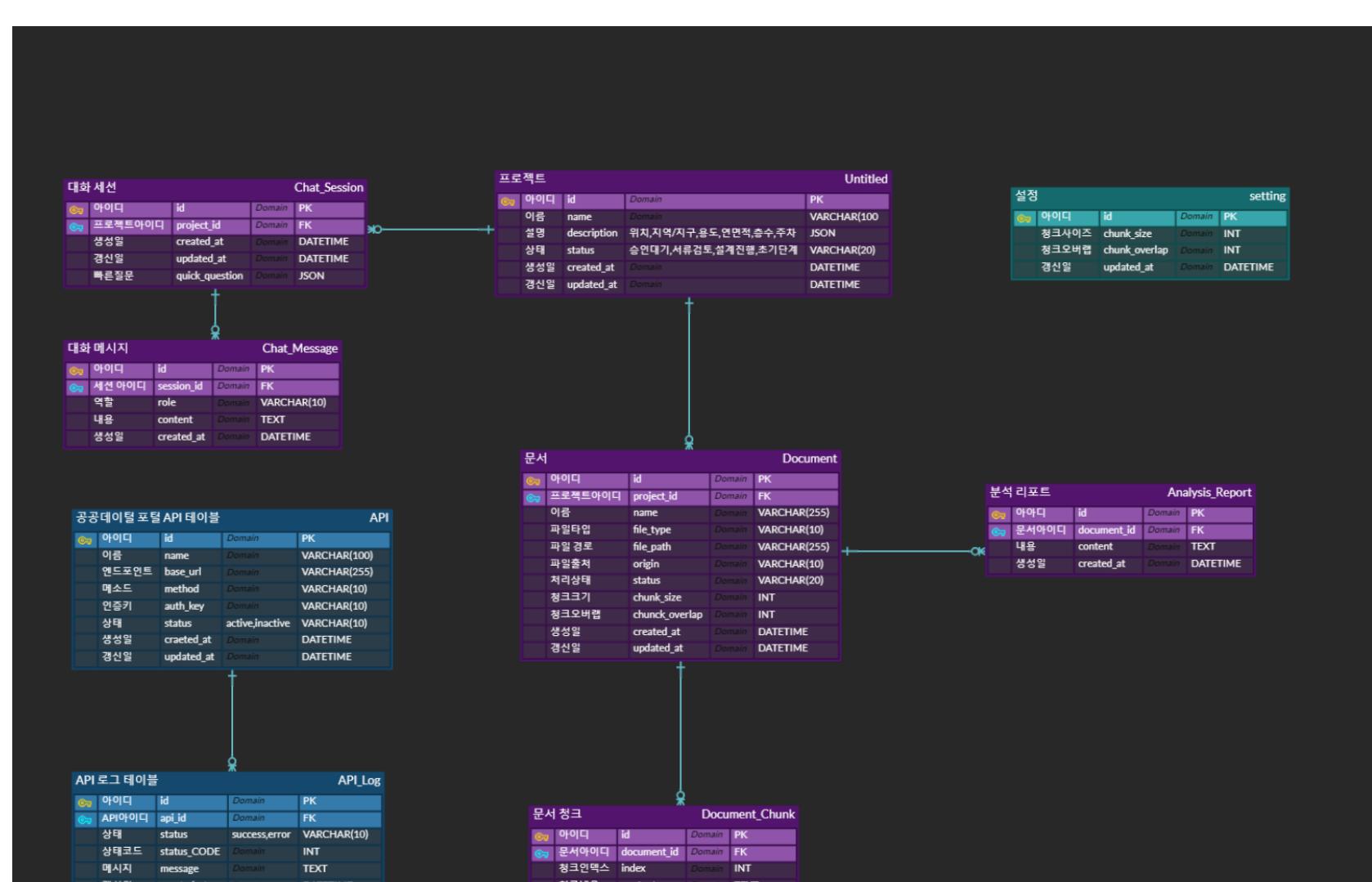
- 건설 도메인에 특화된 7B 파라미터 이하의 경량화 언어 모델 (SLM)을 개발하고, 이를 모바일 환경에서 신뢰성 있게(정확도 90% 이상) 구동하는 시스템을 구축하고자 함

제안 배경



개발 내용

- 데이터 파이프라인 구축: HWPX 등 국내 문서 형식의 텍스트와 메타데이터를 안정적으로 추출하는 전처리 파이프라인을 구축함
 - 도메인 모델 학습: 구축된 건설 질의응답 데이터셋을 활용하여 SLM을 도메인 특화 학습시킴
 - RAG 시스템 통합: 벡터 데이터베이스 및 재랭킹 모듈을 SLM과 통합하여 신뢰성 있는 답변 생성을 구현함
 - 현장 맞춤형 UI 개발: 건설 현장 실무자를 위한 직관적인 모바일 채팅 기반 UI를 개발하고, Firebase를 통해 인증 및 채팅 기록 관리 기능을 연동함



주요 기술

- 신뢰성, 효율성, 현장 가용성을 동시에 확보하기 위해 'Triple Major' 전략을 적용함
 - Construction SLM: 건설 용어 및 법규에 최적화된 경량 모델 (6.5B 파라미터 양자화 모델 기반)을 LoRA 기법으로 Fine-tuning 함
 - RAG System (검색 증강 생성): 사내 비공개 문서(HWPX, PDF 등)를 기반으로 최신 정보를 실시간 참조하여 답변의 신뢰성을 확보함
 - 법률의 ‘조’, ‘항’ 단위 청킹 및 재랭킹 기법을 적용함
 - Mobile / Edge Deployment: 네트워크 단절 상황에서도 사용 가능하도록 온디바이스(On-device) 또는 Edge 서버 구동을 목표로 모델을 최적화함



결과 및 부서

- 목표했던 성능 지표를 성공적으로 달성함
 - 평균 응답 시간은 2.8초로 실시간 사용성을 확보했으며, 질의응답 정확도는 92.5%를 기록함. 또한, 모바일 환경 및 네트워크 제약 하에서도 시스템의 정상 작동을 확인함
 - 복잡한 규정 검색 시간을 5분 이내로 단축시켜 현장 업무 효율을 획기적으로 개선
 - RAG를 통한 규정 준수 실시간 확인으로 휴먼 에러를 최소화하고, 결과적으로 안전 사고 위험을 줄이는 데 기여할 수 있음
 - 건설 도메인 특화 및 HWPX 처리 기술을 통해 국내 건설 산업의 디지털 전환 시장에서 강력한 경쟁 우위를 확보

