실증적SW개발프로젝트 주간보고 (##주차)

작성일: 2025/04/06 팀명: chAt & me

|  |  |  |
| --- | --- | --- |
| 팀 활동  보고 | 활동일시 | 2025.04.01 |
| 장 소 | 동아대 한림도서관 그룹스터디실 |
| 참석자 | 이채영, 김나영, 박수빈, 원민 |
| 특이사항 | - |
| 이번주  진행사항 | 1. 개발내용  **Ollama 설치** - 환경 구성 완료  **Open UI 설치** - 초기 세팅 완료  외부 공개 Dataset 조사 - BenchLLM (회사 내부 Dataset 공개불가)  **멘토 교수님과의 미팅**  - 외부 공개 Dataset BenchLLM으로 다른 논문을 쓰길 권하심  - 회사 내부 Dataset 제공, 석유선박관련 도면  **팀 회의 및 발표 피드백**  - 문제점 재정의 및 해결방안 모색, 향후 방향 재설정  - 매주 2~3장 분량 PPT로 모든 조원 발표 연습  2. 팀원별 활동내용  - 이채영(팀장) : Ollama, Open web ui 작업환경 준비 진행, 교수님과 미팅 주도, github fork  - 김나영(팀원, 프론트엔드담당) : Ollama, Open web ui 작업환경 준비 진행, RAG 논문 조사 및 해석  - 박수빈(팀원, 백엔드 담당) : Ollama, Open web ui 작업환경 준비 진행, LLM 논문 조사 및 해석  - 원 민(팀원, 인공지능 담당) : Ollama, Open web ui 작업환경 준비 진행, LangChain 논문 조사 및 해석 | |
| 개발계획 대비 진행현황 |  | |
| 다음주  계획 | * 공공 데이터로 테스트   OCR 기능(API 또는 모델 상관없음)을 이용해서 PDF에 있는 글자 정보들을 Table 형태로 변환  변환된 정보를 바탕으로 LMM을 학습시켜 RAG(TAG)데이터로 변환  변환시킨 정보를 학습시키고난 전, 후를 비교   * 논문 조사 내용 각자 PPT 발표 * LMM 위주의 Benchmark Dataset으로 baseline 환경준비 (Finetune된 모델의 성능 확인용 및 기업데이터 외부공개 불가 사유) * 현재 데이터 형식의 PDF 파일을 OCR을 이용하여 변환시킨뒤 RAG 또는 TAG로 사용할 수 있도록 변환함. * OCR 모델 또는 간단한 API를 이용하여 구현 * Baseline Knoledge를 공부하기 위해 각자 Paper를 공부하고 Paper Reading을 공유하여 진행함 (LLaMA, MiniGPT) | |
| 주요 결과물 | * Ollama 및 Open UI 설치 완료 보고 * AI 선정 및 학습 중 * 논문 및 프레임워크 조사 자료, 발표자료 | |