

실증적SW개발프로젝트 주간보고 (9주차)

작성일: 2025/05/04 팀명: 작고맵 AI

팀 활동 보고	활 동 일 시	2025.05.02 (금)
	장 소	부민캠퍼스 종합강의동 B04(A)-1405(14층)
	참 석 자	김민수, 박지민, 임성민, 장희웅(전원)
	특 이 사 항	멘토 교수&미컴과 학생 미팅, 데이터 컬럼 확정 및 평가 기준 확립
이번주 진행사항	1. 팀 공통활동내용	
	멘토링 내용	
	(주)슈퍼히어로의 PR 데이터 외에도 밈, 트렌드 등을 반영할 수 있는 방안이 필요	
	언어유희/언어이중성/맥락 등을 고려하여 실제 PR에 활용할 수 있는 고급 한국어 언어모델 구축이 필요	
	* AISAC과 같은 외부 문서 데이터를 파인튜닝에 활용할 수 있음	
	PR 결과 평가 기준 확립	
	기본적으로 (주)슈퍼히어로의 PR 콘텐츠는 종합 점수와 5가지 세부 요소로 평가한다	
	문제해결형 0-100점: 프로젝트 진행 주체자(정부, 기관, 기업)의 니즈 부합. 정보중심의 인지적 요소 부각, 신뢰성/진정성/충성도 기반한 컨셉	
	트렌드반영형 0-100점: 소비자/참여자의 니즈 부합. 시대와 세대(나이, 집단)의 집단인식 부각, 언어 민감성/언어유희적 요소 반영	
	참여유도형 0-100점: 주체자와 소비자의 니즈 동시 부합. 영향력 관점에서 구체적 목표와 결과물이 동시에 부각. 소비자/참여자/구성원과의 상호작용성, 프로젝트 결과로서 행동촉구메시지 요소 부각	
이번주 진행사항	감성추구형 0-100점: 소비자 니즈 부합. 콘텐츠의 영향력 관점에서 느낌/감정의 정서적 요소 부각하고 파급효과를 고려한 스토리텔링 포함. 긍정감정(공감/감동/호기심/행복 등)과 부정감정(죄의식/불안/슬픔/경각심 등)의 명확한 컨셉 반영	
	관심주목형 0-100점: 소비자 니즈 부합. 트렌드 반영형과 감성 추구형에 정확히 분류되지 않지만 관심과 주목을 끌기 위한 다양하고 독특하며 참신한 기법(재미, 놀람, 의문, 어그로 등)을 명확히 보여주는 요소 부각. 콘텐츠의 창의적 가치가 가장 돋보이는 방식으로 인지적 관심을 유도하는 컨셉 반영	
	종합평가 0-100점: 직관적으로 전체의 프로젝트의 결과물이 목적에 부합한지의 적절성과 효율성 파악. 좋은 PR VS 안 좋은 PR. 효과성- 실제 효과가 어떠했는지 정량적 수치 들어갈 수 있는 부분(0/부분 추후논의).	
	PR 데이터 활용 방안 및 과정	
	1) (주)슈퍼히어로의 PR 제작 과정이 담긴 로우 데이터를 제공 받는다.	
	2) LLM 모델을 활용하여 슈퍼히어로의 로우 데이터를 구조화된 형태로 정리하여 정제 데이터를 얻는다.	
	3) 정제 데이터를 미디어커뮤니케이션 학생에게 전달하여 서술한 6가지 평가 기준에 따라 평가하고 각 항목의 점수를 컬럼으로 저장한다.	
	4) 평가된 최종 데이터를 활용하여 벡터 데이터 베이스를 구축하고 RAG을 활용하여 퀄리티 높	

은 PR 데이터 생성에 활용한다.

샘플 데이터 파악

제공 받은 60여 개의 로우 데이터에 대한 확장자, 이미지 유무, 텍스트 형식 등을 파악하여 추후 데이터 추출 및 정제 담당자에게 전달

2. 개별 활동내용

박지민(팀장, UI담당)

- 타이핑 애니메이션 추가 및 디자인 개선
- 복사/평가 관련 버튼 제작(현재, 버그로 추후 수정 및 도입 필요)
- chatbox.js / input.js 코드 통합 및 최적화

김민수(팀원, 백엔드 담당)

- 백엔드 GPT 모델 로딩

임성민(팀원, 파인튜닝 담당)

- Chroma vector DB 및 textspliter 조사
- 정제 데이터 구조화 체계로 메타데이터를 추가하여 DB 저장 구현 중
- RAG 코드 개선

장희웅(팀원, 데이터 마이닝 담당)

- 샘플 로우 데이터 파악 및 텍스트 추출 방안 고안
- 데이터 정제 과정 추가 조사 및 LLM을 통한 데이터 구조화 방안 탐색

개발계획 대비 진행현황

기간	담당자	진행상황	항목명
5주차(3/31~4/6)	전체	완료	기업 미팅 및 요구사항 분석
5주차(3/31~4/6)	전체	완료	개발 방향 설정 및 프로그램 기능 설정
5주차(3/31) ~	장희웅	완료	ppt, hwp, docx 등 문서 파일 데이터 추출 및 정제 코드 개발
6주차(4/7~4/13)	전체	완료	프로그램 설계 및 기술 선정
6주차(4/7~4/13)	전체	완료	기업 데이터 인계 및 분류 기준 선정
6주차(4/7~4/13) ~	전체	완료	LLM 모델 선정
7주차(4/14~4/20) ~	박지민	완료	웹 UI 개발
7주차(4/14~4/20) ~	김민수	지연	LLM API 활용을 위한 테스트 서버 구현
7주차(4/14~4/20) ~	임성민	지연	LLM 모델 파인튜닝 테스트
8주차(4/21~4/27) ~	장희웅	진행 중	데이터 샘플을 이용한 데이터 정제 테스트
8주차(4/21~4/27) ~	임성민	진행 중	RAG와 기업 데이터를 활용한 답변 생성 구현
8주차(4/21~4/27) ~	박지민	완료	로컬스토리지 통한 사용자 데이터 보관 및 활용
8주차(4/21~4/27) ~	전체	완료	중간 발표 준비
9주차(4/28~5/4) ~	임성민	진행 중	정제 데이터 벡터화 구현
9주차(4/28~5/4) ~	장희웅	진행 중	데이터베이스 형식에 맞는 데이터 생성
9주차(4/28~5/4) ~	박지민	진행 중	답변 평가 UI 개발
9주차(4/28~5/4) ~	장희웅	진행 중	학습에 적합한 데이터 샘플 완성
10주차(5/5~5/11) ~	임성민		데이터 샘플을 이용한 LLM 파인 튜닝
10주차(5/5~5/11) ~	김민수		파인 튜닝 모델 API 개발
10주차(5/5~5/11) ~	박지민		언어유희 데이터 추출
11주차(5/12~5/18) ~	박지민		파인 튜닝 API와 UI 연결, 프론트엔드 개발
13주차(5/26~6/1)	전체		3차 발표를 위한 결과물(프로토타입) 완성

<p>다음주 계획</p>	<p>박지민(팀장, UI담당)</p> <ul style="list-style-type: none"> - 복사/평가 버튼 도입 및 평가 정보 전달 코드 구현 - 외부 언어 유희 관련 데이터 추출 및 정리 <p>김민수(팀원, 백엔드 담당)</p> <ul style="list-style-type: none"> - 백엔드 GPT 모델 로딩 <p>임성민(팀원, 파인튜닝 담당)</p> <ul style="list-style-type: none"> - 기존 데이터 DB의 정제 데이터 벡터화 및 벡터 DB 구현 - 기업 데이터의 메타데이터 구체화 후, RAG 효율 테스트 진행 - <p>장희웅(팀원, 데이터 마이닝 담당)</p> <ul style="list-style-type: none"> - 샘플 데이터 60여 개 구조화 및 DB 저장
<p>주요 결과물</p>	<p>자료조사내용 https://github.com/BAAKJIMIN0/PRoAI/tree/main/%EC%9E%90%EB%A3%8C%EC%A1%B0%EC%82%AC </p> <p>UI설계 https://github.com/BAAKJIMIN0/PRoAI/tree/main/UI </p> <p>LLM https://github.com/BAAKJIMIN0/PRoAI/tree/main/LLM </p> <p>데이터 마이닝 https://github.com/BAAKJIMIN0/PRoAI/tree/main/%EB%8D%B0%EC%9D%B4%ED%84%B0%EC%88%98%EC%A7%91%EC%BD%94%EB%93%9C </p>