

김현수

Data Engineer rover0811@hotmail.com [Linktree](#)

Tech Spec

Python, Airflow, Bigquery, ElasticSearch

경력사항

뤼튼테크놀로지스, Data Engineer Intern (2025년 8월 - 2026년 2월)

- Mail Agent 개발하여 뤼튼 내 서비스 출시
 - Langgraph를 활용한 Multi-Agent 워크플로우를 설계하여 복잡한 메일 본문의 맥락을 유지하며 실시간 요약 및 자동 분류 시스템 구현
 - Locust를 활용한 시나리오 기반 부하 테스트 수행하여 평균 30초의 고지연(High-Latency) LLM 추론 플로우를 **60 RPS** 환경에서 **P95 40초** 이내로 방어하며 시스템 안정성 입증.
- Apache Airflow ETL 파이프라인을 설계 및 개발하여 **AI** 기반 검색 기능의 데이터 처리 워크플로우를 최적화
 - 채용, 주식, 웹소설 ETL 파이프라인 운용 ElasticSearch, BigQuery에 적재
 - ElasticSearch 쿼리를 도메인 별 튜닝하여 **Recall@5** 평균 **0.7**에서 **0.9**로 개선
 - RRF, CC, Embedding 차원 및 벤더 등 다변량 수치 조절해가면서 개선
 - LLM tool(채용, 주식) 개발하여 전체 뤼튼 tool call 답변에서 **25%**를 차지하는 일평균 **57000**번의 **tool call** 호출 처리
- Crack 웹소설 프로젝트 참여
 - 나무위키 게시글 지식그래프로 웹소설 생성 파이프라인의 기반 데이터 ingestion
 - 카카오페이지, 노벨피아 등 다양한 소설 플랫폼의 데이터를 표준화된 dbt 스키마로 설계하여 SCD type 2 데이터 일관성 확보
 - 메달리온 단계별 문서간 유망한 간선을 LLM으로 정제 및 변환을 통해 웹소설 생성 파이프라인 구축
 - [Asyncio 기반 Producer, Consumer 패턴 공용 라이브러리](#) 개발하여 Rate limit을 걸면서 동시성 확보

대한민국 공군 지능정보체계관리단, RPA개발병 (2023년 5월 -2025년 2월)

- Apache Airflow ETL 파이프라인을 개발하여 6개 프로젝트에서 연간 **5,000**시간 이상의 업무 시간을 절약
 - 군 내 ERP 내부 데이터를 활용하여 반복되는 데이터 인입, 변환, 적재하는 부서의 업무를 자동화
- 군 내 개발환경 개선
 - Docker Compose Airflow 클러스터 도입 및 보안 인허가 업무, 서버 배포 및 운영
 - Private PyPI 서버, GitLab Runner 기반 CI/CD 파이프라인을 포함한 포괄적인 DevOps 인프라를 구축
 - 문서 자동화(HWPX) 라이브러리, 로깅 시스템, UiPath 연동을 위한 커스텀 Python 라이브러리를 개발하고 표준화된 패키지 개발 워크플로우를 구성

밸리언트데이터, Python Developer Intern (2022년 7월 - 2022년 12월)

- JSON 스키마 준수를 위한 Python 기반 데이터 검증 시스템의 유지보수 및 운용, 대용량 데이터셋 처리

프로젝트

[EATSSU](#), Python Developer (22년 7월 - 현재)

- AWS SAM과 Lambda, EventBridge 스케줄링을 활용한 자동화된 학식 메뉴 스크래핑 시스템을 구축하고, 대학 식당 서비스에서 구조화된 메뉴 데이터를 추출하는 텍스트 파싱 알고리즘을 구현
- 메뉴 데이터 처리를 위한 LLMops 파이프라인을 개발 및 운용하여 프롬프트 엔지니어링, LLM 응답 평가, 프로덕션 모델 모니터링을 통해 일관된 데이터 품질과 시스템 안정성(1년간 task 성공률 99.98%)을 보장

[WhisperLive](#), 오픈소스 기여 (25년 5월)

- 유연한 프로토콜 선택 기능을 포함한 보안 WebSocket(WSS) 지원 기능을 기여하여 다양한 배포 환경에서 프로덕션 수준의 보안 연결을 가능하게 함

인적사항

승실대학교, 소프트웨어학부 **21**학번

2021년 3월 - 현재 (졸업예정: 2027년 2월)

