

직무기술서 (Job Description)

항목	내용
문서번호	HR-JD-2026-017
직무명	AI Engineer
직군	개발
소속	DX부문 SW혁신센터 AI플랫폼팀
직급 범위	CL2 ~ CL3
작성	인사팀 직무관리파트
최종 개정	2026-03-12

1. 직무 목적

사내 업무 시스템과 제품에 적용되는 생성형 AI 기능을 설계·구현·운영한다. 사내 LLM 플랫폼 위에서 RAG·에이전트 기반 서비스를 만들고, 품질과 비용을 측정 가능한 수준으로 관리하는 것이 이 직무의 핵심 책임이다.

2. 주요 업무

구분	업무	비중
서비스 개발	LLM 기반 기능(문서 검색·요약, 상담 보조, 코드 지원 등) 설계 및 구현	40%
데이터·검색	사내 문서 전처리, 임베딩 파이프라인, 벡터 검색 품질 개선	20%
평가·운영	응답 품질 평가셋 구축, 자동 평가, 프롬프트·모델 버전 관리, 비용 모니터링	20%
협업	현업 부서 요구사항 정의, 보안·법무 검토 대응, 결과 리뷰	15%
기술 검토	신규 모델·프레임워크 검증 및 사내 적용 가이드 작성	5%

3. 필수 역량

- Python 기반 백엔드 개발 경험 3년 이상 (FastAPI 또는 동급 프레임워크)
- LLM API(OpenAI, Anthropic, 또는 사내 모델) 연동 서비스를 실제 운영 환경에 배포한 경험
- RAG 구조(문서 청킹, 임베딩, 벡터 DB, 리랭킹) 설계·구현 경험
- 컨테이너(Docker) 및 Kubernetes 환경에서 서비스 배포 경험
- 문제를 정의하고 평가 지표를 세워 개선한 경험을 문서로 설명할 수 있을 것

4. 우대 역량

- 에이전트 프레임워크(Tool use, 멀티스텝 오케스트레이션) 구현 경험
- 오픈소스 LLM 파인튜닝(LoRA 등) 또는 서빙(vLLM 등) 경험
- 대규모 사내 데이터의 접근 권한·개인정보 처리 설계 경험

- 반도체·가전·모바일 도메인 데이터 이해
- 영어로 기술 문서를 읽고 쓰는 데 지장이 없을 것

5. 필요 경력·학력

- 경력: 관련 분야 3년 이상 (CL2) / 7년 이상 (CL3)
- 학력: 학사 이상. 컴퓨터공학·전기전자·수학·통계 등 관련 전공 또는 동등한 실무 경험

6. 사용 기술 (팀 표준)

영역	기술
언어·프레임워크	Python 3.11, FastAPI, LangGraph
모델	사내 LLM 게이트웨이(Claude, GPT, 사내 파인튜닝 모델), Embedding API
데이터	PostgreSQL + pgvector, Elasticsearch, Airflow
인프라	사내 Kubernetes, GPU 클러스터(A100/H100), Terraform
관측	Langfuse, Grafana, 사내 비용 대시보드
협업	GitHub Enterprise, Jira, Confluence

7. 성과 지표 예시

- 담당 서비스 응답 품질 점수(사내 평가셋 기준) 및 사용자 만족도
- 토큰 비용 대비 처리 건수
- 서비스 장애 건수·복구 시간