

신동민 | AI Platform Backend Engineer

Python·Java 기반 백엔드, 데이터 모델링, LLM/RAG 파이프라인, AI 가드레일, MCP 통신 설계/개발, Azure-AKS 운영까지 AI/웹 서비스 전 주기 아키텍처를 설계·개발·운영했습니다. 대규모 빅데이터 웹 서비스에서 총 5억 규모 프로젝트를 3건 완수했으며 담당 서비스들의 성능 개선·운영 자동화·안정성 확보를 이뤄 KT DS CEO 표창, KT 품질혁신상을 수상했습니다.

기술 스택 : Java, SpringBoot, Python, FastAPI, JPA, PostgreSQL, VectorDB, Azure, Docker, Kubernetes, LLM, CursorAI

Careers

KT DS / Platform Backend Engineer @ 솔루션개발팀 (2024.09 ~ 현재)

AI EDU (2025.02 ~ 2025.07): 사내 1,200명 규모 사용 환경에서 월 평균 36만 건 이상 학습 요청 안정적 처리, 무장애 운영

- 백엔드 PM 및 AI 아키텍처로서 사내 대표 학습 플랫폼 런칭(사용자 1,200명) : 커서 AI 기반 6주 내 E2E 개발/주도
- FastAPI, SQLAlchemy, PostgreSQL, ChromaDB 기반 백엔드/DB 전체 설계 및 Docker-AKS 기반 CI/CD 자동화
- 멀티모달 LLM 기반 지능형 콘텐츠 생성 파이프라인 설계/개발 통한 비즈니스 지표 혁신 : 비정형 PDF의 '지능형 데이터 자산화' 파이프라인 설계 → 콘텐츠 생성 비용 50% 절감 및 무손실 데이터 처리, '설계자-저자' 2-Phase LLM 아키텍처 및 프롬프트 강화, 폴백 로직 도입 → 수동 검수·수정 시간 50% 단축, 품질 안정성 99% 확보
- 개인화 학습 기능 : 학습 이력·진행률 저장, 직무/과목 기반 콘텐츠 분류·로드맵 제공 → 학습 연속성/완료율 개선
- 프롬프트 관리 기능 제공 : 개인/조직용 프롬프트 관리 및 콘텐츠 생성 시 유연한 적용 기능 제공

BEAST AI Gateway (2024.09 ~ 2025.12): 멀티 테넌트 API Gateway 환경에서 수백 TPS 수준의 동시 요청을 처리하는 실시간 스트리밍 파이프라인 설계·운영

LLM·MCP 통신 AI 가드레일 개발 (2025.08~2025.12): 99.9% 안전성 보장된 초기 응답 속도 평균 4.2초로 개선

- 유해 요청 차단률 99.9% 이상 유지(운영 로그 기준) : ACS-LLM-패턴(정규식/키워드 기반 블랙/화이트리스트) 적용
- LLM 유해 응답 노출률 0.1% 이하로 관리: Azure PII 8개 카테고리/minimumPrecision 기반 민감정보 마스킹/차단, 슬라이딩 윈도우/Dynamic Buffer 기반 체크 경계 PII 누락 방지, ACS-패턴 기반 유해 응답 차단
- LLM 스트리밍 기반 실시간 병렬 가드레일-SSE 파이프라인 구현 : 토큰 스트림을 동적 버퍼링·재채크하여 가드레일을 병렬 적용하고, Backpressure로 과부하를 제어하며 SSE 응답으로 안정적으로 전송
- 전략 패턴 기반 엔진 모듈화 : Factory-Processor-Contract 구조 기반 SRP/OCP 적용 및 DB 내 정책 실시간 반영
- 다양한 LLM(OpenAI, Claude)-MCP 연동 환경에서 AI Gateway 내부 로직을 이해한 뒤, Provider별 다른 요청·응답 포맷을 포괄하는 가드레일 처리 계층을 적용하여 플랫폼 기능 확장

API 이상패턴알람 실시간 탐지 및 알람 체계 구축 (2024.09 ~ 2025.01): 장애 인지 시간 10분 → 1초로 단축

- 주기·에러코드·실패건수 기반 정책 설정 후 분·시·일 배치 이후 통계 내 에러 패턴 탐지, Spring Boot 기반 정책·통계 API 개발 및 JPA/QueryDSL 기반 조회 최적화, 파티셔닝·인덱싱으로 대용량 기간 조회 성능/안정성 개선

MCP 연동 기능 개발 (2025.07)

- 핵심 컴포넌트 개발 : Spring WebFlux SSE Transport 커스터마이징 → 외부 MCP 서버와의 실시간 스트림 안정화

KT DS / Web Developer & Manager @ 데이터DX개발팀 (2021.01 ~ 2024.08) : 총 5억 규모 프로젝트 3건 리드

KT 빅사이트 관광 고도화 프로젝트 (2022.10 ~ 2023.02)

- 8인 개발팀 스크럼 리더(PO 역할)로 협업 체계 구축 → 일정 준수 및 품질 확보
- 웹 로딩 속도 60% 단축하여 사용자 체감 성능 개선, CPU 효율 20% 향상 : 주요 차트·테이블 쿼리 튜닝, 공통 로직 병렬화, JS async/await 구조 개선으로 페이지 초기 로딩 대폭 단축
- 총 365개 기능 변경 중 234개 리드 개발(64%), 127개 협업 개발 → 팀 매출 3.74억 원 창출

KT 빅사이트 비즈니스 메타데이터 DB / Grafana 실시간 모니터링 시스템 구축 (2023.11 ~ 2024.05)

- PM 보조로서 미사용 테이블/SP 77% 삭제, DB 5.1TB 확보 → 연 1.05억 운영비 절감, KT 품질혁신상 수상

ETC

교육 : 중앙대학교 서울캠퍼스 경영학/컴퓨터공학 학사 | GPA : 3.81/4.5 | EPITECH Paris SW 장학 교환학생 과정 수료

입상/특허 : KT DS CEO 표창 | KT DS 사내 아이디어톤 대상 | AI + 인증 획득 | KT 품질혁신상

자격증 : 정보처리기사 / SQLD / AWS Certified Developer – Associate / AWS Certified Cloud Practitioner

포트폴리오 : [Technical Project Portfolio](#) | [Notion \(Projects & Architecture\)](#) | [GitHub](#)