

경력기술서

AI Engineer

Profile

이름	장 운
포트폴리오	https://woonjang.dev
전문분야	Python(FastAPI) 백엔드 · AI/LLM 서비스 · 데이터엔지니어링 · DevOps
직무경력	(주)케이비엘러먼트 2026.07.01 ~ 재직중 AI 엔지니어 (주)엔이에스 2025.03 ~ 2026.04 (1년 2개월) AI DC 클라우드 운영 PL 스마일봉구 2024.02 ~ 2025.03 (1년 2개월) LLM 기반 데이터 가공 서비스 1인 개발·운영 (주)짐브러스 2022.09 ~ 2024.01 (1년 5개월) 백엔드 / DevOps 엔지니어 (주)티웹 2021.02 ~ 2021.08 (7개월) 클라우드 인프라 엔지니어 (주)와이엔씨씨 2018.11 ~ 2019.08 (10개월) Linux / 클라우드 컨테이너 엔지니어

Skill

AI / LLM	SGLang, vLLM, 오픈웨이트 LLM 온프레미스 서버
Language	Python, Java, JavaScript/TypeScript
Backend	FastAPI, Django, Flask, REST API
Frontend	React, Vite, Next.js
Database	PostgreSQL, MariaDB/MySQL, InfluxDB(TSDB), Redis
System	Linux (Ubuntu/CentOS/Redhat), Windows, Mac OS
Infra / Tool	Docker, Kubernetes, Jenkins, GitLab CI/CD, GitHub Actions, Nginx, Airflow, AWS / NHN Cloud, Git, JIRA

Project

(주)케이비엘러먼트 — AI 엔지니어 2026.07.01 ~ 재직중
[2026.07 ~ 현재] 사내 챗봇 플랫폼 — 온프레미스 LLM 서버 + 풀스택 웹 챗봇 (단독 설계·개발·운영) - 외부 LLM API 사용이 차단된 환경에서 사내 GPU 서버(RTX PRO 6000 × 2) 기반 챗봇 플랫폼을 0→1 단독 구축 - SGLang + Qwen3.8-27B-FP8(TP=2, MTP) 모델 서버, OpenAI-compatible 게이트웨이로 failover/failback-모델 revision 고정 등 운영 설계 - FastAPI 백엔드 + React/TypeScript 프론트엔드, SSE 토큰 스트리밍 기반 실시간 채팅 구현 - 문서 첨부 파이프라인: 업로드 4중 검증 게이트(매직바이트·zip 폭탄·경로탈출 차단), PDF/DOCX/PPTX/XLSX/HWPX 다형식 파싱, OCR 폴백 체인(PaddleOCR-VL → RapidOCR → Tesseract) - 보안·거버넌스: 세션 쿠키+CSRF, TOTP MFA, PBKDF2, metadata-only 감사 로깅, 대화 30일 자동 파기 정책 설계 - 계측 내장: 응답 latency/TTFT/토큰 수집 스키마와  피드백 파이프라인으로 파일럿 품질 측정 기반 마련 - LLM 벤치마크 하네스 개발(부하 sweep, TTFT/ITL percentile 분석, GPU-less CI 회귀 검증)로 동시 사용자 25명 수용 용량 산정 근거 문서화 - ruff/pytest/vitest 기반 단위·통합 테스트와 GitHub Actions CI 자동화 스택: Python 3.12, FastAPI, PostgreSQL, React 19, SGLang, Qwen3.8-27B, PaddleOCR-VL, GitHub Actions
[2026.07 ~ 현재] 보안-우선 RAG 파이프라인 — 제조 문서 검색·응답 (아키텍처 설계 + 골격 구현) - DRM 보호 사내 문서를 승인된 복호화 절차로만 처리하는 보안 게이트웨이(fail-closed)·격리(quarantine)·감사 계약을 포함한 RAG 아키텍처 설계 - 문서 파싱 → page-aware 청킹 → 벡터 DB(Qdrant) 적재 → ACL 사전 필터 검색 → 리랭킹 → vLLM 생성까지 단계별 컴포넌트 계약 설계 및 골격 구현 - PostgreSQL 메타데이터 SoT, docker-compose 기반 local/DMZ/GPU 3종 환경 분리, Prometheus/Grafana 모니터링 스타터 구성 - 제조 데이터 표준(OPC UA, ISA-95) 친화 산업 RAG 아키텍처 설계 문서화 스택: Python, FastAPI, Qdrant, PostgreSQL, vLLM, Docker Compose, Prometheus/Grafana

Project

(주)엔이엑스 — AI DC 클라우드 운영 PL 2025.03 ~ 2026.04 (1년 2개월)

AI 데이터센터 클라우드 인프라 운영 총괄 (SI 신규 구축 + SM 안정 운영 혼재 환경)

- AI 데이터센터 클라우드 인프라 전반의 운영 프로세스 정립 및 운영팀 리드
 - Docker / Kubernetes 컨테이너 기반 서비스 배포 파이프라인 표준화 (Jenkins CI/CD)
 - Git 기반 IaC·구성관리 형상화, 변경관리·릴리즈 프로세스 정착
 - 운영 모니터링 지표 정의 및 SLA 안정성 지속 개선, 장애 대응 Runbook 정비
 - 프로젝트 일정·리소스·리스크 관리 및 이해관계자 커뮤니케이션 총괄
- 스택: Kubernetes, Docker, Jenkins, Git, Linux, Nginx, Python, Bash, 모니터링 스택

스마일봉구 — LLM 기반 데이터 가공 서비스 (1인 개발·운영) 2024.02 ~ 2025.03 (1년 2개월)

GPT/Claude API + LangChain 기반 비정형 → 상품 데이터 변환 SaaS (기획~운영 풀사이클)

- FastAPI 기반 REST API 백엔드 설계·구현 (인증, 비동기 처리, 작업 큐 연동)
 - GPT API·LangChain 기반 비정형 텍스트 → 정형 상품 데이터 변환 Converter 직접 개발
 - Airflow 스타일의 스케줄링 작업으로 대량 데이터 배치 처리 및 결과 적재 자동화
 - PostgreSQL 스키마 설계·인덱스 튜닝, 결제·주문 데이터 정합성 관리
 - Docker 컨테이너 기반 배포, GitHub Actions / Jenkins 스타일 CI 워크플로우 구축
 - Claude Code · GPT(Codex) 등 AI 코딩 도구를 일상 개발 워크플로우에 적극 활용
- 스택: Python, FastAPI, PostgreSQL, Redis, Docker, GPT/Claude API, LangChain, Git

(주)집브러스 — 전기차 충전·배터리 솔루션 백엔드 / DevOps 2022.09 ~ 2024.01 (1년 5개월)

전기차 운행·충전·배터리 데이터 수집·저장·서빙 백엔드 및 DevOps 파이프라인 구축

- Python(Django · Flask) 기반 REST API 설계·개발 및 운영
 - InfluxDB(시계열 DB) 기반 전기차 운행·충전 데이터 수집·저장·API 서비스 구축
 - MariaDB / PostgreSQL 기반 마스터·트랜잭션 데이터 모델링 및 쿼리 튜닝
 - Airflow 스케줄링 기반 데이터 수집·집계·리포트 ETL 파이프라인 운영
 - Docker 컨테이너화 및 GitLab CI/CD 빌드·테스트·배포 자동화 파이프라인 구현
 - AWS · NHN Cloud 인프라 구축·운영, Nginx 리버스프록시·로드밸런싱 구성
 - Grafana 기반 실시간 모니터링 대시보드 구축, React/WebGL 기반 인터랙티브 UI 구현
- 스택: Python, Django, Flask, FastAPI, PostgreSQL, MariaDB, InfluxDB, Docker, GitLab CI/CD, AWS, NHN Cloud, Grafana, Airflow, Git

(주)티웹 — 클라우드 인프라 엔지니어 2021.02 ~ 2021.08 (7개월)

- Linux / CentOS 서버 구축 및 운영 관리, 데이터 분석 시스템 운영·유지보수
 - Docker / Kubernetes 컨테이너 환경 구축·관리·운영
 - PaaS / IaaS 클라우드 기반 서비스 운영
- 스택: Linux/CentOS, Docker, Kubernetes, PaaS/IaaS, Git

(주)와이엔씨 — Linux / 클라우드 컨테이너 엔지니어 2018.11 ~ 2019.08 (10개월)

- Linux / Redhat 서버 운영 및 유지보수, 관리자 시스템 운영
 - 클라우드 컨테이너 서비스 운영
- 스택: Linux/Redhat, Container, Cloud

본 경력기술서의 모든 내용은 사실임을 확인합니다.