

AI Engineer · LLM Systems

On-Prem LLM Infra · Access-Controlled Memory · Security-grounded

Gyeonggi-do, South Korea | bk@bkan.dev | (+82) 10-4676-5884 | GitHub | Portfolio | Threads

경력 요약

기간	소속 / 형태	역할 · 요지
2026	기업 AI 시스템 구축 (프리랜서)	AI 엔지니어 · 개발팀 기술 리드 - 온프레임 LLM 인프라 · 전사 AI 메모리 · DB 설계 · ADR 거버넌스
2025 - 현재	솔로프리너	창업자 / 엔지니어 - AI 프로젝트 기획·개발·출시 (OurFolio · ASHD v2 · BluePy · BK-HQ)
2025.06 - 2026.01	FastCampus AI 부트캠프 14기	수료 - 실전 ML 대회 5회 · MLOps · 멀티에이전트 프로젝트
2021 - 2025	프랜차이즈 운영	대표 - POS 데이터 수요예측·재고최적화 → AI 커리어 전환 준비
2018.02 - 2021.02	CMT EYES	보안 관제 분석가 - SOC / CERT 침해대응

기업 AI 시스템 구축 - 프리랜서

2026

외부 클라우드로 데이터를 내보낼 수 없는 기업 환경에서, LLM 인프라·데이터 모델·전사 지식관리·개발 프로세스까지 AI 시스템 전반을 설계·구축. 소규모 개발팀의 기술 거버넌스 리드를 겸임.

온프레임 LLM 멀티에이전트 인프라

역할 설계·구축·운영 전담 · 기술 Self-Hosted LLM · Tailscale · GGUF/MLX · Multi-Agent · Discord Bot

- 배경 보안 요건상 외부 클라우드 LLM API를 사용할 수 없는 환경 - 사내 자체 서버 인프라 필요
- 수행 Mac Studio 15대에 역할을 나눠 배정하고 Tailscale 독립망으로 묶어 외부 노출 없이 LLM 자체 서버, 디스코드 봇을 단일 진입점으로 구성
- 수행 장비별 로컬 LLM 포트폴리오 설계 - 라이선스·RAM·양자화(GGUF/MLX)·KV cache 헤드룸을 정량 근거로 GLM·Qwen 계열 모델 배정, 제조사 공식 전력 데이터 크로스체크로 장비 시트 오류 발견·정정
- 수행 진행 에이전트·검수 에이전트가 작업을 자동으로 주고받는 2-에이전트 핸드오프 파이프라인 구축
- 성과 재부팅 시 자동 복구되는 상시 가동 클러스터 - 전사 메모리·에이전트 시스템이 이 스택 위에서 구동

통화상담·고객 데이터 통합 DB 스키마 설계

역할 데이터 모델링 전담 · 기술 PostgreSQL · SQL-first Migration (goose) · ETL

- 수행 레거시 시스템 ETL 분석 - 고객 9.4만·주소 8.6만·상담 5.7천 건 이관 설계, CDR 원시 3.2만 행의 IVR·수신 중복 구조를 직접 발견해 실통화 2.3만 건으로 정규화
- 수행 통화상담·고객 데이터를 통합하는 27-table 정규화 스키마 설계 - 자연키 전략, 소프트삭제, 변경이력 추적(SCD), 불변식(invariant) 명세 포함
- 수행 마이그레이션 엔진 비교 검증(Atlas·dbmate·sqitch·Flyway 검증 후 기각 → goose 채택), pgvector/halfvec 용량 한계 외부 검증, SQL-first SSOT + 스키마 불변식 체크리스트(문서 차원 TDD) 수립
- 성과 외부 검토 2회 및 적대적 코드검수 통과

개발팀 기술 거버넌스 · 리드

역할 개발팀 메인 관리자 · 의사결정 기록 담당 · 기술 GitHub Monorepo · ADR/MADR · Git Workflow

- 수행 3인 개발팀의 GitHub 조직·모노레포·브랜치 전략(main/dev/feat) 셋업, ADR(기술)+Decision(비기술) 의사결정 기록 체계 도입·운영 (ADR-001~008)
- 수행 3트랙(판매/광고/CRM) 업무 분담 설계와 팀 온보딩 주도, 멀티세션 AI 협업을 위한 워크트리 격리·Safe Staging 롤 제정
- 성과 계약 종료 시 인수인계 PR 202파일 / +37,000줄 + 비개발자 팀장용 요약본 별도 작성 - 인프라 변경으로 인한 데이터 유실 상황도 중립 기록·재발방지 규칙 제안으로 마무리

팀·전사 AI 메모리 시스템

역할 오픈소스 커스터마이징·시스템 설계 · 기술 Hindsight · Graphify · Access Control

- 배경 에이전트 메모리에 개인 업무기록이 쌓이는 구조 - 개인정보 격리와 누락 없는 축적을 동시에 달성해야 함
- 수행 오픈소스 에이전트 메모리 Hindsight를 커스터마이징해 직원별 personal·team 메모리 bank 구성, Recall·Reflect 권한을 보안 등급으로 제어 - 개인 bank는 본인만 열람
- 수행 1주 트라이얼 211건 운영 실측 - recall 기반 조회의 완전성 한계(재현율 15.6%)를 수치로 확인하고 시간 필터 직조회 원칙 확정, 요약 레이어(reflect)는 정보 오염 방지 근거로 OFF 결정
- 수행 일일·주간 업무보고 자동 합성 파이프라인 구축 - 품질 eval 실측(환각률 0), 임원 가치 큐레이션 롤 운영
- 성과 권한 격리를 유지하면서 누락 없이 쌓이는 전사 지식관리 뼈대 완성 - 개인 → 팀 → 전사 3단계 승격 구조로 설계

기업 AI 시스템 구축 - 프리랜서 (계속)

2026

한국어 임베딩·리랭커 3-way A/B 실험

역할 실험 설계·수행 전담 · 기술 MTEB-ko · FAISS · sentence-transformers · 통계 검증

- 배경 한국어 임베딩 리더보드 점수차(1.38pt)가 노이즈 수준 - 벤치마크만으로는 RAG 스택 모델 교체를 판단할 수 없는 상태
- 수행 arctic-ko vs KURE-v1 vs BGE-m3-ko **3-way A/B 실험** 설계 - 사전 동결 decision rule(Bonferroni $\alpha=0.025$, MRR), 90개 쿼리 6전략, self-retrieval 무결성 검증 포함
- 성과 전 pair 통계 유의성 미달 + 표본 부족 → **underpowered** 정직 판정, "교체 근거 없음"으로 기존 모델 유지 - 리더보드·커뮤니티 시그널·자체 실험 **3중 크로스체크**로 모델 선정 의사결정

코드베이스 지식그래프

역할 구축·검증 전담 · 기술 Graphify · Knowledge Graph

- 수행 코드베이스를 지식그래프로 구조화 - **1,158 노드 / 1,797 관계 / 73 커뮤니티**, 팀 자동화 도구와 결합
- 수행 파이프라인 전수 재검토로 그래프 무결성 버그 **4건 발견·수정**(hyperedge 병합 오류, phantom cross-project 엣지, 노이즈 노드 프루닝, stale 상태) 후 수치 확정

통합 어드민 대시보드 (데모)

역할 단독 개발 · 기술 Spec-Driven Dev (PRD/SDD/TDD) · Accessibility

- 수행 단일 모듈 레지스트리에서 **13개 모듈**을 파생하는 구조로 어드민 구축 - 접근성·다크모드·반응형 포함
- 수행 전수 감사로 UX·접근성 갭 **12건 발굴, 10건 구현** - 통합 검색·WAI-ARIA 키보드 내비게이션·다크모드 차트 동기화·폰트 self-host
- 성과 스펙 주도 개발로 테스트 **54 → 81건 통과**(회귀 0), 13경로 콘솔 에러 0

한국어 문서 파이프라인 · DOCX 변경추적 라이브러리

역할 단독 개발 · 기술 Python · OOXML(ECMA-376) · pandoc · MinerU VLM · typst

- 배경 임원 보고 워크플로가 HWP/DOCX/PDF 중심 - 변경추적·코멘트가 있는 문서를 LLM 파이프라인에 태울 도구 부재
- 수행 DOCX tracked-changes/comments **순수 XML 처리 라이브러리** 개발(LibreOffice 의존 제거, v1~v4 사이클, pandoc round-trip을 TDD 기본 룰로 채택) - 테스트 48건
- 수행 MinerU VLM 한국어 PDF·typst 조합 검증 후 Claude Code **스킬로 패키징**
- 성과 스킴 A/B 파일럿 실험 - 동일 과업 처리 시간 **55% 단축**(67초 → 31초), 도구 호출 수 절반

기타 실무

- 수행 회의 녹음 → 화자분리 → 회의록 자동화 폴스택 리서치(4-cycle 심층 조사), 금융 데이터 검증 MCP 서버(6 tools) 개발·납품 - "**미검증은 미검증으로 명시**" 납품 규율 운영, 사내 웹 유지보수

솔로프리너 - AI 프로젝트

2025 - 현재

기획 → 개발 → 배포 → 운영 판단까지 전 과정을 혼자 수행. 2+ 프로젝트 MVP 출시.

OurFolio - 근거 기반 포트폴리오 플랫폼 [GitHub ↗](#)[ourfolio.bkan.dev/demo](#) · 공개 쇼케이스

역할 1인 개발 (기획·개발·배포) · 기술 Next.js 16 · Neon · Drizzle ORM

- 수행 개발자가 작업물을 근거와 함께 증명하는 포트폴리오 플랫폼 - 공개 프로필 서버 렌더링, GitHub 레포 원클릭 импорт, OG/SEO 동적 생성
- 수행 **Expand-contract 2단계 무중단 스키마 마이그레이션**, Sentry PII 스크리빙(테스트로 강제), 동의 이력 policyVersion 버저닝 - Lighthouse-coverage·마이그레이션 검사 등 **CI 7종** 운영
- 성과 MVP → 검증된 **v1**까지 단독 개발·배포. 시장성 판단 후 수익화를 접고 읽기 전용 라이브 데모가 있는 **공개 쇼케이스로 전환** - 만들고 출시하고 냉정하게 접는 판단까지 수행

ASHD v2 - 영수증·보증서 OCR+LLM AS 관리 [GitHub ↗](#)[ashd.bkan.dev](#) · MVP 출시 → 공개 스냅샷

역할 1인 개발 · 기술 Next.js 15 · Supabase (Auth/RLS) · OCR · LLM · CI

- 수행 OCR로 영수증·보증서를 읽어 롤 추출기+LLM으로 구조화, 이메일·텔레그램 만료 알림 발송 - 결제 연동 실험까지 진행한 실사용 운영, eval·automerge·P0 알림 등 **CI 워크플로 15종**으로 1인 운영 자동화
- 성과 추출 정확도를 수치로 관리 - **KORIE 벤치마크(314예제)** 필드별 F1 score **9% → 100%**·구매일 89% → 99%·금액 98% → 100% 개선, 시드 골드셋 기준점 미달 PR을 CI가 차단하는 **eval 회귀 게이트** 운영

BluePy 2.0 - 오픈소스 인프라 보안 스캐너 [GitHub ↗](#)

v0.1 alpha · MIT

역할 단독 리빌드 · 기술 Python 3.12 · Clean Architecture · CLI

- 배경 2017년 KISA 레거시 보안 스캐너 - 거짓 PASS를 내는 validator가 섞여 있어 결과를 신뢰할 수 없는 상태
- 수행 Python 3.12 클린아키텍처로 재설계, 인과까지 검증한 점검만 기본 노출(Linux 파일권한, macOS SIP·FileVault·자동업데이트) - 거짓 PASS validator 삭제, 미검증 레거시 규칙은 experimental로 분리
- 수행 asyncssh 기반 **SSH 원격 스캔** 아키텍처(Linux/macOS, PASS·FAIL·MANUAL 3분류) - Docker 실행 SSH e2e 픽처 + 테스트 **546건**
- 성과 CLI 우선(JSON·표 출력)·SSH 호스트키 엄격 검증을 갖춘 **정직한 v0.1 오픈소스**로 공개 - 보안 실무 경험을 코드 신뢰성 기준으로 연결

솔로프리너 - AI 프로젝트 (계속)

2025 - 현재

BK-HQ - 1인 AI 회사 운영체제

설계·운영 중

역할 설계·운영 · 기술 Multi-Agent Orchestration · Automation Hooks

- 수행 CEO 1명이 AI 에이전트 **15명**을 4개 팀(기획·법무·마케팅·개발)으로 운영하는 오케스트레이션 플랫폼 - 스킬 31개 · 자동화 혹 19개 · 크루 워크플로우 5개
- 성과 Smart Skill Chain 5종·SNS 자동화·Agent Dashboard·지식 복리 시스템(25+ 도메인·100+ 문서)을 단일 운영체제로 통합 - **1인 운영 체제** 실증

보안 관제 & CERT - CMT EYES

2018.02 - 2021.02

보안 관제 & 침해 대응

역할 보안 관제 분석가 (SOC/CERT) · 기술 FortiGate/UTM · FortiSIEM · WAF/IPS · Wireshark

- 담당 Fortinet 중심 WAF/IPS/UTM/SIEM **20+ 보안장비**, 고객사 **50+곳** 원격 관제·분석 - 야간·주말 교대 관제와 에스컬레이션 체계 운영
- 수행 월 **1,000건+** 이벤트 대응 - 탐지 → 로그/패킷 분석 → 차단 → 정책 개선 → 보고까지 CERT 프로세스 전 과정 수행
- 수행 SIEM 상관분석 룰·IPS/WAF 커스텀 시그니처 튜닝, 이벤트 탐지용 신규 스크립트 다수 개발
- 수행 침해사고 분석 - 로그·패킷 기반 타임라인 재구성, 침해 경로 규명, 사고보고서 작성
- 수행 DDoS 긴급 대응(대역 차단·트래픽 분석), KISA·상급기관 위협정보의 고객사 영향도 파악·긴급 권고
- 수행 다수 대형 고객사 대상 **모의해킹·침해분석 프로젝트** 참여, 피싱 모의훈련 수행
- 성과 정책·탐지 룰 개선으로 오탐·정탐 분류 정확도 **50%+ 개선**
- 성과 반복 리포트 작성을 스크립트로 자동화 - 건당 **30분 → 5분** 단축

보안 엔지니어링 & 고객 지원

- 수행 신규 고객사 관제 온보딩 - 장비 초기 정책 셋업·베이스라인 튜닝, 방화벽 정책 정리·이관
- 수행 고객사 보안장비 엔지니어링 및 Linux/Unix 서버 운용
- 수행 대형 교육 고객사의 모의고사 기간 트래픽 폭주(연례 반복 장애)에 **10G 망 증설 현장 지원** - 네트워크 구축 실작업 참여
- 수행 코로나 시기 SSL-VPN 기반 원격근무 보안 전환 지원
- 수행 월간·분기 보안 보고서 작성과 고객사 정기 브리핑
- 수행 신입 교육 책임 및 육성 커리큘럼 담당

ML 대회 & 프로젝트 - FastCampus AI 부트캠프 14기

2025.06 - 2026.01

실전 ML을 대회 중심으로 수행. 아래 수치는 대회 기간 내 베이스라인 대비 개선폭.

토스 광고 CTR 예측	대규모 광고 로그 기반 CTR 모델 - 시퀀스 피쳐·다운샘플링·튜닝으로 LogLoss/AP 개선, 상위 10%
과학 QA IR (RAG)	Solar 임베딩·랭크 그래프 리팩토링·프롬프트 튜닝으로 검색·랭킹 개선 - MAP/MRR +107%
문서 이미지 분류	17클래스 스캔 문서 분류 - 백본 실험·TTA·증강·후처리로 F1 +470%
대화 요약	한국어 일상 대화 요약 - QLoRA 파인튜닝 + 증강·후처리로 ROUGE +30%
부동산 가격 예측	RMSE -58% 개선

마케팅 AI 멀티에이전트

- 수행 에이전트 **6개**가 트렌드·카피·세그먼트·리뷰·경쟁사·전략 리포트를 자동 생성 - 반복 리서치 업무 대체

영화 평점 예측 MLOps

- 수행 TMDB API 기반 평점 예측 서비스 - 데이터 수집 → 학습 → 배포 → 모니터링 **E2E 자동화** 파이프라인(CI/CD) 구축

LLM 기반 RAG 교육 콘텐츠

- 수행 사내 문서 검색에 RAG를 도입하려는 실무자용 교육 자료 제작 - BM25·Vector·Hybrid 비교, Query Expansion, Hit@k·MRR 평가 실습

참고

프로젝트별 증적(스크린샷·코드·라이브 데모)은 bkan.dev 에서 확인할 수 있습니다. 이력서(1p)는 bkan.dev/resume.html 참고.