

PE.

Product Engineer



개발을 도구 삼아 확실한 가치를 만들어
내는 든든한 메이커가 되겠습니다.

Name: 전소명

Email: callendev20@gmail.com

GitHub: github.com/Callein

LinkedIn: linkedin.com/in/callein

Education

한동대학교 (2020.03~2026.02 졸업예정)

AI·컴퓨터공학 전공 (3.81 / 4.5)

LeTourneau University (2025.08 – 2025.12)

교환학생 (미국, 텍사스)

Introduction

개발자의 본질은 문제를 해결하는 사람이라고 믿는 개발자 전소명입니다. 제게 개발은 사용자의 니즈를 충족하고 불편함을 해소하기 위한 도구이며, 기술을 통해 누군가에게 실질적인 가치를 제공할 때 진정한 즐거움을 느낍니다.

저는 안정적인 시스템 구축을 바탕으로, 8,800여 건의 데이터 중 중복 정보를 70% 압축해 높은 밀도의 정보를 제공하거나, 추천 엔진 연산 비용을 75% 절감해 쾌적한 응답을 제공하는 등 '문제를 해결하는 기술'에 집중해 왔습니다.

개발할 때는 "이게 정말 최선일까?"를 자주 물곤 합니다. 동료와 끊임없이 소통하며 미처 생각하지 못한 엣지 케이스를 검토하고, 더 나은 구조를 탐구하겠습니다. 더 나아가 팀과 비즈니스의 목표를 현실로 만드는 든든한 메이커가 되고 싶습니다.

SKILLS

BACKEND

- Java(Spring Boot)
- Python (FastAPI, Django)
- MySQL
- MariaDB

Infra & DevOps

- Docker
- S3
- RabbitMQ
- GitHub Actions

AI & Data

- FAISS
- Annoy
- Hugging Face
- fastText

Tools & Frontend

- Puppeteer
- FFmpeg
- Selenium
- React

Professional Experience

2024.12~2025.02 구도투자자문 Software Engineer Intern, QA Automation

투자 데이터 정합성 교차 검증 자동화

- 포트폴리오 UI 표출 수치와 실제 재무제표(Statement) 원본 간의 오차를 감지하는 자동화 검증 도구 개발
- 데이터 불일치 항목을 JSONL 로그로 자동 적재하여 투자 정보의 신뢰성 리스크를 사전에 차단

검증 프로세스 최적화 및 업무 효율 개선

- [Problem] 기존 Selenium 기반 테스트의 무거운 실행 속도와 반복 로그인 문제로 점검 효율 저하
- [Strategy] Puppeteer로 기술 스택 전환 및 세션 재사용(User Data Dir)로직 구현으로 초기화 시간 단축
- [Result] 9가지 핵심 점검 항목을 100% 자동화하여 일일 QA 소요 시간 절감

비개발 직군을 위한 사내 QA 도구화

- 개발 지식이 없는 동료도 손쉽게 정합성 검증을 수행할 수 있도록 CLI 메뉴 및 Batch Script(.bat) 실행 환경 구축

Puppeteer Selenium QA Automation Shell Script

Key Projects

RAGvertise Repo

2025. 01 ~ 2025. 08

RAGvertise (광고 포트폴리오 추천 엔진)

산학 연구원, 메인 개발자 (AI 추천 엔진 및 검색 API 개발, 성능 최적화 주도)

KTL 공식 성능 인증 획득 (ISO/IEC 17025)

- 한국산업기술시험원(KTL)을 통해 매칭 알고리즘의 응답 속도와 정확도에 대한 공인 성적서 획득

Weighted Late Fusion 기반 검색 엔진 고도화

- [Problem] 5개 필드 분리 인덱싱 도입으로 정확도는 상승했으나, 개별 쿼리 수행 및 결과 병합 과정에서 연산 비용이 5배 증가하여 레이턴시 지연 발생
- [Strategy] 코사인 유사도의 가중합이 벡터 내적의 선형성을 따른다는 점을 활용해, 가중치를 벡터 공간에 선반영하여 5개의 인덱스를 하나의 **Fused Index**로 통합하는 최적화 수행
- [Result] 검색 정확도(Recall@5) 0.87 달성, 연산 비용 75% 절감하여 레이턴시를 **0.25s**로 단축

Hybrid Embedding 전략 및 의도 분류

- [Problem] 단일 모델(E5) 사용 시, 긴 문맥은 잘 파악하나 짧은 키워드의 의미적 특성을 놓쳐 매칭 정밀도가 떨어지는 한계 직면
- [Strategy] 문맥 파악에 강한 **E5**와 키워드 분석에 특화된 **fastText**를 결합한 이원화 임베딩 파이프라인 구축
- [Result] 짧은 키워드(Style, How)의 의미 소실을 방지하여 입력 데이터 분류 정확도(**ROUGE-1**) **0.98**을 달성

RAG 기반 작업지시서 생성

- [Goal] 검색된 포트폴리오를 기반으로 실무에서 즉시 사용 가능한 광고 작업지시서를 자동 생성해야 함
- [Strategy] 검색 결과의 메타데이터를 시스템 프롬프트 컨텍스트로 주입하여 **RAG 아키텍처** 구현
- [Result] 생성 문장의 품질 지표인 **ROUGE-1 F1 점수 0.63**을 기록

FastAPI MariaDB Python FAISS Vector Search React

Core Repo

Validator Repo

Spotlight Repo

2024. 05 ~ 2025. 06

Handong Feed Platform (교내 정보 큐레이션 플랫폼)

백엔드 개발자 (아키텍처 설계, Core/AI 서비스 개발)

Spring Boot & Python 하이브리드 서비스 아키텍처

- [Problem] 트래픽 처리와 AI 연산의 공존으로 인한 리소스 경합 및 외부 호출에 대한 접근 통제 필요성
- [Strategy] Core (Spring Boot)와 AI Worker(FastAPI, Python)로 서비스를 분리하고, **API Key** 기반 RBAC 게이트 웨이를 도입하여 역할 기반의 요청 격리 및 제어 환경 구축
- [Result] 도메인 특성(Web/AI)에 맞춰 기술 스택 이원화, 리소스 격리를 통해 시스템 안정성을 확보한 서비스 구조 구축

TF-IDF/Annoy 기반 중복 탐지 및 데이터 압축 (Validator)

- [Problem] 다량의 수집 데이터 중 무의미한 중복을 제거하고 유사 정보를 그룹화하여 정보 밀도 향상 필요
- [Strategy] 최근 14일치 데이터를 메모리에 로드하는 월스탁트 인덱싱과 벡터 거리 기반의 분류 파이프라인 구현
- [Result] 건당 **0.9s** 이내의 처리 속도로 8,800건 이상의 전체 데이터 **70%**를 자동 필터링하여 정보 탐색 효율 개선

결합 허용 AI 태깅 및 프롬프트 최적화 (Spotlight)

- [Problem] 정보 탐색 효율을 위해 데이터 태깅이 필수적이었으나, 외부 LLM의 빈번한 호출 제한(Rate Limit)과 비정형 응답으로 파이프라인 안정성 확보 필요
- [Strategy] 프롬프트 최적화(JSON 강제 및 페르소나 적용) 및 태깅 실패 시 격리된 Fail Log DB에 적재하여 다음 주기에 자동 재시도하는 로직 구현
- [Result] 태그 분류 정확도 **87%** 달성 및 외부 API 장애 상황에서도 데이터 유실 없는 복원력 확보

On-Demand Serverless 워크플로우

- [Problem] 간헐적인 태깅 작업을 위해 워커 서버를 24시간 상시 가동하면 인프라 유휴 자원 낭비 발생
- [Strategy] 상시 서버 대신 데이터 검증 완료 시점에만 실행되는 **GitHub Actions Runner** 트리거 패턴 도입
- [Result] 불필요한 대기 서버 비용을 제거하여 운영 및 인프라 비용 최소화

Spring Boot JPA/MyBatis MariaDB FastAPI Vector Search Annoy Docker GitHub Actions

— PE.

Product Engineer

Core Repo

Worker Repo

2025. 03 ~ 2025. 07

CMS & Transcode Worker (비동기 분산 미디어 엔진)

백엔드 리드, 비동기 분산 처리 시스템 설계 및 미디어 파이프라인 최적화

RabbitMQ 기반 비동기 분산 아키텍처 (EDA)

- [Problem] 미디어 변환 작업이 API서버와 결합되면, 대용량 업로드 시 전체 시스템 응답 지연 발생
- [Strategy] RabbitMQ를 도입하여 요청(Producer)과 작업(Consumer)을 분리하고, 트래픽 폭주 시 메시지 버퍼링을 통해 서버 다운 방지 및 안정적 처리량 유지
- [Result] 동기 처리 대비 API 응답 속도를 대폭 개선하고, 워커의 수평 확장 구조 및 자동 재연결 로직으로 무중단 서비스 가능성 확보

독립적 파일 생명주기 관리 및 데이터 정합성 보장

- [Problem] 비즈니스 데이터 삭제 시 S3 객체가 남거나, 업로드 중단 시 스토리지에 더미 데이터가 쌓이는 고아 파일 (Orphan File) 문제 발생
- [Strategy] MinIO(S3) 객체의 상태를 추적하는 독립된 파일 레지스트리(TbFileList)를 구축하고, 삭제 트랜잭션 내 비관적 락을 적용하여 DB와 스토리지 간 동기화 강제
- [Result] 파일의 물리적 존재 여부를 논리적으로 제어하여 스토리지 비용 누수를 원천 차단하고 데이터 신뢰성 확보

FFmpeg 파이프라인 최적화 (Single-Pass)

- [Problem] 기존 멀티 프로세스 방식(동일 원본 파일 중복 Read)의 과도한 디스크 I/O 및 처리 시간 비효율
- [Strategy] Filter Complex를 활용한 Single-Pass 인코딩 파이프라인을 구축하여, 메모리 상에서 스트림 복제 후 480p/1080p 동시 변환 수행
- [Result] Disk Read I/O 50% 절감 및 CPU 컨텍스트 스위칭 비용을 최소화하여 인코딩 성능 최적화

Spring Boot, RabbitMQ, Python, Celery, FFmpeg, Kubernetes, MariaDB, MinIO

Awards & Activities

2025. 09

한동대학교 스마트 어플리케이션 공모전 우수상 CMS & Transcode Worker

교내 동아리의 효율적인 운영과 활동 기록 지식의 영속적 보존을 위한 Web App 및 미디어 엔진 개발

2025. 08 ~ 12

제3기 한미 첨단분야 청년교류 지원사업 장학생 (산업통상자원부 / 한국산업기술진흥원)

미국 LeTourneau University (Texas) 파견 (Software Engineering, Network Security 수학)
Project: LLM 기반 로컬 뉴스 콘텐츠 자동 선별 및 요약 시스템 설계 및 개발 (LocuSum)

2025. 08

KTL 소프트웨어 품질 성능 인증 (ISO/IEC 17025, 한국산업기술시험원)

RAGVertise (AI 추천 엔진) 공식 성능 시험 성적서 획득
매칭 알고리즘 Latency 0.25s, Recall@5 0.87 달성을 인증

2025. 01

COSS 전북대 올림피아드 (LLM 정확도 향상) 최우수상 (빅데이터 혁신융합대학 7개 대학 연합)

초기 정확도 30%인 sLLM을 RAG/FAISS/Sentence Transformers로 96.1%까지 향상
참가 20개 팀 중 1위 (Top Award) 수상

2024. 11

한동대학교 스마트 어플리케이션 공모전 대상 Handong Feed Platform

교내 정보 통합을 위한 Polyglot 아키텍처 기반 큐레이션 플랫폼 개발

2021. 12

한동대학교 스마트 어플리케이션 공모전 대상 HUT

불규칙한 대중교통 문제를 해결하기 위한 교내 택시 등승 매칭 O2O 플랫폼 (Flutter, Firebase)