정성헌

010-6506-0260 / masondev1024@gmail.com

포트폴리오 / GitHub

BI 엔지니어로 2년 5개월 일하며 SQL과 로그로 리포트 수치와 적재 오류를 확인했습니다. 이후 팀·개인 프로젝트에서 법령 검색과 데이터 조회 기능을 만들고, 오류가 났을 때의 재처리와 백업 복구를 확인했습니다.

경력

㈜비아이비즈앤 · BI 엔지니어

2022.04 - 2024.08 / 2년 5개월

MicroStrategy · SQL · Oracle · Linux

- 한국투자증권·신세계백화점·이마트24 등 고객사의 MicroStrategy BI 플랫폼 운영과 리포트·데이터 모델 개발 및 유지보수를 담당했습니다.
- 리포트 수치가 맞지 않으면 SQL 실행 결과와 원천 데이터를 대조했습니다. 시스템 로그와 Trace 기록도 함께 살펴 오류가 난 구간과 서비스 상태를 확인했습니다.
- 신세계백화점 Enterprise Manager 통계 적재 오류 대응에 참여했습니다. 수동 로드에서도 USER·PROJECT·SERVER DEFINITION 오류가 반복되는 현상을 재현하고, 접속·연동 설정과 재시도 결과·오류 로그를 정리해 기술지원의 후속 원인 분석을 지원했습니다.

프로젝트

청년 법률상담 AI 서비스

질문을 분류하고 관련 법령을 검색하는 기능을 만들었습니다. 팀 프로젝트 이후에는 상담 준비 카드를 만들고, 기록을 보관하거나 철회할 수 있도록 기능을 추가했습니다.

2026.06 팀 개발 / 2026.10 개인 후속 개발 / 금융 Agent·검색·Supervisor 분류 담당

Python · LangGraph · Bedrock Claude · Chroma/pgvector · RAGAS · AWS · Terraform

- 개인 작업에서 공식 안내로 상담 준비 카드를 만들고 인증과 관찰 기록 저장을 연결했습니다. 철회하거나 보관 기한이 지난 기록은 삭제하도록 했습니다.
- EC2·SQLite의 합성 자료를 S3 백업에서 새 서버로 복구하는 시험을 1회 진행했습니다. 약 3분 55초가 걸렸고, 보존 6건의 해시 일치와 DynamoDB 삭제 이력에 따른 철회 자료 제외를 확인했습니다.
- 합성 자료를 쓴 별도 EC2 내부 HTTP 시험에서는 연결 풀을 직접 갱신하고 복구 처리(reconcile)를 실행한 뒤 상태 API의 200→503→200 응답을 확인했습니다.
- 팀에서 Supervisor와 금융 Agent를 맡아 법령 XML을 조문·항·호로 나누고 BM25·임베딩을 결합한 검색을 구현했습니다. 개인 후속 검증에서는 개발용 120문항으로 hit@3 0.975, MRR@3 0.840을 확인했습니다.
- 별도로 고정한 12문항에서 같은 검색 결과·모델·채점 조건으로 답변 지침을 비교했습니다. RAGAS 근거 충실도 0.783→0.903, 질문 관련성 0.778→0.791을 기록했지만, 필요한 조항이 빠진 경우가 있어 기존 지침을 유지했습니다.

구현 코드 / 평가 원문 JSON / 복구 기록 / HTTP 권한 시험

Agentic Marketing Data Pipeline

판매자 데이터를 연결하고, 질문에 맞는 도구로 데이터를 조회하는 서비스를 만들었습니다.

2026.10 / 개인 프로젝트 / 조사 Agent·API·데이터 검증

Python · PostgreSQL · FastAPI · LangGraph · NVIDIA Nemotron Super · RDF/OWL · SHACL

- Nemotron Super가 질문을 분류하고 허용된 읽기 전용 도구로 조회하도록 했습니다. 호출 전 예산을 예약하고 완료 결과는 저장본을 다시 사용했습니다. 결과를 알 수 없는 모델 호출은 자동으로 다시 보내지 않도록 했습니다.
- 공개 Olist 리드 8,000건·계약 842건의 스키마·키·참조관계를 검사했습니다. 주문 표본과 연결되지 않은 계약 462건을 판매 0건과 구분했습니다.
- 10월 12일 AWS 시험 기록에서 같은 RDS의 별도 DB로 복원한 14개 테이블·뷰의 행 수·해시 일치와 API 응답을 확인했습니다. 실행 기록은 한 호스트의 SQLite에 저장했습니다.
- 이후 RDF/OWL·SHACL로 지표의 단위·버전·필수 근거를 검사하고, 맞지 않으면 답변을 보류하도록 했습니다. 이 부분은 10월 18일 로컬에서 별도로 검사했습니다.

AWS 실행·복구 요약 / 호출 중단 기록 / 의미 검증 범위

ASK Seoul

API 호출 제한과 중복 적재에 맞춰 수집 방법을 고쳤습니다. 개인 작업으로 기상 데이터를 AI가 조회하는 기능을 추가했습니다.

2026.06 - 현재 / 팀 기상·교통 수집·변환 담당 / 2026.08부터 개인 Weather·Agent 개발

Python · Airflow · Iceberg · dbt/Trino · FastAPI SSE · Gemini · MCP

- 팀 기상·교통 API에서 429 오류가 나 요청 간격과 재시도를 조정했습니다. 예보 8만 건 이상을 적재하다 SQL 크기 제한에 걸려 PyIceberg 직접 적재로 바꾸고, 삭제·추가를 한 번에 커밋했습니다.
- 개인 작업으로 기상 데이터 4종을 조회·집계하도록 만들었습니다. 검색·미리보기와 인증이 필요한 MCP 조회를 로컬에 구현하고, SSE로 근거와 갱신 시각을 제공했습니다.
- 단계별 허용 도구를 제한하고 정상·도구 오용·타임아웃 등 30개 모의 회귀 검사를 CI에 연결했습니다. 실제 Gemini 10문항 평가는 별도로 진행했습니다.

Weather 코드 / Agent 코드

GitOps 배포·장애 검증

카나리 배포가 자원 부족으로 멈췄을 때의 원인과 수동 조치 기록을 정리했습니다.

팀 프로젝트 이후 개인 후속 개발 / 2026.10 / 10월 6일 AWS 시험 기록

Kubernetes · Argo CD/Rollouts · Prometheus · CloudWatch

- 10월 6일 기준선·카나리 각 13분의 요청 기록을 비교했습니다. CPU 할당량 6→8코어 수정과 수동 reconcile 후 Healthy, readiness 150/150, 추가 응모 10/10을 확인한 기록입니다.

배포·수동 복구 기록 / 할당량 원인·변경 PR #42

교육·학력·자격

SK플래닛 T Academy ASAC 데이터 엔지니어 과정 / 2026.02 - 09 / 960시간 수료

상명대학교 컴퓨터과학과 / 2016.02 - 2022.08 / 졸업

정보처리기사 2026.09.11 / 데이터분석 준전문가(ADsP) 2022.11