

김민지 Minji Kim

Backend Developer / Data Engineer

 GitHub  Tableau

Summary

Airflow, GCS, RDBMS 등을 활용해 대규모 데이터를 안정적으로 수집·가공·전송하는 데이터 파이프라인을 구축하고 운영해왔습니다. 특히, 주기적 스케줄링과 장애 대응이 포함된 자동화 구조 설계를 통해 시스템의 지속 가능성과 운영 안정성을 높이는 데 집중했습니다. 비즈니스 로직을 구조화된 데이터 아키텍처로 전환하고, 사용자가 신뢰할 수 있는 정제된 데이터를 제공하는 일에 높은 책임감을 가지고 임하고 있습니다. 분산 처리와 예외 상황에서도 견고하게 작동하는 시스템을 설계하는 데 관심이 많으며, 앞으로는 데이터 마트를 통해 조직의 전략적 의사결정을 지원하는 역할로 성장하고자 합니다.

Work Experience

R&D Center Backend Chapter / Backend Developer Aijinet Seoul, South Korea	2022.02 ~ Present
<ul style="list-style-type: none">Airflow 및 GCP 기반의 병렬 처리·스케줄링 시스템 설계 및 운영 자동화보험 약관·보험료 등 정형·비정형 데이터를 대상으로 한 ETL 파이프라인 재구축 및 운영Bodoc 앱의 보험 진단·추천 기능을 위한 엔진 로직 설계 및 TypeScript 기반 API 서버 개발PostgreSQL, Nest.js 기반 SaaS 아키텍처 도입 및 대용량 보험료 데이터 처리 개발Spring Boot 기반의 Bodoc 어드민 페이지 및 보험사 대상 B2B API 개발 및 운영 고도화	

DataLabs / Data Researcher Aijinet Seoul, South Korea	2021.08 ~ 2022.01
<ul style="list-style-type: none">KOICD 질병코드 간 연관성 분석 및 매핑을 통한 약관 데이터와 보상 내용 정제 작업보험 공시 데이터를 기반으로 한 고객 보유 상품의 가성비 분석 로직 기획 및 엔진 개발 지원보험료 및 공사실 데이터 수집을 통한 엔진 메타데이터 구성 및 Pandas 기반 가공	

Projects

B2B SaaS 보장분석 서비스 개발 <i>RabbitMQ 기반 메시지 처리 및 보장 리포트 템플릿 자동화</i>	2025.07 ~ Present
<ul style="list-style-type: none">요구사항 기반 DLQ 로 포함한 메시지 큐 서버 구성 및 장애 대응 구조 설계Spring Boot 기반 Producer/Consumer API 개발 및 큐 상태 실시간 모니터링 처리Thymeleaf 템플릿 엔진 기반 보장 분석 리포트 HTML 바인딩 및 정적 페이지 생성정적 HTML 페이지 생성 후 S3 업로드 자동화	
Skill Keywords: Java, Spring Boot, RabbitMQ, Thymeleaf, S3	

B2B 통계 메타데이터 자동화 시스템 구축 <i>Airflow 기반 대규모 보험업권 마이데이터 중심의 통계치 추출 자동화</i>	2025.06
<ul style="list-style-type: none">최대 1,900만 건 보험 데이터를 벡터화 연산 기반으로 처리하기 위해 Pandas 최적화 적용, 처리 속도 개선Pandera 기반 스키마 검증 로직을 도입하여 데이터 유입 시 타입 안정성과 유효성 보장고객사 분기별 통계 리포트 추출 작업을 자동화하여 수작업 제거 및 운영 안정성 확보	
Skill Keywords: Python, Pandas, Pandera, Airflow, GCS, Amplitude API, Slack API	

Amplitude 사용자 정보 일괄 처리 시스템 개발 <i>Airflow 기반 병렬 처리 ETL 파이프라인 설계 및 GCS 연동 구조 개발</i>	2025.06
<ul style="list-style-type: none">expand 기반 병렬 처리 시스템 신규 구축 및 최적화로 6만 건 데이터 처리 시간 84% 개선Task와 GCS 간 매핑 구조를 설계하여 에러 발생 시 추적 가능성 향상 및 운영 효율 증대Pydantic 기반 입력 데이터 검증 로직을 통해 타입 안정성과 유효성 보장Amplitude API 호출 시 Exponential Backoff 로직 적용으로 장애 대응력 및 재시도 안정성 확보	
Skill Keywords: Python, Pydantic, Airflow, GCS, Amplitude API, Slack API	

전사 KPI 대시보드 데이터 적재 자동화 <i>전사 KPI 시각화를 위한 일별 누적 데이터 ETL 파이프라인 설계 및 자동화</i>	2025.05
<ul style="list-style-type: none">Airflow 기반 DAG 설계 및 MySQL-Google Sheet API 연동을 통한 데이터 적재 자동화Google Calendar API 연동하여 공휴일 여부 판단Task branch를 통해 DAG 실행 흐름을 제어로 리소스 사용 최적화통합 트랜잭션 처리 방식 도입으로 데이터 적재의 일관성과 안정성 확보MSSQL과 연동된 Looker Studio 대시보드에 데이터 노출 및 시각화 환경 구성	
Skill Keywords: Python, Airflow, MSSQL, MySQL, Google Calendar API, Google Sheet API, Slack API	

보험료 기반 추천 데이터 자동화 시스템 구축 <i>보험료 기반 추천 점수 생성을 위한 대용량 데이터 ETL 자동화를 Airflow 기반으로 구축</i>	2025.03 ~ 2025.04
<ul style="list-style-type: none">Airflow 도입 및 GCS 연동을 포함한 ETL 파이프라인 자동화 DAG 설계약 150만 건의 API 요청을 비동기 처리로 구현, 5분 이내 처리 완료하여 생산성 향상내부 엔진 빌드 시 최신 메타데이터 자동 참조 구조로 배포 효율 및 데이터 신뢰도 강화Slack API 연동을 통한 실시간 모니터링 체계 구축으로 운영 안정성 확보	
Skill Keywords: Python, Airflow, GCS, PostgreSQL, Slack API	

Bodoc 어드민 페이지 보안 강화 시스템 개발 <i>금융감독원 보안 가이드 대응을 위한 인증 및 계정 보안 고도화</i>	2025.02
<ul style="list-style-type: none">Spring Security 기반 마이페이지 API 설계 개발비밀번호 인증 및 변경 로직 구현으로 보안 정책 구현JUnit 기반 단위 테스트 커버리지 96% 달성 및 API 명세 자동화(Redoc 기반) 적용	
Skill Keywords: Java, Spring Boot, Spring Security, JUnit, MySQL	

통계 데이터 생성을 위한 배치 시스템 구축 <i>MySQL Event Scheduler 기반 통계 데이터 자동 추출 및 최신화 체계 구축</i>	2025.01 ~ 2025.02
<ul style="list-style-type: none">통계 처리 전용 스키마 설계 및 일괄 처리(Batch) 자동화 구조 구현기존 운영 스키마 분석을 바탕으로 인덱스 재설계 및 최적화 적용인덱스 적용을 통해 통계 데이터 추출 속도 약 90% 개선기존 수동 통계 작업을 주 단위 자동화하여 데이터 최신성 및 운영 효율성 향상	
Skill Keywords: MySQL, Event Scheduler	

보험 약관 크롤링 시스템 재구축 <i>반복적인 수동 크롤링 API 요청 기반 수집 작업 자동화 및 데이터 파이프라인 재구축</i>	2024.04 ~ 2024.12
<ul style="list-style-type: none">전체 ETL 파이프라인 재설계 및 오류 복구 가능한 구조 구현HTML 파싱 및 API 응답 기반 이중 데이터 수집 방식 적용으로 크롤링 범용성 향상월간 오류 발생률 142건에서 15건으로 89% 감소Exponential backoff 기반 요청 재시도 체계 강화로 운영 안정성 확보전체 처리 기간 7일에서 3일로 단축하여 수집 주기 개선분기 단위의 약관 업데이트를 주 단위로 Cron 스케줄링하여 운영 효율성을 91% 향상	
Skill Keywords: Python, GCS, PostgreSQL, Django ORM (transaction.atomic, bulk insert), Slack API	

앱 백엔드 서버 마이그레이션 지원 <i>유지보수성과 구조적 일관성을 위한 Laravel → Spring Boot 기반 백엔드 API 마이그레이션 수행</i>	2024.08 ~ 2024.09
<ul style="list-style-type: none">Laravel 기반 백엔드 서버를 Spring Boot 환경으로 특정 API를 마이그레이션하여 시스템 일관성 기여중복 로직 제거 및 JPA Repository 기반 쿼리 구조 개선을 통해 공통 모듈화 진행모듈 및 쿼리 재사용을 통해 유사 API 개발 소요 시간 5일 → 3일로 단축 (약 40% 감소)	
Skill Keywords: Java, Spring Boot, MySQL, PHP	

간편보험료 기반 보험 상품 비교 시스템 개발 <i>보험설계사들의 수기 설계안 작성을 자동화하는 보험 비교 엔진 구축</i>	2024.01 ~ 2024.03
<ul style="list-style-type: none">Next.js 기반 보험료 계산 및 데이터 제어 API 설계 및 성능 최적화보험설계사의 상품 비교 소요 시간 80% 단축, 계약 전환율 4%p 향상TypeORM 도입 및 쿼리 최적화를 통해 응답 속도 90% 개선데이터 검증용 API 추가 개발로 정합성 및 서비스 신뢰도 확보	
Skill Keywords: TypeScript, Node.js, Express.js, MongoDB, GCP Cloud Run, Redoc, Memcached, Python	

SaaS 보험 엔진 고도화 및 개발 <i>TypeScript 기반 보험 진단·추천 엔진 고도화 및 개발</i>	2022.05 ~ 2025.07
<ul style="list-style-type: none">Bodoc 앱 진단·추천 탭의 비즈니스 요구사항 반영 및 안정적 유지보수 수행건강 연동 여부에 따라 엔진 엔드포인트 분기 처리하여 건강 맞춤형 보험 결과 제공 로직 구현GCP Cloud Run 환경에서 B2C/B2B 서버 분리 및 Github branch 체계 구축환경별 배포 파라미터를 반영한 GCP Cloud Run 배포 명령어를 스크립트화하여 배포 절차 자동화자동화 문서화 도구 Redoc 도입으로 API 문서 관리 생산성 향상메타데이터 캐싱을 위한 Memcached 적용으로 초기 로딩 성능 30% 개선유사도 기반 Python 스크립트를 도입하여 상품 메타데이터 갱신 업무 시간 70% 단축, 성공률 4%p 개선비개발 부서 대상 엔진 시뮬레이션 API 설계로 정책 변경 변경 테스트 환경 제공증권사 마이데이터 기반 보장 분석 기능에 TypeGuard 및 암호화 토큰 기능 설계 참여마이데이터 및 신용정보원 보험 데이터 비교 분석하여 진단·추천 엔진 로직 기획 및 구현 참여	
Skill Keywords: Node.js, TypeScript, Express.js, MongoDB, Redoc	

Skills

</> Languages & Runtimes Python, Java, TypeScript, JavaScript, Node.js	🔧 Frameworks & Libraries Spring Boot, Nest.js, Express.js, Django
--	---

☁ Infrastructure & Databases GCP, Airflow, MySQL, PostgreSQL, MSSQL, MongoDB, Docker	🔧 Tools & IDEs Bash, Zsh, Git, Gitlab, Github, Jira, Confluence, Notion, IDEA / PyCharm / DataSpell / Datagrip / Webstorm
--	---

Education

경영학과 Business Analytics 융합연계전공 세종대학교 Seoul, South Korea	2017.02 ~ 2022.02
--	-------------------

Total GPA of 3.84 / 4.5

Major GPA of 3.94 / 4.5

Awards

지역 데이터를 활용한 데이터 분석 및 시각화 "광진구 시장 활성화 프로젝트" 세종대학교/서울특별시 - 대학총장상	2021.12
--	---------

Skill Keywords: Tableau, Python, Geopandas

대학 팀프로젝트 경진대회, "삼성전자 주식 예측 보고서" 세종대학교 - 최우수상	2020.02
--	---------

Skill Keywords: Python, scikit-learn

대학생 영화 수요 예측 경진대회, "개봉 전후 영화 흥행 예측 분석을 통한 의사결정" 소니전자 - 장려상	2020.01
--	---------

Skill Keywords: Tableau, Python, scikit-learn

전국 대학생·일반인 고조선 독서논술대회 겨레열살리기국민운동본부 - 2위 은상	2019.10
--	---------

교내 독서경시대회 "사소한 것들" 세종대학교 - 3위	2018.10
---	---------

전국 대학생·일반인 고조선 독서논술대회 겨레열살리기국민운동본부 - 장려상	2018.10
--	---------

신입생 입시 합격수기 공모전 세종대학교 - 최우수상	2017.04
--	---------

Certificates

정보처리기사 한국산업인력공단 2024.06	TEPS 358점 서울대학교발전재단 2024.09
--------------------------------------	--

Interests

Sports

Pilates, Golf, Weight Training, Running