

프로필

웹 개발로 커리어를 시작해 공공·민간 데이터 플랫폼에서 수집, 적재, 모델링, 오케스트레이션, BI까지 데이터의 전체 라이프사이클을 경험했습니다.

화려한 기술 자체보다 비효율의 원인을 찾고, 운영 가능한 표준과 자동화로 바꾸는 문제 해결 과정에 강점이 있습니다.

핵심 역량

배치 최적화

1,000+ dbt 모델의 워크로드 격리·배치 표준화

플랫폼 전환

Oracle·SAS 기반을
Airflow·dbt·Redshift로 전환

데이터 신뢰성

품질 검증·모니터링·장애 복구 체계 자동화

경력

누리꿈소프트(주)

데이터 엔지니어 · 프로

2021.07 - 2026.05 · 4년 11개월

루나랩주식회사

웹 개발자

2021.01 - 2021.07 · 7개월

핵심 기술

Python

SQL

Apache Airflow

dbt

Amazon Redshift

AWS S3

AWS Glue

Amazon Athena

01 주요 프로젝트 경험

민간 프로젝트

2025.06 - 2026.03

배치 최적화 매니저

GS리테일 HBU 데이터 플랫폼 전환

기여도 100% · 배치 최적화 담당 영역 전반 주도

Oracle·SAS 기반 홀소핑 데이터를 Airflow, dbt, Redshift 기반 플랫폼으로 전환하고 배치 시간, 비용, 품질 거버넌스를 개선했습니다.

Apache Airflow 2.10 · dbt · Amazon Redshift · AWS ECS Fargate · Grafana · Bitbucket · Bamboo · Confluence

워크로드 격리와 비용 최적화

문제	1,000개 이상의 dbt 모델이 특정 시간대에 Provisioned 클러스터의 쿼리 큐잉과 리소스 경합을 유발했습니다.
실행	manifest.json 의존성을 파싱하고 헤비 모델을 Redshift Serverless로 분리 실행하는 태그 기반 파이프라인을 설계했습니다.
성과	상시 클러스터 증설 없이 피크 타임 병목을 제거하고 불필요한 비용을 방어했습니다.

Stateful Backfill 체계

문제	Hourly 배치의 소급 범위가 4시간으로 하드코딩되어 장시간 장애 시 복구에 제약이 있었습니다.
실행	Airflow Variable로 마지막 성공 data_interval_end를 추적하는 상태 관리 로직을 도입했습니다.
성과	안전 소급 범위를 24시간 이상으로 확장해 데이터 유실 위험을 낮췄습니다.

대규모 오케스트레이션 표준화

문제	단일 DAG에서 대규모 모델을 파싱할 때 렌더링 지연과 워커 슬롯 낭비가 발생했습니다.
실행	도메인별 Sub DAG와 Control DAG 구조로 분리하고 deferrable 실행을 템플릿화했습니다.
성과	파싱 부하와 워커 점유를 줄이고 다수 수행사가 함께 사용할 수 있는 운영 표준을 만들었습니다.

01 주요 프로젝트 경험 · 계속

민간 프로젝트

2025.06 - 2026.03

배치 최적화 매니저

GS리테일 HBU 데이터 플랫폼 전환

기여도 100% · 배치 최적화 담당 영역 전반 주도

데이터 품질 가드레일

문제	Redshift가 PK 유일성을 강제하지 않아 멀티 수행사 환경에서 중복 데이터 인입 위험이 있었습니다.
실행	핵심 모델의 dbt test 의무화와 배치 후행 전수 중복 검증 DAG를 구축했습니다.
성과	사람의 주의력에 의존하지 않는 이중 데이터 품질 검증 체계를 확립했습니다.

메타 DB 부하 개선

문제	Grafana의 비효율적인 메타 DB 쿼리가 Airflow 스케줄러 지연을 유발했습니다.
실행	인덱스를 고려한 경량 쿼리로 대시보드를 재설계했습니다.
성과	스케줄러 응답을 정상화하고 태스크·Pool Slot·지연을 모니터링을 안정화했습니다.

공공 프로젝트

2023.09 - 2024.12

공통 지원 선임

차세대 KICS 데이터 플랫폼 구축

기여도 70% · 공통 표준·자동 수집 핵심 담당

다수 사법기관의 이기종·비정형 데이터를 안정적으로 수집하기 위한 공통 표준과 자동 수집 프로세스를 구축했습니다.

TeraStream · Hive · Hadoop · Python · Shell

이기종 데이터 표준화

문제	4개 기관별 데이터 형식 차이로 BLOB 적재 오류가 반복됐습니다.
실행	ETL 공통 개발 가이드와 데이터 표준을 수립했습니다.
성과	적재 실패를 줄이고 파이프라인 안정성을 높였습니다.

법령 수집 무인화

문제	법령 정보 수집이 수작업에 의존하고 있었습니다.
실행	메타데이터의 수집 여부를 판단하는 Python 전처리·수집 프로그램을 개발했습니다.
성과	반복 작업을 자동화하고 법령 데이터의 최신성을 보장했습니다.

공공 프로젝트

2022.05 - 2023.08

OLAP 개발 선임 · Admin

차세대 지방재정관리시스템 데이터 플랫폼

기여도 70% · OLAP·Admin·외부 연계 핵심 담당

행정안전부와 전국 지자체가 사용하는 정책지원서비스의 통계·BI 화면과 외부 데이터 연계 파이프라인을 구축했습니다.

BI-Matrix · InnoQuartz · SAP HANA · Oracle · Python

One-Table 모델 도입

문제	화면별 물리 테이블 방식으로는 200여 개 통계 화면의 유지보수 비용이 급증할 가능성이 컸습니다.
실행	화면 ID를 파티션 키로 사용하는 통합 테이블 모델과 공통 규칙을 설계했습니다.
성과	200여 개 화면을 기한 내 완수하고 신규 화면 개발 생산성을 높였습니다.

행정 비즈니스 로직 전산화

문제	자치단체 유형분류와 재정위기관리 산출이 외부 용역과 수작업에 의존했습니다.
실행	담당 공무원과 요구사항을 조율하고 내부 데이터와 Open API를 결합한 지표 산출 파이프라인을 구현했습니다.
성과	유형분류·위기관리 심화 화면 30여 본을 성공적으로 오픈했습니다.

02 추가 프로젝트 경험

민간 프로젝트

2021.07 - 2022.05

데이터 분석 및 ETL 선임

LG화학 전사 데이터 레이크 구축 1차

기여도 70% · 원천 분석·적재 전략 핵심 담당

전사 정형 데이터를 AWS 데이터 레이크로 이관하기 위한 원천 분석과 적재 전략 수립을 담당했습니다.

데이터별 적재 전략 최적화

문제 모든 원천 데이터를 Full Load하면 처리 시간과 클라우드 비용이 불필요하게 증가할 수 있었습니다.
실행 볼륨, 휘발성, 예상 증가량, 갱신 주기를 프로파일링해 Full-Incremental 전략을 세분화했습니다.
성과 초기 데이터 레이크 파이프라인의 경제성과 처리 효율을 확보했습니다.

2021.03 - 2021.05

한화솔루션 ESH Portal

임직원 건강관리와 환경 공정시설을 통합 관리하는 포털을 개발했습니다.

기여도 70% · Backend-Frontend 주요 구현

Vue.js · Spring Framework · MSSQL

2021.02 - 2021.03

UBASE 관리자 인사평가 시스템

평가 히스토리 관리를 도입하고 인사평가 메뉴의 전체 로직을 재설계했습니다.

기여도 70% · 인사평가 메뉴 전면 재설계

Java · Spring Framework · JavaScript · Oracle

03 기술 및 자격

기술 역량

프로그래밍·쿼리 언어	Python · SQL · Java
오케스트레이션·변환	Apache Airflow · dbt
데이터 웨어하우스·DB	Amazon Redshift · Oracle · PostgreSQL/pgvector · MySQL · MSSQL · SAP HANA
클라우드·빅데이터	AWS S3 · AWS Glue · Amazon Athena · Hadoop · Hive
개발 운영·CI/CD	Git · Bitbucket · Bamboo · Docker · GitHub Actions
문서화·협업	Confluence · Jira (진척도 확인)
AI·에이전트	HyperCLOVA X · LangGraph · RAG

학력

국가평생교육진흥원
컴퓨터공학 · 2019.01 - 2025.02

서일대학교
생활가구디자인과 · 2016.03 - 2018.03

어학

영어 · OPIC IM2
2026.07

자격증

	AWS Certified Solutions Architect - Associate Amazon Web Services · 2026.09.01
	AWS Certified Data Engineer - Associate Amazon Web Services · 2026.08
	빅데이터분석기사 한국데이터산업진흥원 · 2025.07
	SQLD 한국데이터산업진흥원 · 2023.12
	정보처리기사 한국산업인력공단 · 2024.09
	리눅스마스터 2급 한국정보통신진흥협회 · 2024.1
	컴퓨터활용능력 2급 대한상공회의소 · 2024.1

직무 교육 및 프로젝트

NIPA-NAVER Cloud Sovereign AI PBL 기반 프로젝트

2026.08 · 4인 팀 프로젝트 · 우수상 · 기여도 40%

HyperCLOVA X, RAG, LangGraph를 결합한 영화 추천 멀티 챗봇 에이전트와 NCP 기반 서비스 인프라를 구현하고 최종 평가에서 우수상을 수상

Python·R 기반 빅데이터 플랫폼 개발자 과정

2020.04 - 2020.10

빅데이터 플랫폼 요구사항 분석, 수집·저장·처리·분석 시스템 개발과 품질 관리