

입수부터 변환·품질·서빙·인프라까지, 대규모 서비스 로그와 금융 데이터를 믿고 쓸 수 있는 구조로 만드는 데이터 엔지니어입니다.

카카오 추천팀과 CryptoQuant를 거치며 대규모 사용자 로그와 금융 데이터를 다루는 파이프라인과 데이터 플랫폼을 설계·운영해 왔습니다.

카카오에서는 특·선물하기·광고·다음 뉴스 등 여러 서비스의 대용량 사용자 로그를 실시간으로 통합·전처리하는 파이프라인을 운영했고, Apache Airflow 2.0 도입을 주도해 팀의 배치 오케스트레이션 체계를 정비했습니다.

CryptoQuant에서는 미국 증권사 등 기관 고객을 대상으로 한 금융 데이터 프로젝트를 담당했습니다. dbt로 BigQuery 변환 파이프라인을 표준화하고 대규모 연산을 Spark 분산 처리로 전환해 레이턴시를 줄였으며, PostgreSQL 중심의 API 서빙 구조를 Databricks·Delta Lake 기반으로 옮기고 Delta UniForm으로 Iceberg 카탈로그를 제공해 고객이 Snowflake·BigQuery에서 데이터를 직접 사용할 수 있는 채널을 열었습니다. HA Airflow, EKS/Karpenter 마이그레이션 등 파이프라인을 받치는 인프라도 직접 구축·운영했습니다.

일하면서 중요하게 여기는 것은 요청을 하나씩 처리하는 것보다, 같은 종류의 요청이 다시 생기지 않도록 구조를 만드는 것입니다. 데이터 소비 방식이 다른 고객이 늘어날 때마다 API를 덧붙이는 대신 Open Table Format으로 소비 채널 자체를 바꾼 것도 같은 맥락이었습니다.

개인적으로는 암호화폐 파생상품과 DeFi를 활용한 퀀트 트레이딩 시스템을 직접 구축·운영하며, 데이터가 실제 의사결정과 손익으로 이어지는 과정을 체감하고 있습니다. 최근에는 AI 코딩 에이전트를 엔지니어링 업무에 적극 적용하고 있고, 사람뿐 아니라 AI 에이전트도 이해하고 활용할 수 있는 데이터 구조에 큰 관심을 두고 있습니다.

업무 경험

CryptoQuant

2023.08 ~ **재직 중**

Web3 데이터 엔지니어

블록체인·마켓 데이터 입수부터 변환, 기관 고객 대상 데이터 서빙, 파이프라인 인프라까지 데이터 플랫폼 전 구간 담당

데이터 플랫폼 현대화 — Open Table Format 기반 서빙 전환

- 미 증권사 등 기관 고객의 다양한 데이터 소비 니즈에 대응하기 위해, PostgreSQL 중심 API 서빙 구조를 Databricks 기반 Delta Lake로 마이그레이션
- Delta UniForm으로 Iceberg 카탈로그를 제공해 고객의 Snowflake·BigQuery 직접 연동 지원, API 외 데이터 세일즈 채널 확보

분석 파이프라인 표준화 및 고도화

- BigQuery 기반 분석 플랫폼의 데이터 파이프라인 관리 및 최적화
- dbt로 데이터 변환 파이프라인 표준화, 대규모 연산을 Spark 분산 처리로 전환해 데이터 레이턴시 단축

XRP Ledger 데이터 프로젝트

- XRP Ledger 및 마켓 데이터를 PostgreSQL·BigQuery로 수집하는 파이프라인 개발
- 실시간 업데이트와 트랜잭션 데이터 디코딩을 포함해 경쟁 우위의 데이터 확보

워크플로우 오케스트레이션 및 인프라

- ArgoCD와 Kustomize-Helm으로 관리되는 HA Airflow 구축·운영으로 파이프라인 안정성과 운영 효율성 확보
- EC2 서비스를 EKS로 전환하고 Karpenter 기반 오토스케일링 도입, SPOT 인스턴스와 Kubernetes 재시작 로직으로 배치 워크로드를 빠르고 유연하게 프로비저닝

카카오

2019.12 ~ 2023.08

추천팀 · ML/데이터 엔지니어

카카오·다음 주요 서비스의 추천 시스템과 이를 받치는 대용량 로그 파이프라인 개발·운영

대용량 사용자 로그 통합 파이프라인

- 카카오톡·선물하기·광고, 다음 뉴스·스토리 등 서비스의 대용량 사용자 로그를 실시간으로 통합·전처리하는 파이프라인 개발 및 운영
- 다수 서비스의 추천 시스템 개발 및 유지보수로 사용자 참여도와 경험 향상에 기여

Airflow 2.0 도입 주도

- Apache Airflow 2.0 도입을 주도해 태스크 오케스트레이션의 확장성과 운영 효율성 향상 (if(kakao) 발표)

검색 키워드 광고 추천 파이프라인

- 다음 검색의 쿼리·소비 이력으로 pCTR을 예측하고 키워드 광고를 추천하는 데이터 파이프라인 구축
- VectorDB(ANN) 기반 서빙으로 50ms 이내 응답 달성

프로젝트

퀀트 트레이딩 시스템 구축 및 운용

2025.01 ~ 진행 중

개인 프로젝트

암호화폐 파생상품과 DeFi를 활용한 체계적 트레이딩 전략을 직접 설계·구현·운용하는 개인 프로젝트

- Uniswap LP 포지션과 선물 포지션을 합성한 델타 헤지 마켓 메이킹 전략 설계 및 실전 운용
- 선물·옵션 기반 합성 전략 구현, 포지션 관리와 리스크 모니터링 자동화
- 온체인/오프체인 마켓 데이터 수집·분석 파이프라인을 구축해 시그널 생성 및 전략 백테스팅에 활용

전문 분야

대표 스킬

dbt

Airflow

Spark

스킬

Python

SQL

BigQuery

Databricks

Delta Lake

Apache Iceberg

Kafka

PostgreSQL

Kubernetes

AWS

ArgoCD

Helm

Karpenter

Prometheus

Grafana

Redis

MongoDB

산업 지식

금융

콘텐츠

광고·마케팅

학력

연세대학교
산업공학과 학사

2013.02 ~ 2019.08

활동

Decipher 블록체인 리서치 그룹

2023.08 ~ 2024.08

스터디, 컨퍼런스 발표, 리서치 글 기고

- 라이트닝 네트워크 노드 운영 리서치 (YouTube, 한국어)
- 모놀리식 체인의 스케일링 솔루션 (아티클, 한국어)
- 온체인 데이터 실전 분석 (한국어·영어)
- Interplanetary Money (아티클, 한국어)

외국어

영어
일상 회화