

Chloe Minkyung Kim

Singapore | (+65) 8923 7528 | chloeminkyung@gmail.com | chloeminkyung.github.io
Python · Java · Snowflake · Cloud Data Platforms · Distributed Systems

Education

Hong Kong University of Science and Technology

Sep 2016 – Jan 2020

Bachelor of Engineering in Computer Science

Experiences

Goldman Sachs, Singapore

Apr 2024 – Present

Associate, Data Platform Engineering

- Leading design and implementation of the **Lakehouse** platform — a global unified cloud data warehouse architecture combining data lake flexibility with enterprise-grade performance, governance, and analytics capabilities.
- Building enterprise-scale ETL pipelines and data transformation frameworks using Java, Python, SQL, and **Pure**, integrating Snowflake cloud data warehouse with **Legend** semantic layer and **Alloy** execution engine to process billions of records daily with strict SLAs.
- Architected and delivered real-time and near-real-time data pipelines, reducing end-to-end latency from multi-day batch processing to sub-10-minute streaming ingestion for globally distributed datasets, enabling time-sensitive analytics and operational reporting at scale.
- Designed comprehensive data governance frameworks including fine-grained access controls, row-level security, data lineage tracking, and audit capabilities to ensure regulatory compliance and establish the platform as a trusted single source of truth.
- Driving operational readiness and production rollout strategy, defining SLAs, monitoring dashboards with Prometheus and Grafana, alert mechanisms, and coordinating phased deployment across Americas, EMEA, and APAC regions.

Dozer, Singapore

Dec 2022 – Mar 2024

Founding Software Engineer

- Architected and built a high-performance real-time data platform in Rust achieving 4.92 ms average query latency under 10,000 concurrent requests, designed specifically for low-latency analytics and data exploration workloads with sub-second response times.
- Developed production-grade streaming data connectors for Postgres, Snowflake, Ethereum blockchain, Kafka, and AWS S3, building a custom SQL query engine on RocksDB that executed complex joins across 100M+ record datasets in under 10 minutes with optimized query planning.
- Designed distributed snapshot and checkpoint mechanisms for cloud object stores (S3, DeltaLake), implementing exactly-once semantics and fault tolerance for high-throughput streaming data pipelines processing millions of events per hour.
- Built cloud-native deployment infrastructure on AWS EKS and Kubernetes with full CI/CD automation, containerized microservices architecture, and integrated observability using Prometheus and Grafana for production monitoring.
- Represented the company at major data engineering conferences (Big Data & AI World Singapore, Big Data Conference Europe, APIdays Paris), delivering technical presentations on real-time data architecture and engaging with enterprise customers.

Goldman Sachs, Singapore

Jul 2020 – Dec 2022

Analyst, Data Platform Engineering

- Designed, implemented, and open-sourced a **temporal data versioning library** for Snowflake and H2 using Java and SQL, enabling bi-temporal data tracking and point-in-time queries for large-scale enterprise data warehouse applications supporting audit and regulatory requirements.
- Contributed to the Legend open-source data platform ecosystem (FINOS Foundation), enhancing the **Studio** modeling tool and persistence service to improve developer productivity, reduce data pipeline development time, and strengthen data quality testing capabilities.
- Optimized the Alloy Streaming ETL platform built on Kafka and Kubernetes: implemented comprehensive monitoring with Prometheus and Grafana, built automated alerting and recovery systems, and reduced P95 query latency by 40% through Java performance tuning and resource optimization.

- Operated production Kafka-based distributed data pipelines on Kubernetes clusters, managing high-volume message ingestion (1M+ events/sec) with exactly-once processing guarantees, zero data loss, and 99.9% uptime SLAs across multiple data domains.

JP Morgan Chase & Co., Hong Kong

Jun 2019 – Aug 2019

Summer Analyst, Technology

- Built a personalized recommendation engine for the Private Banking platform using Python, Java, Flask, and Angular, deployed on JPMC's internal cloud (Gaia) with collaborative filtering and user-clustering algorithms, processing customer interaction data to deliver personalized content experiences.

Imago.AI, Hong Kong

May 2018 – Aug 2018

Machine Learning Engineer

- Developed neural Question-Answering models using Bi-Directional Attention Flow with self-attention mechanisms, trained on SQuAD and MS MARCO datasets for production information retrieval systems serving real-time query workloads.
- Built a Chinese Named Entity Recognition model with BiLSTM + CRF architecture for Hong Kong address extraction, improving downstream data processing accuracy by 35% and enabling automated data enrichment pipelines.

Projects & Technical Leadership

News Data Processing for Fake News Detection, Final Year Project

May 2018 – May 2019

- Architected a multi-modal deep learning pipeline combining text, image, and query-based modules with attention mechanisms for fake news classification, achieving 89% accuracy through ensemble modeling and feature engineering.
- Implemented a fact-checking system using BiDAF and AllenNLP models with a claim-and-evidence verification approach, reducing false positives by 25% and enabling real-time content verification at scale.
- Deployed trained models as production-ready REST APIs with scalable infrastructure, enabling integration with external applications for real-time news verification and automated content analysis workflows.

BundlePort, Startup Project

Jan 2018 – May 2019

Co-Founder / CTO

- Co-founded and led engineering for a full-stack logistics management platform built with Node.js, MySQL on AWS (EC2, S3, Elastic Beanstalk, CloudFront), serving retail merchants with cloud-based workflow automation and real-time inventory tracking.
- Owned end-to-end product architecture, cloud infrastructure design, technical roadmap, and database schema optimization, managing development cycles and coordinating cross-functional teams between engineering and business operations.
- Improved transaction processing efficiency by 60% through automated ETL workflows, optimized SQL queries, and UI/UX enhancements, reducing manual operations and improving data accuracy for partner merchants.

Technical Skills

Programming Languages Java, Python, Rust, C++, SQL

Machine Learning PyTorch, TensorFlow

Cloud Data Platforms Snowflake, BigQuery, PostgreSQL, MySQL

Cloud & Infrastructure AWS, Kubernetes, Docker

Spoken Languages English (Fluent), Korean (Native), Mandarin (Intermediate)

Honors and Awards

- JPMorgan Chase & Co. Internal Innovation Week Hackathon 2019 Winner (JPMC Trends), 2019
- SmarTone Hackathon 2018 "Smart Properties", Interactive Property Chatbot - Top 3 Project, 2018
- Targeted Scholarship under HKSAR Government Scholar Fund, Full Scholarship, 2016 - 2020
- Admission Scholarship from HKUST, Full Scholarship, 2016 - 2020
- Mensa Membership from Malaysia Mensa (Mensa International), High IQ society, 2015