

Kenshin Tanaka

778-222-1023 | kemkemg0@student.ubc.ca | github.com/kemkemG0 | linkedin.com/in/kemkemg0

EDUCATION

University of British Columbia
BSc in Computer Science, 4th year

Vancouver, BC
Expected Graduation: April 2025

EXPERIENCE

RestUp

Software Engineer (Contract)

San Francisco, CA

July 2024 – Present

- Developed and deployed an **ETL pipeline** using **Kafka**, **AWS Firehose**, **Lambda**, and **Redshift Serverless** to process and analyze large datasets.

MemoryLab

Software Engineer (Part-Time)

Tokyo, Japan

August 2023 – Present

- Enhanced a search engine using **Qdrant** on **AWS EKS**, increasing query accuracy by implementing **RAG** (Retrieval-Augmented Generation).
- Optimized a conversion pipeline for transforming natural language queries into **vector embeddings**, improving the relevance of search results.
- Designed and implemented a high-performance **REST API** in **Go**, enabling seamless integration with a machine learning server through event queuing. Utilized **interfaces** to create a **loosely coupled** implementation, allowing smooth transition from **v1** to **v2**.
- Coordinated the deployment of microservices using **AWS ECS**, ensuring seamless integration and performance within the search engine infrastructure.
- As a **Lead Engineer**, managed tasks and monitored progress using **Kanban** for effective project management. Ensured smooth workflow by prioritizing tasks, assigning responsibilities, and tracking the team's progress. This approach facilitated transparent communication, quick identification of bottlenecks, and timely delivery of project milestones.

PLAID

Software Engineer (Internship)

Tokyo, Japan

May 2023 – August 2023

- Implemented a **Go-based Dataflow job** for efficient event processing from **Cloud Pubsub** within **GCP**.
- Integrated **Sentry** and **Datadog** for real-time system monitoring, maintaining system stability.
- Optimized the **E2E CI/CD pipeline** on **GitHub Actions** by using **git diff** to apply only the differences in **Cloud Spanner** migrations instead of reapplying from scratch, resulting in over a **50%** reduction in **E2E execution times**.

Contract Freelance Engineer

Full-stack Web Developer

Remote

November 2022 – December 2023

- Led the development of a **KYC web application** for streamlining client onboarding in the legal sector, from concept to deployment.
- Built the front-end using **Nuxt.js** and the back-end using **TypeScript** and **Express**, orchestrated with **Docker Compose**.
- Deployed infrastructure on **AWS ECS** and **AWS Amplify**, utilizing **Terraform** for efficient management and deployment.

SOAT Corp

Backend and ML Engineer (Part-Time)

Tokyo, Japan

June 2021 – May 2022

- Developed an advanced anomaly detection program using **Python** and machine learning techniques, which included feature extraction, data preprocessing, and model selection. This system was implemented to identify defects in manufacturing processes, resulting in a **40%** increase in defect detection accuracy and a **25%** reduction in false positives.
- Developed an **OCR tool** for extracting actionable data from financial documents, streamlining data extraction processes.
- Contributed to the concurrent development of multiple large-scale web applications, each serving thousands to tens of thousands of users, utilizing a range of technologies from modern SPA frameworks like **React** and **TypeScript** to traditional MVC frameworks like **Express**, **Django**, and **Laravel**.

OPEN SOURCE SOFTWARE (OSS) CONTRIBUTIONS

Qdrant | *Rust-based high-performance distributed vector database*

- Separated the implementation and interface of snapshot storage, which previously supported only the local file system, allowing for flexible selection of storage solutions such as **S3**, **Google Cloud Storage**, and **Azure Storage**. This feature was released in **Qdrant 1.10.0**. Detailed documentation can be found [here](#).
 - Link to PR: github.com/qdrant/qdrant/pull/4150
- Resolved an issue where processes would crash upon reaching disk capacity limits. To avoid throughput degradation caused by querying storage capacity on each write, I proposed and implemented a heuristic approach that adjusts the frequency of capacity checks based on remaining storage, introducing **DiskMonitoringManager**.
 - Link to PR: github.com/qdrant/qdrant/pull/4165
- Earned approximately **\$700 USD** in **bounties** for these contributions.

PROBLEM SOLVING & PERSONAL HISTORY

- Ranked in the top **15%** in **competitive programming**, showcasing strong algorithmic and problem-solving skills.
- Started programming at the **age of 13**. The first computer purchased was a **RaspberryPi**, where I began working with **Python**, specifically **Python2** rather than **Python3**. At the **age of 14**, I built own PC and created games using **DirectX** with **C++**.