

Tony Su

SF Bay area, CA | boyusu18@outlook.com | (447) 902-7657 | <https://www.linkedin.com/in/boyusu>

SKILLS

- **Programming Languages:** Java, Go, Javascript, Typescript, Kotlin, Python, C/C++, PHP, SQL, HTML/CSS, Lua
- **Framework and Databases:** React, MySQL, PostgreSQL, MongoDB, Kafka, ClickHouse, RESTful API, Flask, Django, Gin, Next.js, Charts.js, Tailwindcss, Pytorch, SQLite, Neo4j, Redis, Apache Airflow, RabbitMQ, Celery
- **Tools:** Claude Code, CodeX, Grafana, Prometheus, Protobuf, Git, Kubernetes, Docker, Jenkins CI/CD, Jira, Postman, Azure, GCP, AWS, Heroku, Figma (UI), Selenium

EDUCATION

University of Illinois Urbana-Champaign

Master of Computer Science

Jan. 2024 - May 2025

GPA: 3.9/4.0

National Yang-Ming Chiao-Tung University, Taiwan

B.S. in Computer Science, Cross Discipline Specialty in Finance Technology

Sep. 2018 – Jun. 2022

GPA: 3.9/4.3

WORK EXPERIENCE

NewsBreak

Software Engineer

Mountain View, CA

July 2025 – present

- Built a self-service ClickHouse-backed ad-index snapshot explorer with a streaming Java pipeline and ClickPipes S3 ingestion, cutting query latency from 40s to 300ms, eliminating JVM OOM, and removing the SRE dependency
- Instrumented the ad-index pipeline with a Protobuf/Kafka + Schema Registry filter log, exposing why ~97% of ad sets are dropped across a 9-stage funnel and enabling self-serve drill-down that replaced manual engineering investigations
- Re-architected a Kotlin, S3, MySQL, MongoDB media processing pipeline with Grafana monitoring, achieving 90% latency improvement, scaling to 2,500+ daily tasks while ensuring the stable serving of 176K existing GIF ads
- Automated the NetSuite billing workflow to eliminate manual accounting operations, reducing daily data entry time from 8 hours to 10 minutes (~98% reduction) and enabling the accounting team to focus on high-value financial analysis

OTSO Fintech

Software Engineer

Taipei, Taiwan

Aug. 2022 – Feb. 2023

- Developed 5 Go Gin microservices for a social trading platform with a 2 billion trading volume, including a system logging service, trading signal query service with Azure, Docker, Kubernetes, gRPC, PostgreSQL, and Drone (CI/CD)
- Re-factored a trading performance calculation system, improving server response time by an average of 120% across 13 RESTful APIs for 5k trading signals through incremental computation and Redis cache
- Led the development of a trading bot deployment service using Python and React; coordinated with a 5-member crossfunctional team (including Engineering, Product Management, and UI/UX) within a Jira Agile environment
- Engineered a Flask backend service with a quote buffering system, trading bot managing system, and ordering system, enhancing reliability across 3 broker exchanges, using Websockets, RESTful API, and Multithreading
- Designed 7 quant trading bots with technical indicators package TA-Lib and ML models such as XGBoost and LGBM

QuantX

Co-Founder and Lead Developer

Remote

Jul. 2021 – Dec. 2024

- Founded QuantX, an algorithmic trading startup with a monthly trading volume of 800K across 2320 equities
- Architected an investing management service handling 1M+ transactions with Python FastAPI featuring RESTful API, JWT and OAuth2.0; developed the dashboard's frontend using Figma, React, Typescript, MUI, and Next.js
- Constructed an automated deployment pipeline using Docker and Gitlab CI/CD on GCP, reducing deployment time by over 90% through optimized integration of Cloud Engine, Cloud Storage, and Cloud SQL
- Built a distributed data pipeline using Apache Airflow, RabbitMQ, and Celery, processing 60M+ financial data into MySQL; constructed an automatic trading program with chatbot interface, integrating multiple brokers' APIs

Trend AI

Software Development Intern

Taipei, Taiwan

Jul. 2024 – Aug. 2024

- Engineered an LLM-based vulnerability prioritization workflow using LangChain, GPT-4o, Azure, and Selenium, solving inefficiency in rules creation by 70%; deployed 2 threat detection rules over 2.25M customer's edge devices
- Developed an LLM-based (GPT-4o, LLaMA3-70b) AI Pull Request review system using CodiumAI PR-Agent and GitHub Actions for CI/CD, improving code quality, security, and team review efficiency by 75%