Keith Chew

Software Engineer

Singapore | (+65) 9789 5623 | keithchewzk@gmail.com | linkedin.com/in/keithchewzikai

Summary

Results-driven Software Engineer with a proven track record in designing, developing, and scaling enterprise-grade applications. Proficient in backend systems, cloud architecture, and developing generative AI applications. Notably engineered a multimodal LLM assistant serving 35,000+ global users, where leading the development of a core feature resulted in a 15% increase in user activity, and achieved a 40% reduction in hosting costs through strategic modernisation. Adept at end-to-end software development, from building robust full-stack web applications to implementing CI/CD pipelines and enhancing system observability. Driven by a passion for solving complex technical challenges that address business needs and deliver value.

Skills

Languages: Python, TypeScript, JavaScript, SQL, HTML, CSS
Developer Tools: Docker, Git, GitHub Actions, AWS, PostgreSQL, Kubernetes, Terraform, OpenTelemetry
Frameworks: FastAPI, ReactJS, NextJS, React Native
Technical Concepts: AI/ML, LLMs, RAG, CI/CD, Observability, NLP, Cloud Architecture, FinOps

Professional Experience

Software Engineer – MSD   Jan 2024 — Present

Engineered and scaled the Python/FastAPI backend of an enterprise-grade multimodal LLM assistant serving **35,000+ monthly users** across 100+ countries and 15 global divisions, enabling GenAI-powered productivity, automation, and decision support across the organisation.

- Led the design of an orchestration framework integrating web search, document retrieval, and image processing in a unified chat interface, simplifying user interaction and improving response quality over legacy workflows, leading to a **15% increase in average monthly user activity**.
- Implemented a feedback mechanism for LLM-generated responses, capturing **2,000+ user inputs** that informed product quality analysis and guided roadmap decisions.
- Facilitated infrastructure migration from Amazon EC2 to EKS, enabling a scalable microservice architecture that improved system reliability, enabled **zero-downtime** releases and **reduced hosting costs by 40%** through Fargate serverless integration.
- Built CI/CD pipelines with GitHub Actions for automated testing, vulnerability scanning, and Kubernetes deployment; maintaining **80%+ test coverage** and meeting monthly release targets.
- Developed a proof-of-concept integration of internal news feeds and technical documentation into the LLM assistant, laying the foundation for enterprise knowledge retrieval from proprietary data sources.
- Integrated OpenTelemetry into backend services, establishing observability capabilities designed to facilitate future reductions in mean time to resolution (MTTR) for performance and reliability issues.

Built and maintained a full-stack enterprise URL shortening and management tool with React, powering **14,000+ registered links** across **3,400+ users**, and improving link governance, campaign tracking, and stakeholder visibility.

- Engineered comprehensive features that enhanced platform functionality and user experience, including secondary link ownership, ownership transfer, and advanced keyword search capabilities.
- Integrated customisable QR code generation for short links, featuring logo uploads, colour theming, and size adjustments, leading to greater adoption by marketing teams and increased campaign-tracking usage.

Cloud Engineer – MSD
Jul 2023 — Dec 2023

- Architected and implemented cloud automations mapping **95%+ of cloud resource cost** to company applications, generating targeted cost-optimisation reports and automated alerts for untagged resources, promoting enterprise-wide accountability.
- Developed AWS QuickSight dashboards through SQL data manipulation, uncovering **US $1.3M in monthly savings opportunities**, enhancing cost visibility and driving targeted optimisation initiatives.

Education

Bachelor of Science in Computational Biology (Graduated with Distinction)
May 2023

National University of Singapore

Side Projects

heyitskeith
Personal portfolio website built to showcase professional experience, projects, and technical expertise.

- Website: https://keithchewzk.github.io/heyitskeith
- GitHub: https://github.com/keithchewzk/heyitskeith
- Tech Stack: React, TypeScript, Tailwind CSS

ReportScribe
AI-powered web application that automates the generation of personalised student reports for teachers, streamlining workflow and significantly reducing manual effort.

- Website: https://reportscribe.up.railway.app
- GitHub: https://github.com/keithchewzk/reportscribe
- Tech Stack: Python, FastAPI, React, TypeScript, Docker, Google Gemini API

FareShare
Collaborative web application for tracking and splitting shared car usage and expenses. Enables users to log trips, calculate costs automatically, and settle balances transparently within groups.

- Website: https://fareshare.up.railway.app
- GitHub: https://github.com/keithchewzk/fareshare
- Tech Stack: Python, FastAPI, React, TypeScript, Docker, PostgreSQL, Google Maps API

Certifications

Kubernetes and Cloud Native Associate (KCNA)	May 2025
AWS Cloud Quest: Solutions Architect	Jan 2025
AWS Certified Cloud Practitioner	Sep 2023