Zhengxuan Liu

Backend Engineer | Go · Rust · TypeScript | APIs, auth, streaming

Shenzhen, China · UTC+8 · yinaaaliu@outlook.com · +86 18002511485

Summary

Backend engineer working across Go, Rust and TypeScript. I have built OpenAI-compatible gateway services (auth, quota, usage metering, streaming), agent runtime HTTP services (sandbox lifecycle, signed requests, compatible sandbox endpoints), and a Fastify control plane with WebSocket sessions and a deterministic capacity/cleanup state machine. Comfortable owning API contracts, persistence and production debugging. Services I built serve 10k+ B2B tenants at ~80M requests/day with P95 180ms.

Skills

- **Backend**: Go (Gin/GORM), Rust, Node.js / TypeScript, Fastify, REST, OpenAPI / JSON Schema, SSE / WebSocket, JWT / auth, SQLite / PostgreSQL / Redis
- **Infrastructure**: Docker / Compose, Linux, Shell, Nginx, GitHub Actions, CI/CD, APM / health checks, Reverse tunnel (frp), Firecracker / microVM, Capacity planning
- **Architecture**: SDK abstraction, Module boundaries, Unidirectional data flow, Plugin/extension points, API contracts, State management

Experience

Xiaoe Inc. — SaaS for online education — Software Engineer (Client & Platform)

2023.10 – Present

- Contributed to internal platform services in Go and Rust: API gateway and LLM proxy, agent-runtime HTTP wrappers, service governance contracts, and Docker-based delivery for 10+ components.
- Served 10k+ B2B tenants with 10+ Go/Rust services handling ~80M requests/day (peak 4k QPS), keeping core API P95 at 180ms.
- Evaluated and integrated multiple LLM providers behind an OpenAI-compatible gateway; implemented quota/usage metering, streaming and failover for internal agent workloads.
- Built customer-service and data-analysis agents for 10k+ B2B tenants — intent routing, tool calling, human handoff and evaluation — automating ~70% of routine tickets, with first response under 5s and 92% answer accuracy on the internal eval set.

PacPort (Japan) — remote SaaS (part-time) — Full-Stack Engineer (iOS + Backend)

2023.10 – 2024.10

- Built iOS features and Go/Node APIs for a Japan-based SaaS product, owning features end to end from API contract to App Store release.
- Added backend endpoints, auth and third-party integrations; kept API contracts documented and covered by smoke tests.

Projects

OpenAI-compatible LLM gateway & API service · 2025

- Go service with API-key auth, quota and usage metering, admin APIs, SQLite WAL persistence, CORS and streaming passthrough to multiple upstream models.

- Local OpenAI-compatible proxy packaged as a launchd service, with request routing, failover and streaming verification scripts.

AI Agent runtime & sandbox platform · 2025 – 2026

- Built Rust services (SDK, CLI and HTTP API) for sandbox lifecycle, API keys, image pre-cache and compatible sandbox endpoints, with HMAC-signed requests and shared typed models.
- Containerized 10+ services with Docker/Compose and contract-first OpenAPI + JSON Schema governance; added SSE for streaming status and health/readiness probes.

Coding-agent session service · 2026

- Node/TypeScript zero-dependency service that wraps local coding agents: streaming deltas, tool-call cards, a permission-approval loop, slash commands and multi-session isolation by working directory.
- Session persistence and reconnect replay, per-turn watchdogs, and browser–server–browser topology for remote access.

Education

South China Agricultural University — B.Eng. in Computer Science · 2020.09 – 2024.06 · GPA 3.8/5 · Top 10%

Awards

- ByteDance iOS Advanced Camp — team ranked #1, "Best of the Best" award (2022)
- Top 10 student startup team, South China Agricultural University (2023)