Alisher Sharipov

AI / Backend Engineer (Python, Go) | LLM Integrations, REST APIs, ERP and 1C

Astana, Kazakhstan (open to remote) | alishar2005@gmail.com | github.com/hikkian | linkedin.com/in/hikkian

Summary

Backend engineer in Python and Go. In production: the HR and payroll module of an ERP with integrations to 1C, and a retrieval-augmented generation (RAG) assistant. Open source: an OpenAI-compatible LLM gateway and an evaluation harness for AI coding agents. Also takes frontend tasks in React and TypeScript.

Experience

ABS LAB Astana, Kazakhstan

Full Stack Developer *Feb 2026 – Present*

- **Lead developer of the HR and payroll module** of an ERP for a road-construction group (team of two); accepted by the customer and in daily use since Sep 2026. **Python**, **Frappe** (Python web framework), **MariaDB**, **SQL**.
- Payroll recalculation and Excel export match the customer's accounting system (**1C**) exactly, to the cent; built a connector to 1C:ZUP, the payroll system, over **OData** and synchronised data from seven 1C databases of subsidiary companies.
- Designed the **OpenAPI** contract for the mobile app; 95 **pytest** test files, including an end-to-end payroll test against 1C.
- Shipped a **RAG** assistant to production: **embeddings** with **Qdrant** vector search and the **Gemini API**, answering recurring product questions in the website chat; cut repetitive pre-sales workload by **15%**.
- Built the pipeline from web forms and the AI chat into **Bitrix24** CRM and email; deployed a multilingual **React** site on **Strapi** and **PostgreSQL** with **Docker** and **Nginx**, with **Microsoft 365 OAuth 2.0** sign-in.

Projects

Shura: local LLM platform | Python, C++, TypeScript Open-source, self-hosted

- **OpenAI-compatible LLM gateway in Python:** checks free memory before loading a model, loads models on demand, saves and restores long chat sessions (100k tokens in **0.5 s** instead of 156 s) and recovers from crashes; **580+ tests**, **GitHub Actions** CI on three OSes.
- **Evaluation harness:** 111 coding tasks checked by hidden tests, with run-to-run noise measured so changes ship only on clear results; it caught a gateway bug that timed out every task of one model before release.
- Extended **llama.cpp** (C++) to run a 35B model on one 16 GB GPU at **75 tok/s**; **MCP** tool integration and a Linux sandbox for AI coding agents.

Open Go Academy | Go, React, TypeScript, Docker Open-source project, MIT

- Learning platform with a **Go** backend that compiles and runs user code in **isolated Docker containers** and grades it with tests; React frontend; optional LLM tutor.

Barid | Python, JavaScript Open-source project, MIT

- Task board for running several AI coding agents in parallel: queue with dependencies, locks on shared resources, human review gates, web dashboard.

Technical Skills

Backend: Python, Go, FastAPI, Frappe, REST API design, OpenAPI, SQL, PostgreSQL, MariaDB, OAuth 2.0
AI / LLM: retrieval-augmented generation (RAG), embeddings, vector databases (Qdrant), prompt engineering, LLM agents, tool calling, MCP, OpenAI-compatible APIs, Gemini API, llama.cpp
Integrations and infrastructure: 1C (OData), Bitrix24, Docker, Docker Compose, Nginx, Linux, Bash, Git, GitHub Actions (CI/CD)
Testing and frontend: pytest, Playwright, React, TypeScript
Languages spoken: Kazakh and Russian (native); English (B2 / upper-intermediate)

Education

Astana IT University | BSc in Computer Science 2023 – 2026

Bronze, International Eurasian Olympiad in Informatics; Bronze, Kazakhstan National Olympiad in Informatics