

Jennifer Umoke

Software Engineer, Backend & AI Systems, First Class LLB

iamjenniferumoke@gmail.com | github.com/jenniferumoke | linkedin.com/in/jenniferumoke | jcumoke.com/blog

Profile

Software engineer with 3+ years of experience building backend systems and agents. I obtained a First Class Law degree in 2024 and I am currently studying (part-time online) for an MSC in Computer Science with AI at The University of York. I ran an AI startup for 2 years and I currently maintain an open source SDK that provides a unified interface to both OpenAI Responses API and Anthropics Messages API. During my startup, I built and shipped applications that securely handle user data, provide secure authentication via Auth0 and ran on Azure Kubernetes Service, backed by Cloudflare.

Skills

- **Backend & data:** Python, TypeScript, Node.js, REST APIs, SDK design, data pipelines, Pandas, NumPy
- **Secure systems:** Secure storage of user data, Authentication with Auth0, Access management with azure entra id, credential manangement with Azure keyvault
- **Infrastructure & operations:** Azure Kubernetes Service, Docker, Istio, Azure ML, Azure Monitor, GitHub Actions CI/CD, incident response, automated testing
- **AI systems:** Building coding agents, tool calling, context management, OpenAI and Anthropic APIs, using huggingface and vllm

Work Experience

AI Engineer and Primary Maintainer, Tessaract

Jun 2026 – Present

Tessaract: A unified LLM API for building agents

Link: <https://jenniferumoke.github.io/tessaract/>

- Built an SDK that gives agents one interface to OpenAI and Anthropic models, with full support for reasoning, tool calling, caching and streaming.
- Designed an extensible adapter and provider architecture, so new providers can be added without changing core code.
- Own the project end to end, from design to release: published an alpha on PyPI with CI that tests across 4 Python versions and automates releases and documentation.

Sabbatical: Independent Engineering

Jan 2026 – Jun 2026

jvdb: Vector database built from scratch in Python

Link: <https://github.com/jenniferumoke/jvdb>

- Built a vector store for embeddings with semantic search and disk persistence, keeping stored records and their index consistent through save and load.
- Benchmarked two storage backends and documented the trade-offs (NumPy roughly 400x faster) in a published technical write-up.

tiny-code: CLI coding agent in Python and TypeScript

Link: <https://github.com/jenniferumoke/tiny-code>

- Built a coding agent using OpenAI Responses API in both python and typescript.
- Designed the tool calling loop to catch errors and feed them back to the model, so the agent is able to handle errors in its environment.

AI Engineer & Founder, Brilliant AI

Jun 2023 – Jan 2026

BrilliantCode: Open-source coding agent harness

Link: <https://github.com/jenniferumoke/brilliantcode>

- Built a full-stack desktop app where an AI agent writes code, runs terminals and tests its work in a built-in browser.
- Designed approval gates and least-privilege controls for autonomous agents: several permission modes and filters that block risky actions until a user explicitly approves them, keeping human review and accountability at the centre of the automation.
- Built a multi-provider runtime (OpenAI, Anthropic, local models) with ChatGPT-subscription sign-in as an alternative to API keys.
- Implemented Trajectory Compaction: older turns and large tool outputs are summarised before reaching the main model, while every user message is kept intact.
- Set up failure-tracking dashboards in Azure Monitor and handled incidents in production, shipping fixes that stopped issues recurring and cut failure rates.
- Built a release pipeline that automates builds, tests, code signing and macOS notarisation on three platforms.

LlamaCloud: OpenAI-compatible inference platform for open-source AI models

Link: <https://youtu.be/pBSWesRW3m8>

- Built and operated a backend platform serving Llama 3, Mistral Nemo and Flux Schnell through an OpenAI-compatible API.
- Deployed and served Wan Video Model using diffusers running on Model GPU sandboxes, also built a client SDK for users to securely access the video API.
- Deployed workloads across Azure ML and Scaleway GPU infrastructure, weighing cost, latency and availability.
- Owned production operations end to end: tracked request failures on Azure Monitor dashboards and shipped fixes that stopped incidents recurring and cut failure rates.
- Published a Python SDK that handles authentication, routing and error handling for developers.

Internships

Legal Intern, DHL

Sep 2022

- Supported the IT legal services team on GDPR compliance for technology procurement contracts, working on Standard Contractual Clauses for EU-US data transfers and data processing agreements with technology suppliers.
- Drafted and reviewed data protection clauses in technology contracts, covering processor obligations, security measures and audit rights.

Vacation Scheme, Mayer Brown International LLP

Jun 2023

- Rotated through the Intellectual Property and Tax practice groups at an international law firm.
- Contributed to IP due diligence on technology matters, scoping and reviewing records to identify legal risk.
- Offered a training contract on the strength of my performance on the scheme.

Data Analytics Intern, Kesar & Co Solicitors

Aug 2022

- Worked with the practice manager to turn open legal-operations questions into a clear analysis plan, then reported findings on immigration, civil and criminal case outcomes.
- Built a data pipeline for 56 clients' survey responses: digitised the forms, cleaned and standardised them with Pandas, ran sentiment analysis with Hugging Face classifiers, and produced Power BI dashboards for staff.

Education

MSc Computer Science with Artificial Intelligence, University of York (online, part-time)

2026 – 2028

LLB (Hons) Law, First Class, University of Greenwich

2021 – 2024

Nanodegree, Data Analysis, Udacity

2022

Publications

- Building a Vector Database in Python, jcumoke.com June 2026
<https://jcumoke.com/blog/building-a-vector-database-in-python>
- How to Build a Coding Agent, jcumoke.com April 2026
<https://jcumoke.com/blog/how-to-build-a-coding-agent>
- How to Build an AI Agent with TypeScript and the OpenAI API, jcumoke.com February 2026
<https://jcumoke.com/blog/how-to-build-an-ai-agent-typescript-openai-api>