Miriam Ho Ai Ning

Senior Software Engineer · Production AI/LLM Systems · AWS
miriamhowork@gmail.com · miriamho.dev · Selangor, Malaysia
linkedin.com/in/miriam-ho-ai-ning-67b3711b1 · github.com/myrrym

PROFESSIONAL SUMMARY

Designed and built property-ai at IQI as its sole architect, a production AI microservice (Python/FastAPI) serving 7M+ property listings across 111 countries, with OpenAI API query understanding backed by a deterministic regex fallback that keeps search fully functional through LLM provider outages. Winner of the coder track at the IQI AI Hackathon 2026. I ship with the Claude API daily (code review automation, RAG chatbots with no managed AI infrastructure) and bring 6+ years across production systems on Laravel, React, PostgreSQL, and AWS. Multilingual: English, Mandarin, Malay, Cantonese.

TECHNICAL SKILLS

AI / LLM Systems: OpenAI API and Claude API in production, RAG architecture, hand-rolled vector similarity (no vector DB, no managed AI infra), deterministic fallback and rules-engine design, grounded generation (LLM phrases computed figures, never prices), prompt and context design, LLM cost controls, agentic coding workflows (Claude Code daily), ElevenLabs

Languages & Frameworks: PHP (Laravel), Python (FastAPI), JavaScript/ES6+, TypeScript, React, Next.js, Flutter, Tailwind CSS, Laravel Filament, Leaflet.js, Java (Joget)

Data & Infrastructure: PostgreSQL, MySQL, MariaDB, Supabase, AWS (EC2, S3, Lambda, CDK), Linux (Ubuntu/CentOS), Git, CI/CD, RESTful APIs, geospatial routing (self-hosted Valhalla), geocoding, XML data pipelines

PROFESSIONAL EXPERIENCE

Full-Stack Developer (AI Systems) | IQI Holdings Sdn Bhd Dec 2025 - Oct 2026

Sole architect of property-ai, a cross-platform AI microservice serving 7M+ listings across 111 countries on iqiglobal.com and juwai.asia.

- Architected property-ai end-to-end (Python/FastAPI): semantic property search, lead-capture RAG chatbot, geospatial routing, and multi-domain API integrations across 2 production domains.
- Built OpenAI API query understanding for natural-language property search, with a deterministic regex fallback (bedrooms, property type, price) that keeps core search fully functional during LLM provider outages and serves covered queries at zero LLM cost.
- Engineered a RAG chatbot over the platform's service-content corpus using hand-rolled vector similarity - no vector database, no managed AI infrastructure - meeting retrieval quality targets within existing cost constraints.
- Owned the team's Claude API code review service: automated diff analysis with risk-tiered review lanes, rolled out in waves against a measured baseline, with a per-token cost model built before enforcement.
- Designed a content compliance service: versioned rulepacks and a deterministic, Chinese-aware evaluation engine with a staged rollout.
- Owned the isochrone commute-time search feature end-to-end on self-hosted Valhalla (AWS), a first-of-kind capability for the platform.
- Built an AI podcast generator (Claude API + ElevenLabs) that turns property briefing PDFs into two-host audio episodes.
- Won the coder track at the IQI AI Hackathon 2026 with a property-price transparency tool built on public transaction data.

Freelance Full-Stack Developer (selected engagements) | Independent Sep 2021 - Nov 2025

36+ projects delivered between 2021 and 2025; clients included major Malaysian telcos.

- Architected and deployed real-time streaming platforms for major Malaysian telcos using Laravel, PostgreSQL, MariaDB, and GitLab on Linux, including custom analytics dashboards, CI/CD pipelines, and DevOps practices supporting 200+ concurrent viewers.
- Delivered national-level HR management and permit & licensing platforms using Joget, Java, and MySQL.
- Managed full project lifecycles independently: scoping, contracts, delivery, and client relationships.

Associate Software Developer | Mokxa Technologies LLC (US, remote) May 2024 - Jun 2025

Remote role; collaborated with global IT services and systems integration partners on enterprise-scale platforms.

- Led development of a national competency standard platform for a federal agency (Java + Joget) in partnership with two global IT services firms, owning Event and Course Management modules.
- Built a permit and licensing management system for deployment across US and Canadian jurisdictions, handling complex multi-region regulatory workflows.
- Redesigned employee portal UI/UX, achieving an 80% reduction in maintenance time through optimised CMS workflows and automated content updates.

Product Specialist | Joget, Inc. (part-time) Oct 2022 - Apr 2024

Part-time contractor role during final year of degree; contributed to official product content published on Joget's company channels.

- Produced official Joget DX tutorial content (60,000+ views on Joget's YouTube channel) and knowledge-base documentation used as primary onboarding material for the global user base; developed and debugged Joget plugins in Java.

Junior Full-Stack Developer | Cake Experiential Communications Jul 2020 - Feb 2021

- Supported disaster recovery (DR) for a Malaysian bank, covering web application failover and database mirroring across environments.
- Junior developer on the real-time streaming platforms for three Malaysian telcos (Laravel, PostgreSQL, MariaDB, GitLab on Linux), later carried forward as freelance engagements.

PROJECTS

sanctuary - AI journaling app (2026): shipped solo to 17 closed-beta users on Google Play (Flutter, Supabase, Claude API) with tiered pricing and PDPA/GDPR compliance on a ~USD 70/month budget; built to learn the mobile stack and store-release pipeline end to end.

token counter - live event installation (2026, freelance via an events agency for a German luxury automaker): 8 ESP32 stations with IR sensing, SD-card persistence and wired Ethernet, reporting to a FastAPI hub on a Raspberry Pi with a live WebSocket dashboard. Ran unattended for 4 days with crash-safe counting. Built the firmware, the circuits and the dashboard.

CERTIFICATIONS

AWS Certified Cloud Practitioner 2026

EDUCATION

B.Sc. (Hons) Software Engineering · Universiti Tunku Abdul Rahman (UTAR) 2023