

Nihal Yusof

Senior Solutions Engineer · AI Solutions Architecture & Delivery

Singapore · hello@nihalyusof.com · nihalyusof.com · linkedin.com/in/nihal-yusof

SUMMARY

Solutions engineer at OptioAI since Nov 2024, promoted to Senior in Jul 2026. Delivered 8 production AI systems (chatbots, agentic assistants and analytics platforms) for enterprise and government clients across Southeast Asia and the GCC, owning them from pre-sales scoping to handover with teams of up to 15–20 engineers.

CORE SKILLS

- **AI:** Agentic AI, RAG, LLM-as-judge evaluation, Letta, OpenRouter (GPT-5.x, Claude, Gemini), self-hosted Llama 3.1, pgvector, Qdrant, ML solvers and schedulers
- **Engineering:** Python (FastAPI), TypeScript (React, Next.js, NestJS), Flutter, Temporal, Kafka, Celery, Redis, PostgreSQL
- **Cloud & DevOps:** Google Cloud (GKE, AlloyDB), Kubernetes, Terraform, Docker, GitHub Actions, Linux
- **Integrations:** AMPECO, Microsoft Graph, Gmail API, Google Health Connect
- **Leadership:** Solution architecture, pre-sales scoping, teams of up to 15–20 engineers, graduate mentoring

EXPERIENCE

OptioAI Technologies — Singapore

Senior Solutions Engineer

Jul 2026 – Present

- Own architecture and delivery for AI engagements in Singapore, Malaysia, the UAE, Saudi Arabia and Qatar, leading teams of up to 15–20 engineers from pre-sales scoping to handover.
- **Newton**, agentic AI platform (Letta, OpenRouter, FastAPI, React): led production readiness; load tests proved one 8-vCPU node serves 60 concurrent conversations, and prompt and Redis caching cut LLM costs by over 60%.
- **SkillBit Module Creator:** AI pipeline that researches, writes and fact-checks security-training courses, with LLM-as-judge review (FastAPI, React, OpenRouter).
- **EV fleet charging optimiser** for a Singapore charge point operator, built on AMPECO (NestJS, Next.js).
- Advise companies in India and the USA on AI strategy; built the graduate intern-to-engineer pathway.

Solutions Engineer

Nov 2024 – Jul 2026

- Delivered 5 production chatbots and agentic systems and 3 analytics systems (sports analytics for Aspire Zone, 2 governance monitors) for clients across Southeast Asia and the GCC.
- **Rally**, Aspire Zone Foundation's AI assistant (Qatar): HR, learning and email AI with Gmail/Outlook sync on Celery workers, pgvector RAG on AlloyDB and self-hosted Llama 3.1 on GKE GPUs. Recognised with Google Cloud's Transformative AI Application Excellence Award (Doha, 2025).
- **Newton:** agentic AI platform; designed the backend and co-built the front end and CI/CD, then extended it with RAG and analytics tools for a UAE ministry.
- **Euler:** Temporal-based workflow automation engine with Microsoft Graph and Kafka triggers.
- Built ML scheduling solvers, EV-charging integrations and brand-personality social media bots.

INDEPENDENT PROJECTS

- **Home server:** Unraid server (51 containers, 3 VMs) hosting AI projects and nihalyusof.com since 2023.
- **Kairos:** health-insights app that syncs Fitbit data through Google Health Connect.
- **KopiYo** (in development): Flutter app with LLM-summarised daily briefings on news, markets, weather and traffic.

LEADERSHIP

Singapore Police Force — Inspector (Full-time National Service)

Aug 2018 – Aug 2020

- Commissioned through the Officer Cadet Course (70th intake); Division Security Officer at the Police Coast Guard.
- Selected for the Guard of Honour contingent at the National Day Parade 2019.

EDUCATION

BSc Computer Science, First Class Honours — University of London (SIM)

Feb 2022 – May 2024

Diploma in Computer Engineering — Singapore Polytechnic

2015 – 2018

ADDITIONAL

- **Part-time sales**, Malabar Gold & Diamonds (Sep 2022 – Jun 2024), alongside full-time degree studies.
- **Languages:** English and Tamil (native), Malayalam (professional working)