

# Irfan Naqieb Baharuddin

+60 12 503 0332 / + 82 10 2707 0323 | [linkedin.com/in/irfannaqieb](https://www.linkedin.com/in/irfannaqieb) | [irfannaqieb@gmail.com](mailto:irfannaqieb@gmail.com) | [irfannaqieb.pages.dev](https://irfannaqieb.pages.dev)

## EXPERIENCE

### Enygma AI

Full Stack AI Engineer

Kuala Lumpur, Malaysia (Remote)

January 2025 – November 2025

- Designed and optimized modular REST APIs for an AI-powered enterprise asset-management platform, improving query latency through SQL optimization, enabling scalable AI-powered features.
- Integrated automated testing (unit & integration) into CI/CD pipelines using GitHub Actions, improving release stability, deployment confidence, and test coverage for AI-backed microservices.
- Performed large-scale data migration and production deployment on AWS, using Terraform for infrastructure provisioning and GitHub Actions for CI/CD automation, ensuring zero downtime and 99.9% system reliability.
- Developed an end-to-end ML pipeline for freight-train anomaly detection, combining YOLO-Seg for segmentation, DINOv3 Vision Transformer (LVM) for feature extraction, and FAISS with k-Nearest Neighbors for similarity-based anomaly retrieval.
- Implemented heatmap visualization with OpenCV and optimized inference using PyTorch quantization and FAISS, achieving real-time anomaly detection performance suitable for edge deployment.

## PROJECTS

### Image Preference / GenAI Benchmarking Platform | [link](#)

Yonsei University

Nuxt.js, Tailwind CSS, Python (FastAPI), OpenAI, Gemini, Hugging Face

October 2025

- Designed and automated GenAI benchmarking and verification framework to compare state-of-the-art LLM models' outputs, integrating PyTorch-based evaluation scripts for latency, diversity, and quality metrics
- Developed a FastAPI backend to receive votes and expose REST endpoints for the frontend. Deployed via FastAPI with CI/CD automation

### Waybill | [link](#)

Remote

Python, FastAPI, OpenAI, Chroma DB, Docker, LangChain, LlamaIndex

October 2025

- Developed AI agent backend using FastAPI + LangChain + LlamaIndex, enabling Retrieval-Augmented Generation (RAG) workflows grounded in company documentation.
- Deployed as Dockerized services on cloud and integrated ChromaDB + RAG to ground OpenAI endpoints in company documentation for higher accuracy.
- Shipped secure, production-ready APIs with context engineering and PII guardrails (spaCy) to meet data-protection requirements.

### 360° Panoramic Visualization Tool | [link](#)

Enygma AI

Nuxt.js, TypeScript

August 2025 - September 2025

- Developed an internal tool using Nuxt (Vue + Typescript) to calibrate orientation and link 2,000+ 360° images of oil rig interiors, producing a Google Street View-like panoramic navigation system that enabled seamless exploration of complex industrial environments.

## EDUCATION

### Yonsei University – College of Computing

Seoul, South Korea

BS in Computer Science

January 2022 - December 2025

- Relevant Courses: Data Structures, Deep Learning, Computer Networks, Reinforcement Learning, Computer Architecture
- Awards & Achievements: CS 2024-1 Merit Scholarship; MARA (Majlis Amanah Rakyat) Scholarship Recipient

## SKILLS

**Frontend:** React, Vue, Next.js, Nuxt.js, TypeScript, Tailwind CSS, Shaden UI, JavaScript

**Backend:** Python (FastAPI, Django), Node.js, SQL, Redis, PostgreSQL

**Cloud & DevOps:** AWS (EC2, S3), Terraform, CI/CD (GitHub Actions), Docker

**AI/GenAI:** LLMs (Open AI, Anthropic, Hugging Face), Vector DBs (Chroma), Context Engineering, PyTorch, LangChain

**Languages:** English (Fluent), Malay (Native), Korean (Intermediate)