

# PETER HUYNH

## AI Engineer

Melbourne VIC | Student visa (subclass 500), expires 30 Aug 2027  
giatinhuynh2612@gmail.com | (+61) 478 719 799  
linkedin.com/in/peterhuynh2612 | github.com/giatinhuynh | shinya.live

### PROFESSIONAL SUMMARY

AI engineer who puts agents into production rather than demos. Built a **79-tool agentic system** (LangGraph, AWS Bedrock, pgvector) with governed autonomy, gating every mutating action behind approval and confidence-based escalation so agents could act safely near regulated infrastructure. Founding engineer and co-technical lead at Orbitify, authoring ~65% of a 10-repository platform spanning RAG, GPU inference, geospatial AI and the cloud infrastructure beneath it. Melbourne Campus Winner at RMIT CyberHack 2025.

### TECHNICAL SKILLS

**AI and GenAI:** Agentic AI (LangGraph, LangChain), AWS Bedrock (Claude, Cohere), Google Gemini, RAG with pgvector, Model Context Protocol, LLM observability (Langfuse), adversarial red teaming, YOLOv8, Mask R-CNN, SAM3, BERT

**Languages:** Java 17, TypeScript, Python, SQL, JavaScript, Bash, HCL

**Backend:** Spring Boot 3, FastAPI, Express, microservices, CQRS, REST, Server-Sent Events, PostgreSQL, PostGIS, pgvector, Redis, Firestore, Flyway, JPA and Hibernate

**Cloud and DevOps:** AWS (ECS Fargate, EC2 GPU, RDS, ElastiCache, S3, CloudFront, Bedrock, SageMaker), GCP (Cloud Functions v2, Firebase), OpenTofu and Terraform, Docker, GitHub Actions with OIDC

**Observability:** OpenTelemetry, Grafana Cloud, PostHog, CloudWatch, GA4, SLOs and alerting

**Security:** JWT and RBAC, multi-tenant isolation, IAM least privilege, Workload Identity Federation, secrets management, API and CORS hardening

**Frontend:** React 19, Next.js 16, Vite, Tailwind CSS, Three.js

**Geospatial:** GeoServer, NodeODM and ClusterODM, Mapbox, COLMAP, orthomosaic pipelines

**Spoken:** Vietnamese (native), English (professional), Japanese and German (elementary)

### PROFESSIONAL EXPERIENCE

**Orbitify**  
**Founding Software Engineer and Co-Technical Lead** **Mar 2025 to Present**  
*Full time, remote. The agentic operating system for critical infrastructure: O-EYE detects, O-ARC analyses, O-OPS holds the record, O-AGENT decides and acts. Backed by NVIDIA Inception, AWS and Microsoft for Startups.*

- Made autonomous agents safe to run against regulated infrastructure by designing a governed-autonomy model that gates all **27 mutating actions** of a 79-tool agent behind approval workflows and confidence-based escalation (LangGraph, AWS Bedrock, pgvector).
- Collapsed detection-to-dispatch from a manual handoff into **one automated flow** by building the anomaly to incident to work-order loop, matching assignees against technician credentials and raising inventory and procurement requests against the asset.
- Delivered the platform as founding engineer by authoring **~65% of all commits and pull requests** across 10 repositories: a Spring Boot core API, a React operator console and a dozen services on AWS ECS with GPU autoscaling.
- Eliminated long-lived credential risk across the platform by moving **100% of CI and human access** onto OIDC and task roles, with tenant-scoped RBAC, least-privilege IAM and managed secrets, shipped with zero downtime.
- Made the cloud estate reproducible and auditable by taking **100% of it into OpenTofu** as sole author, with plan-on-pull-request, nightly drift detection and a staging environment from the same modules, holding zero drift.
- Took long-running geospatial jobs from unreliable to survivable at scale by instrumenting the pipeline, then shipping split-and-merge processing and ledger reconciliation with an **8x GPU capacity increase** for concurrency.
- Standardised how the team builds and responds by writing the **engineering handbook**: architecture decision records, C4 diagrams, runbooks and the incident process.

## RAISE Hub, RMIT University

### Software Engineer

Feb 2026 to Jun 2026

Full time, hybrid. One semester RMIT university placement, carried alongside the Orbitify role.

- Delivered Internbot to **2,000+ students** as tech lead of a 5 person team, hitting 100% completion across 5 sprints and 176 story points.
- Cut policy lookup from a manual search to a cited answer by shipping **interbotRAG single handedly**: an AI microservice on Cloud Functions v2 (Gemini 2.0 Flash, Supabase pgvector) serving semantic RAG with direct citations.
- Reached **87% test coverage** behind an 80% CI/CD gate by implementing CQRS and 4 layer clean architecture with Firestore unit of work transactions.
- Removed **100% of long-lived service account keys** by automating zero-downtime GCP deploys through Terraform and GitHub Actions with OIDC Workload Identity Federation.
- Built a reusable full-stack boilerplate adopted across RMIT's capstone course, supporting **400+ students each semester**.

## Project Pluto

### Software Engineer, Contract

Nov 2024 to Aug 2025

Full time contract, hybrid, Ho Chi Minh City. Held concurrently with the founding Orbitify role from Mar 2025.

- Reduced manual workload by **40%** by leading GreenBook and automating project setup, deployment and GA4/PostHog analytics on Next.js, Supabase/Prisma, Clerk and Vercel.
- Delivered 5 production web products on that stack reaching **80,000+ users**: KOTO Media Center 12,000 on Payload CMS, Newing 10,000, SIC 2025 1,000, and Alumni Impact Showcase 2025 at 50,000 views and 24,000 returning.

## PROJECTS

---

### RMIT GenAI and Cyber Security Hackathon | RMIT-CyberHack2025

#### Blue Team Lead and Red Team Contributor. Melbourne Campus Winner.

Oct 2025

- Built an ensemble of DistilBERT, RoBERTa and DeBERTa with a hybrid TF-IDF architecture to detect unsafe prompts across a 5,000 sample dataset.
- Red teamed Azure OpenAI safety filters, **bypassing 2 of 5 advanced safety prompts** and documenting the vulnerabilities found.

### Colon Histopathology AI Classifier | GitHub · Kaggle

#### Machine learning engineer. PyTorch, scikit-learn.

Apr 2026 to Jun 2026

- Exceeded diagnostic targets at **0.91 Macro F1** on cancer detection and 0.75 on 4-class cell typing against genuinely unseen patients, evaluating 12 models across classical ML and custom convolutional networks in PyTorch over strict patient-level splits that prevent staining and scanner leakage.

### OurList | anime-duo-delight.vercel.app

#### Solo. React, TypeScript, Supabase, PWA.

Nov 2025 to Present

- Built a unified tracker for anime, film, books, music and games with live MyAnimeList, TMDB and Spotify sync and explainable recommendations.
- Built a **Model Context Protocol server** with OAuth and BYOK: one tool layer serving the in app chat, Claude, Cursor and ChatGPT Apps through the same audited surface.

## AWARDS AND LEADERSHIP

---

**1st Place (Individual)**, SentinelOne ThreatOps Challenge | Sep 2026

**Student Fellow W26**, Startmate

**2nd of 50+ teams**, CSIT x SWITCH Hackathon | Oct 2024

**1 of 10 students selected across RMIT's global campuses**, Global Citizen Leadership Program (McCall MacBain) | Nov 2024

**Technology Department Executive**, RMIT Vietnam FinTech Club | Jul 2023 to Feb 2024

## EDUCATION

---

**Bachelor of Information Technology, RMIT University** | Oct 2022 - Expected graduation: 2027

GPA 3.8 / 4.0 | Academic Merit Scholarship, Dean's List 2024 and 2025