# Hoang Thai Huy

+84 902 650 802 | hoangthaihuy.work@gmail.com | LinkedIn | GitHub

## WORK EXPERIENCE

**DevOps Engineer | PMAX - Total Marketing Partner** **Sep 2026 – Present**

- Provisioned and operated self-hosted GitLab Runners (Docker executor) on VPS and AWS Lightsail, giving pipelines dedicated build capacity with Docker layer caching and no shared-runner minute limits.
- Built and maintained GitLab CI/CD pipelines (build Docker image, image pushed to AWS ECR, deploy) for staging services across PMAX's functional platforms (CreativeOS, MediaOS, CompanyOS, FinOS), replacing manual builds with automatic deploys on merge and cutting release from **∼20** to **∼5 minutes**.
- Built a Google Chat bot for the company workspace using Google Cloud APIs, with role-based access control limiting which users can write data to the AWS PostgreSQL database and an audit log in Google Sheets recording each interaction (user, action, status); hosted on an AWS Lightsail instance.
- Built Grafana dashboards for Amazon RDS PostgreSQL using CloudWatch metrics and logs (covering CPU, connections, storage, latency and slow queries), validated in a staging environment ahead of production rollout.
- Wrote runbooks and incident notes in English covering deployment, rollback and Runner recovery.

**Analytics Engineer Intern | GiaoHangNhanh (GHN)** **Mar 2026 – Jun 2026**

- Designed BigQuery SQL queries (multi-table joins, CTEs, window functions) for ad-hoc analysis, and automated data exports to Google Sheets using Python, reducing report preparation time from **3–4 hours** to significantly less for non-technical stakeholders.
- Built and maintained daily ETL/ELT pipelines with Apache Airflow and Docker, structuring raw logistics data into OLAP-friendly dimensional tables to support reliable, timely data delivery.
- Built AppSheet applications to manage user access and report visibility, and developed interactive Looker Studio dashboards (with pivot-based summary views) to monitor logistics KPIs and support reporting needs across multiple departments, replacing manual reporting workflows.

**Data Processing Intern | FPT Information System HCMC** **Oct 2025 – Mar 2026**

- Processed and organized backend data for a **HCMC Metro System** and government websites (during provincial merger), writing SQL joins and aggregations across OLTP source tables to prepare clean, query-ready datasets.
- Supported the development of a **Metro Passenger Analytics** system by aggregating passenger volumes by station and 30-minute time intervals from 5:00 AM onward, and building pivot-style reports to analyze passenger patterns across stations and time periods.
- Supported the build of statistical and predictive models for passenger volume estimation, validating outputs against historical ridership data.
- Developed a web application to visualize and validate predictions, giving stakeholders a self-serve view of forecasted vs. actual volumes.

## EDUCATION

**HO CHI MINH CITY OPEN UNIVERSITY (HCMCOU)** Sep 2022 – Jul 2026

Bachelor of Computer Science (Major in Data Science) GPA: 3.29/4.0

Relevant Coursework: DSA, Statistics, Database, Database Management,
Systems Analysis and Design, Cloud Computing, Data Analysis, Data Mining

## ACHIEVEMENTS

**Scholarships:**

- 5 out of 11 semesters from HCMC Open University (70% - 100% of tuition fees each semester)

**Honour:**

- Awarded "Outstanding Student of the Year 2025" by HCMC Open University
- Awarded "Certificate of Merit for Outstanding University Entrance Examination Performance" by Ho Chi Minh City Civil Servants' Trade Union (Công đoàn Viên chức Thành phố Hồ Chí Minh)

## SKILLS

**Programming:** Python, Java, SQL
**CI/CD & GitOps:** GitLab CI/CD, GitHub Actions, GitOps
**Cloud:** AWS (EC2, Lightsail, S3, IAM, VPC, Lambda, RDS), Cloudflare (DNS, TLS, proxy); familiar with GCP, Azure
**Containers:** Docker, Docker Compose
**OS**: Microsoft Windows, Linux administration
**Scripting & Config:** Bash, Shell, YAML, JSON, SQL
**Data:** Apache Airflow, MySQL, SQL Server, PostgreSQL
**Tools:** Jira, Claude Code, OpenAI Codex, AI-assisted tools

## PROJECTS

**SaaS ETL Platform** Mar 2025 – Apr 2025
*Python, Apache Airflow, MySQL, MinIO, Docker, AWS(EC2, VPC)*

- Developed an end-to-end SaaS ETL pipeline using Apache Airflow to automate scheduled data ingestion and loading workflows, storing raw files in MinIO (S3-compatible) and processed (OLTP) data in MySQL.
- Containerized Airflow, MySQL, MinIO, and the Flask app using Docker, and deployed the platform on AWS EC2, validating data flow from ingestion to storage and web application.
- Designed the platform to support user-defined data sources and database structures, allowing users to customize data extraction and database design, receiving a final project grade of **10/10**.

**House Price Prediction in HCMC (Team Leader)** Mar 2025 – May 2025
*Python, Pandas, Google Sheets, Selenium, Scikit-learn, AWS EC2*

- Collected and processed more than **100,000 real estate records** from Nha Tot using Selenium and Python, then cleaned and transformed the data with Pandas and Google Sheets, including feature engineering to improve model performance.
- Built a machine learning model to predict house prices from property features (location, size, amenities), achieving **79%** overall accuracy with F1-scores of 0.84/0.82 for high/low-price classes.
- Deployed the trained model as a web application on AWS EC2 for real-time price prediction and user testing.

**Sports Facility Management System (Team Leader)** Apr 2026 – Jul 2026
*Java, Spring Boot, JUnit 5, Docker, GitHub Actions, Azure (ACR, App Service, SQL Server)*

- Led the team as Project Manager: assigned tasks, managed the codebase and version control, and ran weekly progress meetings to keep the team on track and meet deadlines.
- Designed the relational (OLTP) database schema on Azure SQL Server, defining entities, relationships, joins, and constraints to support booking, scheduling, and customer management.
- Built a CI/CD pipeline with GitHub Actions that runs JUnit tests, builds Docker images for the frontend and backend, pushes them to Azure Container Registry (ACR), and deploys them to two Azure App Service services.
- Analyzed functional specifications to design and prioritize test cases covering core backend modules, and created technical documents covering system architecture, database design, testing, and system functions, contributing to a final project grade of **9.3/10**.

## CERTIFICATIONS

**Big Data Course Certificate** Samsung Innovation Campus | Aug 2024
**EF SET English 74/100 (C2)** EF SET | May 2026
**TOEIC L&R (815/990)** IIG Vietnam | May 2026
**Data Engineer Certificate** DataCamp | Jun 2026