

Dam Hong Phuc

+84 394240555 | damhongphuc.4104@gmail.com | [LinkedIn](#)

Education

FPT University, Ha Noi
B.Tech in Artificial Intelligence

Sep 2022 - Dec 2025
GPA: 8.49/10

Startup Product

Wizzdom

July 2025 – December 2025

Technical Leader & Founding Member

- Architected an AI-driven EdTech platform targeting Math educators. The platform transforms inefficient manual assessment habits into a 10x faster workflow, delivering personalized student feedback and eliminating hours of repetitive grading.
- Spearheaded end-to-end development, rapidly prototyping and pivoting iterations to accelerate product-market fit. Engineered complex AI evaluation pipelines using **Gemini** models and successfully managed the scalable cloud deployment on **AWS**.

Professional Experience

FPT Software

January 2026 – Present

AIops & Platform Engineer

- Project: Enterprise HR Conversational AI Platform**
 - Led the migration of conversational AI directing HR workflows for **Covestro**, a multi-billion dollar German chemical corporation. Engineered autonomous **Multi-Agent** systems using **LangGraph** and **RAG**.
 - Shifted legacy monoliths to event-driven microservices on **AWS** (Lambda, Bedrock, EventBridge, SQS, OpenSearch, S3, RDS, ElasticCache). Resolved bottlenecks using semantic caching strategies to handle high traffic throughput.
 - Implemented rigorous **LLM-as-a-judge** evaluation using Langfuse for factual accuracy. Engineered root-level zero-trust policies and real-time **PII redaction** for compliance.
- Project: Enterprise AI Contract Intelligence System**
 - Upgraded the legacy system to a **Skill-based framework** to parse unstructured contracts. Employed prompt chaining and vector **RAG (Qdrant)** to slash hallucinations and boost extraction accuracy by 35%.
 - Dynamically routed multi-model traffic via **AWS Bedrock** to scale for thousands of documents. Synchronized latency via **SSE streaming**.
 - Executed end-to-end migration from local deployments to an enterprise-grade **AWS** infrastructure to scale heavy batch algorithms horizontally.

HTI Group

April 2025 – December 2025

AI Engineer & LLMops

- Project: Enterprise Agentic Chatbot & LLMops Infrastructure** 
 - Rapidly prototyped **Multi-Agent** solutions with **LangGraph** and **LlamaIndex** following **ReAct Pattern** to unify institutional knowledge. Orchestrated cloud-native microservices using **AWS Lambda**, **ECS Fargate**.
 - Automated **CI/CD pipelines** via IaC (**AWS CDK/TypeScript**) for repeatable deployments. Implemented distributed observability with **Langfuse** to trace agent trajectories and ensure hallucination-free outputs.
 - Designed a **Data Ingestion Pipeline** to normalize multi-format sources (Excel, Docs, MD, PPT, PNG, CSV). Applied **Hierarchical Semantic Chunking** with **Auto Merge** and **LLM Contextualization** for child chunks.
 - Combined diverse retrieval strategies: **Vector Search**, **Full-text Search**, **Graph Search**, and **Pandas Retrieval**; integrated tool calling to external services (Discord, Google Calendar).

HG Group

May 2023 – March 2025

AI Engineer

- Project: HG Chatbot - Enterprise Conversational AI**
 - Engineered robust data pipelines to parse and normalize unstructured content across diverse formats (PDF, Docx, PPT, Excel, PNG). Architected an advanced search engine utilizing **Hybrid Retrieval** and **Cross-Encoder reranking** to maximize context relevance.
- Project: HG Render - AI Media Recommendation Engine**
 - Architected a recommendation engine over 1M+ multimodal assets (images, videos, SFX) using dense retrieval (**Qdrant**) and **GPT** for combinatorial optimization to assemble production-ready videos.
 - Engineered async background task processing with **Celery**, **Redis**, and **asyncio** to handle long-running video assembly jobs at scale.
- Project: HIG (HG Image Generation) - Internal AI Asset Creation**
 - Translated SOTA text-to-image diffusion papers into operational internal workflows. Engineered robust visual data pipelines to curate and preprocess large-scale datasets, establishing high-quality training baselines.
 - Fine-tuned **FLUX.1** models with LoRA to highly specific industry design styles. Managed the generation of **custom visual assets**  (*Lofi-cartoon aesthetics*) to drive enterprise media campaigns.

Technical Skills & Certifications

AI & Agentic Paradigms: Python, PyTorch, LangChain, LangGraph, LlamaIndex, FLUX.1 LoRA, vLLM
Distributed Systems: FastAPI, Event-driven Microservices, gRPC, Async Systems, Redis, Model Context Protocol
Databases & Messaging: PostgreSQL, MongoDB, OpenSearch, Qdrant, Kafka, RabbitMQ, AWS SQS, EventBridge
Cloud, DevOps & LLMops: AWS, Docker, Ray, K8s, Langfuse, LangSmith, GitHub Actions, CI/CD
Certifications: [AWS SAA-C03](#), [AWS AIP-C01](#), [Project Manager](#), [ML & DL Specialization](#), [IBM Full Stack](#), [IELTS 7.0](#)