



## Pham Dinh Tien

AI Engineer  
Ha Noi, Vietnam

phamdt203@gmail.com

Mobile: 0393725596

GitHub — LinkedIn

---

## Summary

AI Engineer with 2+ years of experience designing and shipping production-grade AI systems across LLMs, AI Agents, Computer Vision, and MLOps. Hands-on experience building multi-agent platforms, RAG systems, eKYC/face-recognition pipelines, and real-time video analytics, then deploying them at scale on cloud-native infrastructure (GCP, Kubernetes, Terraform). Passionate about turning cutting-edge research into reliable, user-facing products.

---

## Work Experience

- **Vin Dynamics** Ha Noi, Vietnam  
*AI Engineer* *Jun 2026 – Present*
  - Building AI Agent systems for robots, integrating LLMs and multi-agent architectures to enable autonomous perception, reasoning, and task execution.
  - Designing scalable, cloud-native pipelines for training, optimizing, and deploying robotics AI models in production.
- **Savvycom Solution JSC** Ha Noi, Vietnam  
*AI Team Lead / AI Engineer* *Apr 2025 – Jun 2026*
  - Led a small team of AI engineers, owning architecture decisions and mentoring across multiple production projects.
  - **Savvy Agent:** Built an internal enterprise assistant with **LangChain** and **LangGraph** that automates company-policy Q&A (RAG), generates business reports, and integrates directly with timesheet systems.
  - **eKYC Pipeline:** Engineered an end-to-end face-verification system, training and integrating **RetinaFace** (detection), **MiniFASNet** (liveness/anti-spoofing), and **ArcFace** (recognition) for high-security authentication.
  - **Video Analytics:** Maintained and optimized **NVIDIA DeepStream** pipelines for real-time CCTV streaming over RTSP; implemented **YOLO**, **CLIP4Clip**, and **ClipVis** for object detection and semantic video search.
  - **Paygate AI Agent:** Built a multi-agent system for Paygate (a foreign-exchange trading firm), integrated into Mattermost.
  - **CV Scanning Tool:** Built an internal hiring tool for the TA team to source candidate profiles from LinkedIn and TopCV, automating the hiring workflow (email outreach, stage transitions) in an all-in-one platform.
  - **Agent SDLC:** Architected a multi-agent system where each agent owns a Software Development Life Cycle role (Business Analyst, Designer, Developer, Tester) using **GoClaw** as the core agentic framework.
- **1BitLab Technology** Ha Noi, Vietnam  
*AI Engineer* *Sep 2024 – Mar 2025*
  - Researched SOTA AI through papers and blogs and prototyped solutions to keep the team's stack current.
  - Built a RAG system using **LangChain**, Streamlit, and FastAPI; containerized with Docker and deployed on GCP.
  - Applied OCR (incl. Mistral OCR) to extract, process, and clean document data across formats (PDF, DOCX).
  - Configured Nginx reverse-proxy paths and domains; integrated external APIs (OpenAI, Google AI, Groq, DuckDuckGo Search).

---

## Projects

- **eKYC Application** *May 2024 – Jul 2024*  
*AI Engineering, Full-Stack Development, DevOps*
  - Developed a web application with face detection, comparison, search, and liveness detection (anti-spoofing).
  - Used WebSocket for real-time communication between frontend and backend, streaming face-recognition results.
  - Designed and deployed CI/CD pipelines with GitHub Actions to automate testing and deployment of facial-recognition models.
  - **Tech stack:** Python, WebSocket, FastAPI, DeepFace, ArcFace, RetinaFace.

- **LLMOps System for Legal Question Answering**

Apr 2025 – Jun 2025

*MLOps Engineer*

- Designed an MLOps system deployed on Google Kubernetes Engine (GCP).
- Set up CI/CD pipelines for automated build and deployment to the production environment.
- Built a logging namespace with the ELK Stack, set up metrics in Prometheus, visualized in Grafana, and routed alerts via Discord.
- Managed and automated cloud infrastructure with Terraform (deployment, configuration, and management of GCP resources).
- Configured an API gateway and model routing with OpenWebUI and LiteLLM; designed semantic caching with Redis and user-data storage in PostgreSQL.
- **Tech stack:** GCP, Kubernetes, Docker, ELK Stack, Nginx API Gateway, LiteLLM, Redis, PostgreSQL, Prometheus/Grafana, Jenkins.

- **Savvy Agent**

Apr 2025 – Jun 2025

*AI Engineer, MLOps Engineer*

- Designed Savvy Agent (RAG Agent + Timesheet Agent) for Savvycom Solution JSC.
- Implemented AI-agent logic with LangChain and LangGraph in a highly scalable architecture.
- Set up CI/CD pipelines for automated build and deployment of the chatbot application to production.
- Built message, prompt, and request tracing/logging with LangSmith.
- Managed and automated cloud infrastructure with Terraform on GCP.
- Configured an API gateway, semantic caching with Redis, and user-data storage in PostgreSQL.
- Set up the Milvus vector database and integrated a rerank model for improved retrieval performance.
- **Tech stack:** GCP, Docker, LangSmith, Nginx API Gateway, Redis, PostgreSQL, GitLab CI/CD, Terraform, LangChain, LangGraph, Milvus.

## Technical Skills

---

- **Core Languages & Scripting:** Python (Strong), SQL, Bash/Shell, Go, C/C++, JavaScript.

- **AI Engineering & LLM Stack:**

- **RAG & Agents:** LangChain, LangGraph, LangFuse, LangSmith, LangFlow, LightRAG.
- **Agentic Frameworks:** GoClaw, OpenClaw.
- **Model Serving & Inference:** Ollama, LiteLLM, vLLM, LM Studio, Triton Inference Server.
- **Model Optimization:** ONNX, TensorRT.
- **LLM Providers:** OpenAI, Whisper, Grok, DeepSeek, Gemini, GLM.
- **Face Tech Stack:** DeepFace, InsightFace, Silent-Face-Anti-Spoofing, FaceNet, MiniFASNet.

- **Data Engineering:** Airflow, MLflow, Kafka, RabbitMQ, Pandas, Polars, Spark.

- **DevOps:**

- **Cloud:** Google Cloud, AWS.
- **Containerization & Orchestration:** Docker, Kubernetes.
- **CI/CD:** Jenkins, GitLab CI/CD, GitHub Actions.
- **Infrastructure as Code:** Terraform, Ansible.
- **Logging:** Logstash, Elasticsearch, Beats, Vector, Kibana.
- **Monitoring:** Prometheus, Grafana.

- **Databases & Storage:**

- **Relational:** PostgreSQL, MySQL, SQLite.
- **Caching:** Redis.
- **NoSQL:** MongoDB.
- **Vector / Graph DB:** ChromaDB, Milvus, Qdrant, Neo4j.

- **Data Lake:** Google Cloud Storage, MinIO, Amazon S3.
- **Data Warehouse:** Google BigQuery.
- **Backend:** FastAPI, Flask, Django, Node.js.
- **Frontend:** React.js, Next.js.
- **Low-code & Workflow:** n8n automation workflows.
- **AI-Assisted Development:** Cursor, Claude Code, Antigravity, Stitch, Agentic Skills, SuperPowers, Auto Claude.
- **Version Control:** GitHub, GitLab.
- **Languages:** Vietnamese (Native), English (Intermediate), Korean (Beginner).

## Education

---

- **Hanoi University of Industry** Ha Noi, Vietnam  
*B.Eng. in Software Engineering, Major in Artificial Intelligence* 2021 – 2025
  - GPA: 3.03 / 4.0.

## Honours & Awards

---

- **National Student Informatics Olympiad – Open Source Contest** *Dec 2024*
  - Represented Hanoi University of Industry in the open-source competition at the National Student Informatics Olympiad.
  - Won the consolation prize, organized by the Vietnam Informatics Association and VFOSSA.
  - Built a Langflow-based RAG model for the healthcare sector within the SOS Connect application.
- **VFOSSA Open Source Contest at HaUI** *Nov 2024*
  - Achieved first prize in the university-level open-source contest.
  - Built a RAG-based question-and-answer system for legal consulting support.
- **HCMC AI Challenge 2024** *Oct 2024*
  - Ranked top 10 in the preliminary round.
  - Built a multi-user Vietnamese news-video search engine with the CLIP ViT-B/32 model and Streamlit.

## Leadership & Activities

---

- **HaUI Information Technology Club** Ha Noi, Vietnam  
*Data Engineer Lead* *Apr 2026 – Present*
  - Design lessons and explore trending Data Engineering concepts and tech stacks.
  - Host seminars and invite experienced data engineers to share knowledge.
  - Mentor students in high-level design, problem-solving, and data-pipeline optimization.
- **HaUI Artificial Intelligence Lab** Ha Noi, Vietnam  
*Commissioner of the AI Lab* *Sep 2023 – Dec 2025*
  - Participated in weekly seminars on technology advancements.
  - Studied new algorithms in LLMs, Computer Vision, and AI for scientific research.
  - Joined the Hackathon team as a technology consultant, guiding younger members.
- **HaUI Information Technology Club** Ha Noi, Vietnam  
*AI Engineer Lead* *Feb 2024 – Dec 2024*
  - Prepared and taught AI from scratch – Machine Learning and Deep Learning – to club members.
  - Ran workshops on AI tools and general AI/ML knowledge for students.