

Nguyễn Quốc Huy

Nam | 22 tuổi (2004) | (+84) 878-955-654 | ngquochuy010104@gmail.com
[linkedin.com/in/ngquochuy0101](https://www.linkedin.com/in/ngquochuy0101) | github.com/ngquochuy0101

TÓM TẮT

Sinh viên Công nghệ Thông tin tại Đại học Sài Gòn với định hướng Python Backend & AI Engineer. Có kinh nghiệm thực tế xây dựng và triển khai dự án như: Chatbot cho trường đại học Sài Gòn, Trợ lý AI cấp phường xã, Hệ thống Quản lý đặt lịch hẹn cho trại tạm giam, và Hệ thống tổng hợp yêu cầu cuộc họp tự động. Sử dụng tốt Python, FastAPI, Angular, Docker.

HỌC VẤN

Đại học Sài Gòn TP. Hồ Chí Minh, Việt Nam
Chuyên ngành: Công nghệ Thông tin; GPA: 3.09/4.0
KINH NGHIỆM LÀM VIỆC Tháng 10 2022 – Đã hoàn thành chương trình

Công ty TNHH Công nghệ và dịch vụ VPS TP. Hồ Chí Minh
Thực tập sinh / Lập trình viên Tháng 6 2026 – Tháng 9 2026

- Dự án tham gia:** Hệ Thống Quản Lý Đăng Ký Thăm Nuôi cho trại tạm giam (*ASP.NET Core 8, Angular, SQL Server*)
- Xây dựng giải pháp quản lý thăm gặp toàn diện tại trại giam, gồm đăng kí online, Kiosk tự phục vụ và Quản lý cho cán bộ.
- Nghiên cứu và hiện thực module nhận dạng phiếu OMR (Optical Mark Recognition) tự động nhận diện lựa chọn canteen từ giấy scan; tích hợp quét QR / CCCD và Face Auth tại Kiosk.
- Phát triển giao diện Admin Portal & Kiosk responsive bằng Angular Standalone + Angular Material + Tailwind CSS.

DỰ ÁN CÁ NHÂN

Hệ Thống Chatbot RAG Hỏi Đáp Đại Học Sài Gòn | *Python 3.12, FastAPI, Angular 22, Docker* 2026

- Hệ thống giúp sinh viên Đại học Sài Gòn hỏi đáp thông tin trực tuyến
- Thiết kế pipeline RAG nâng cao: Hybrid Search (Dense + BM25/RRF) trên Qdrant, Re-ranking (**bge-reranker**), HyDE và Semantic Cache.
- Xây dựng pipeline Crawling & Data Ingestion đa luồng (AsyncIO + multi-threading), tự động thu thập hàng nghìn trang web/PDF; tích hợp OCR xử lý công văn scan và phục hồi dấu tiếng Việt.

Hệ Thống Tổng Hợp Yêu Cầu Cuộc Họp | *Python, FastAPI, Redis, Celery, React* 2026

- Hệ thống hỗ trợ trích xuất nội dung cuộc họp tổng hợp nội dung xuất báo cáo.
- Thiết kế pipeline xử lý âm thanh thời lượng dài (5–8 tiếng) bằng cơ chế cửa sổ trượt (Sliding Window) kết hợp **faster-whisper** STT, đảm bảo trích xuất văn bản tốc độ cao và tối ưu VRAM.
- Tích hợp mô hình Ngôn ngữ Lớn (LLM) để tự động phân tích ngữ nghĩa, bóc tách chính xác các yêu cầu sửa đổi, phản hồi của khách hàng và tóm tắt nội dung cốt lõi của cuộc họp.

NGHIÊN CỨU KHOA HỌC

Mô hình Dual-View CNN cho Hệ thống IDS Mạng | *Huấn luyện Đối kháng* | [GitHub](#) Tháng 7 – Tháng 11 2025

- Bài báo khoa học: Nguyễn Quốc Huy, Phạm Duy Khánh, Xuất bản trên tạp chí REV-ECIT, Tháng 12 năm 2025.
- Thiết kế kiến trúc Dual-View CNN for Rare-class (DVRCNN) với cơ chế trích xuất các phụ thuộc chuỗi thời gian kết hợp với các tương quan không gian từ lưu lượng mạng và dung hợp đặc trưng dựa trên cơ chế chú ý để phát hiện tấn công mạng.
- Cải thiện điểm Macro F1 lên **+5.72%** và đạt độ chính xác **82.43%** trên tập dữ liệu chuẩn NSL-KDD bằng cách sử dụng kiến trúc DVRCNN kết hợp với việc giảm thiểu mất cân bằng dữ liệu bằng thuật toán ADASYN và Học có trọng số kết hợp Focal Loss.

GIẢI THƯỞNG

Bán kết | *Cuộc thi Đổi mới Sáng tạo Công nghệ cấp Thành phố InnoX (ĐH Văn Hiến & SIHUB)* Tháng 4 2026
Giải 3 / 25 Đội - Hạng mục Trợ lý Du lịch | *Hackathon SGU 2025 - 2026* Tháng 11 2025

KỸ NĂNG CHUYÊN MÔN

Ngôn ngữ & Backend: Python (FastAPI), C/C++, SQL, HTML, CSS, JavaScript/TypeScript, Redis
AI & Framework: RAG Pipeline, PyTorch, Qdrant, pgvector, Gemini, Angular, React/Vite
Hệ thống & Mạng: Windows, Ubuntu Linux, Docker, LAN/WAN, xử lý sự cố máy tính/phần mềm văn phòng