Nguyen Quoc Khanh

(+84) 039-269-7777 nqkhanh2002.github.io nqkdeveloper@gmail.com linkedin.com/nqkhanh2002

Summary

Applied AI Engineer with hands-on experience building LLM-powered applications Machine Learning, MLOps, Data Engineering. Aspiring highly motivated MSc/PhD candidate in Computer Science.

Education

University of Information Technology - VNUHCM June 2020 - Nov 2024
*Bachelor of Science in **Information System** (GPA: 3.19 / 4.00)* *Ho Chi Minh, Viet Nam*

Nguyen Du High School for the Gifted 2017 - 2020
Major Subject in Physics (GPA: 3.5 / 4.00) *Dak Lak, Viet Nam*

Experience

Silver Tiger Technology LLC - Dubai, UAE Aug 2025 – Present
AI Engineer *LangChain, LangGraph, Tensorflow Serving, LLM systems*

- Design and deploy RAG architectures, multi-agent systems, and LLM-powered applications for enterprise clients (Chatbot, Assistant, Phishing Detection).
- Self-hosted and optimized large-scale foundation models on dedicated 8×NVIDIA H200 GPU infrastructure, enabling high-throughput, low-latency inference for enterprise AI applications.
- Evaluate new technologies and tools to improve scalability, cost efficiency, and model lifecycle automation.

Speech Processing Department - VinBigdata - HCMC, Vietnam Dec 2024 – Jun 2025
AI Engineer *Data Engineering, MLOps, ASR, TTS*

- **Data Engineer** Developed and implemented **Data crawling, Labeling, Transcript Forced Alignment (TFA)** workflows within the SpeechWorld Data Framework. Improved Grapheme-to-Phoneme (G2P) models and acoustic models to enhance forced alignment accuracy and reliability.
- **Machine Learning Engineer/MLOps** Enhanced SpeechBase system with new features including model tuning, backend improvements, and online/active learning capabilities.

Project

AI Solution Platform - SilverTiger | *LLM, RAG, NLP, GraphRAG, Neo4j, Python, FastAPI*

- Developed modules of enterprise LLM-powered assistant with RAG-based knowledge retrieval, integrating foundation models for real-time multi-language question answering across 100+ knowledge sources.
- Designed ML evaluation and data generation framework including golden test cases, RAG benchmarking, and systematic model quality assessment, ensuring consistent performance across multi-language inputs (Vietnamese, English, and others).

Oten Trust - AI Security Platform - SilverTiger | *LLM, XAI, NLP, Python, FastAPI*

- Led AI development of LLM-based phishing detection system with natural language reasoning, implementing explainable AI (XAI) for transparent trust scoring and cross-language scam detection across Vietnamese and English inputs.

Sign Language Translation - Graduation Thesis | *TypeScript, HumanGAN, LLM, Skeleton Viewer, Mobile/Web*

- Developed a real-time sign language translation solution with multi-language support, leveraging cutting-edge machine learning for accurate sign-to-text and text-to-sign conversions.
- Enhanced user interaction through an intuitive interface for both desktop and mobile, enabling speech-to-text, text normalization, and internationalization across 107 languages.
- Innovated with 3D avatar animations and Human GAN technology for realistic sign language visualization, complete with video output features for sharing and accessibility.

Publication & Manualscript

Adapting WavLM for Vietnamese Speaker Diarization in Real-world Conversations 05/2025
Thang Tuan Duy, Do Tri Nhan, Nguyen Van Huy, Nguyen Quoc Khanh, Phan Trung Kien, Mac Dang Khoa ***MAPR 2025***

An Orchestrated Framework for Automated Speech Data Processing and Alignment 05/2025
Nguyen Quoc Khanh, Nguyen Van Huy, Phan Van Tuan, Do Tri Nhan, Mac Dang Khoa ***Manualscript***