

Thien An NGUYEN

Hanoi, Vietnam | nguyenthienan1162003@gmail.com | +8434 288 4797 | <https://github.com/TGraceAn>
<https://www.linkedin.com/in/thien-an-nguyen-5325a61a7/>

Introduction

I am a graduate student pursuing a Master's degree. I love to do research and am passionate about Generative Models, Audio/Speech Models, NLP, Computational Creativity, and On-device Models. I aim to contribute meaningfully to the academic community while furthering my intellectual growth and professional development. Moreover, I also want to create different software that can benefit the public.

Education

University of Science and Technology of Hanoi , MSc in Data Mining for IoT	Sept 2024 – Sept 2026 (Expected graduation date)
<ul style="list-style-type: none">• GPA: 15.04/20 as for now (French grading systems) ~3.57/4.0 (US grading systems conversion)• Coursework: Machine Learning, Deep Learning, Modeling Techniques, Information Systems, Security and Ethics for Data, Systems Architecture, HPC Programming, etc.	
INP Toulouse , MSc in Data Science (Double Degree program with USTH)	Sept 2024 – Sept 2026 (Expected graduation date)
University of Science and Technology of Hanoi , BSc in ICT	Sept 2021 – Aug 2024
<ul style="list-style-type: none">• GPA: 16.24/20 (French grading systems) ~3.61/4.0 (US grading systems conversion)• Coursework: Machine Learning, Deep Learning, Web Application Development, Mobile Application Development, Digital Signals Processing, Natural Language Processing, Distributed Systems, Cryptography, Database, etc.	

Experience

Research Assistant , ICT Lab - University of Science and Technology of Hanoi	Feb 2025 – Present
<ul style="list-style-type: none">• Working on on-device open-set KWS.• Implemented distillation from large model for better on-device KWS model.• Implemented RL for multi-teacher distillation.	
AI Engineer , eUp Group (Full-time)	Jan 2025 – Present
<ul style="list-style-type: none">• Used RAG and text embeddings for personalized dictionary lookup. (Mazii)• Developed AI for explaining Japanese questions in multiple languages. (Migii-JLPT)• Developed AI for generating Japanese exercises which match difficulty levels. (Migii-JLPT)• Working on Japanese speaking evaluation using both speech and transcript. (JOPT)• Researching on on-device translation models from speech to speech.• Used frameworks such as n8n to build Agent for Japanese data labeling.	
Teaching Assistant , University of Science and Technology of Hanoi (Part-time)	Oct 2024 – Mar 2025
<ul style="list-style-type: none">• Teaching Course(s): Signals and Systems, NLP.• Prepare exercises.• Help students understand lectures.	
Generative AI Developer , Amela Technology (Full-time)	Nov 2024 – Jan 2025
<ul style="list-style-type: none">• Researched about OCR for Japanese handwriting text.• Researched about Japanese handwriting text generation.• Learned about VLM, RAG, and Diffusion Models.	
Research Intern , L3i Lab - La Rochelle University – La Rochelle, France. (Full-time)	Apr 2024 – Jul 2024
<ul style="list-style-type: none">• Researched on Computational Creativity in Music Generation.	

- Proposed methods for handling Music Semantics over notes, dynamics, and time. (Which Note, at which Velocity, and when is the Time to play the note)
- Processed data from MIDI files to create tokens for Symbolic Music.
- Designed models for Symbolic Music Generation using Transformer.
- Utilized Relative-Attention mechanism to improve long dependencies learning.
- Evaluated objectively the performance of Symbolic Music Generation models assistive and autonomously.
- Proposed a method to evaluate the performance of the models subjectively.

Projects

DaiJazz

Apr 2024 - Jul 2024

Experimented and built an AI model to generate music autonomously and assistive.

- Members: 5.
- Position: Researcher, Main Developer.
- Tasks: Process MIDI data from MIDI Files, find the best text representation, build a tokenizer for the Transformer-based model, and build Transformer-based sequential models to handle different types of tokens. Experiment with different models and give evaluations.
- Tools Used: Python, PyTorch, json, MIDI Files, Transformer, Relative Transformer, GPT-2, etc.

Indoor Assistant for Blind People

Nov 2023 - Jan 2024

The project aimed to develop an affordable system utilizing technology and artificial intelligence to assist people with vision impairment in navigating inside their houses. Additionally, it includes a feature to describe the surrounding scenery to aid them in navigating around objects.

- Members: 6.
- Position: Developer.
- Tasks: Run inference and integrate the models into a system.
- Tools Used: yolov8, MiDaS, Python, PyTorch, etc.

USTH Type-1 Project

Feb 2024 - Aug 2024

Create an AI model for Vietnamese Keyword Spotting. Develop a model for few-shot learning and deploy it on edge devices.

- Members: 7.
- Position: Developer, Tester.
- Tasks: Train and test the AI model.
- Tools Used: Python, AIoT AutoCar II, NaoV6, Keyword Transformer, yaml, etc.

Technologies

Programming Languages: Python, C/C++, Java, JS, PHP, HTML/CSS.

Frameworks: PyTorch, TensorFlow, Cuda jit, etc.

Technologies: Git, shell, Jupyter Notebook, GPU Clusters, etc.

Operating Systems: MacOS, Linux, Windows.

Additional Skills

Languages: Fluent in English and Vietnamese; German at B1 level; French at A2 level.

Soft skills: Self-study, teamwork, problem-solving skills

Musically: Bass player for NewCreation Band. Featured on VTV.

References

Hoang Tung Tran, PhD - Deputy Director

Department of Information and Communication Technology, University of Science and Technology of Hanoi.

Email: tran-hoang.tung@usth.edu.vn

Giang Son Tran, Assoc. Prof. - Director

Department of Information and Communication Technology, University of Science and Technology of Hanoi.

Email: tran-giang.son@usth.edu.vn

Additional: Provide when needed.