

# Hieu Tran

0935 957 250 | [tthieu22.06@gmail.com](mailto:tthieu22.06@gmail.com) | [linkedin.com/in/tthieu2206](https://www.linkedin.com/in/tthieu2206) | [github.com/heraclex12](https://github.com/heraclex12)

## EXPERIENCE

---

### NLP Engineer

*M\_Service (MoMo)*

August. 2021 – Present

*Ho Chi Minh, Vietnam*

- Build contextual AI chatbots for Customer Support on MoMo e-wallet platform
- Deploy chatbot models on different environments using Kubernetes, GCP
- Implement and train Transformer-based models on multiple GPUs

### AI Research Collaborator

*VietAI Research*

May. 2021 – Present

*Ho Chi Minh, Vietnam*

- Research machine translation models for English and Vietnamese that outperform previous SOTA by 4.5 BLEU

### Undergraduate Research Student

*AISIA Research Lab*

September. 2020 – July. 2021

*Ho Chi Minh, Vietnam*

- Conducted studies on the advantage of Transformer-based models for Vietnamese Punctuation Prediction

### AI Software Engineer Intern

*Zalo Group - VNG Corporation*

April. 2019 – August. 2021

*Ho Chi Minh, Vietnam*

- Constructed RDF-triplestore Knowledge Base for Kiki - first Vietnamese AI virtual assistant
- Applied robust Transformer-based models for Intent/Topic Classification, Named Entity Recognition
- Invented a state-of-the-art Vietnamese Spelling Correction model
- Developed rule-based approach for domain-specific Vietnamese Question Answering system

## EDUCATION

---

### VNUHCM - University of Science

*Bachelor of Computer Science, GPA: 8.3/10*

Ho Chi Minh, Vietnam

*Aug. 2017 – Jul. 2021*

## PROJECTS

---

### VLSP2020: Fake News Detection | *Python, Pytorch, Transformers*

- Fine-tuned a variety of Transformer-based models to identify reliable information shared on social network sites

### HackEng | *Python, Flask, Tensorflow, Git*

- Built a standardized English test website utilizing T5 model to automatically generate reading comprehension questions and correct answers from reliable text sources

### NLP2SPARQL | *Python, Pytorch, Transformers, Fairseq*

- Employed Transformer-based Encoder-Decoder models (e.g., BERT2BERT) to translate between natural language and SPARQL query language on Question Answering over Knowledge Graphs

## HONORS AND AWARDS

---

**Outstanding Undergraduate Thesis Award**, *Undergraduate Thesis*, VNUHCM - University of Science

**First Prize**, in *MindX Grand Hackathon 2020*, MindX Technology School

**Second Place**, in *Apkathon 2019*, Aptech Computer Education

## TECHNICAL SKILLS

---

**Languages:** Python, Java, SQL, SPARQL

**Developer Tools:** Git, Docker, Google Cloud Platform, MongoDB, PostgreSQL, Kubernetes, Elasticsearch, Linux

**Libraries:** Pytorch, Tensorflow, Transformers, scikit-learn, Pandas, Numpy, Flask

## ACTIVITIES

---

**MTet: Multi-domain Translation for English and Vietnamese** (blog), talk at Google I/O Extended Hanoi 2022

**How MoMo Builds an AI Chatbot**, published on MoMo Technology Blog (Medium)

Participated in **BigScience Research Workshop**, mainly organized by HuggingFace

## PUBLICATIONS

---

- Laurençon et al. 2022. The BigScience ROOTS Corpus: A 1.6TB Composite Multilingual Dataset. In *Proceedings of NeurIPS 2022 Track Datasets and Benchmarks*
- Chinh Ngo, Trieu H. Trinh, Long Phan, **Hieu Tran**, Hieu Nguyen, Tai Dang, Minh Nguyen and Minh-Thang Luong. 2022. MTet: Multi-domain Translation for English and Vietnamese. In *reviewing EACL 2022*
- Long Phan, **Hieu Tran**, Hieu Nguyen and Trieu H. Trinh. 2022. ViT5: Pretrained Text-to-Text Transformer for Vietnamese Language Generation. In *Proceedings of NAACL SRW 2022*
- Hieu Tran**, Long Phan, James T. Anibal, Binh T. Nguyen and Truong-Son Nguyen. 2021. SPBERT: An Efficient Pre-training BERT on SPARQL Queries for Question Answering over Knowledge Graphs. In *Proceedings of ICONIP 2021*, Acceptance rate  $\sim 20\%$
- Long Phan, **Hieu Tran**, Daniel Le, Hieu Nguyen, James T. Anibal, Alec Peltekian and Yanfang Ye. 2021. CoText: Multi-task Learning with Code-Text Transformer. In *Proceedings of NLP4Prog @ ACL-IJCNLP 2021*
- Hieu Tran**, Cuong Viet Dinh, Long Phan and Son Nguyen. 2021. Hierarchical Transformer Encoders for Vietnamese Spelling Correction. In *Proceedings of IEA/AIE 2021*
- Hieu Tran**, Cuong Viet Dinh, Quang Hong Pham and Binh T. Nguyen. 2021. An Efficient Transformer-Based Model for Vietnamese Punctuation Prediction. In *Proceedings of IEA/AIE 2021*

**Google Scholar:** <https://scholar.google.com/citations?hl=en&user=atONCyEAAAAJ>