

# Hieu Tran

[hieutt@umd.edu](mailto:hieutt@umd.edu) | [scholar.google.com](https://scholar.google.com) | [linkedin.com/in/hieutt0](https://linkedin.com/in/hieutt0) | [heracler12.github.io](https://heracler12.github.io)

## RESEARCH INTERESTS

---

**LLM Interpretability:** representation engineering & activation editing (steering vectors) for reliable control

**Agentic LLMs & Multi-Agent Systems:** deep research, coding & web agents, coordination & tool use optimization

**Machine Translation:** multilingual transfer & synthetic data augmentation for low-resource MT

## EDUCATION

---

### University of Maryland, College Park

*PhD Student in Computer Science, supervised by Marine Carpuat*

- Achievements: Dean's Scholarship

Maryland, USA

*Aug. 2024 – Present*

### VNUHCM - University of Science

*Bachelor of Computer Science*

- Achievements: Outstanding Undergraduate Thesis Award

Ho Chi Minh, Vietnam

*Aug. 2017 – Jul. 2021*

## WORK EXPERIENCE

---

### Senior Machine Learning Engineer

*BLACKBOX AI*

- Pioneered core features as an early engineer for flagship products, including the VSCode Coding Agent, Web Agent, Auto Completion, scaling the user base from their early stages to over **30M users** and **4M VSCode** downloads
- Owned the end-to-end development of the code search/web search system, from researching ranking algorithms to implementing the final retrieval pipeline used in production
- Trained CodeLlama-34B on synthetic instruction-following datasets, achieving **86.6%** on HumanEval, surpassing GPT-4 (Nov 2023) which scored 85.4%

Jun. 2023 – Jul. 2024

*Remote (Canada)*

### Lecturer & Reseach Collaborator

*VietAI*

- Taught Foundations of Deep Learning, Advanced NLP, and LLMs courses to over **200 students**
- Worked directly with Dr. Minh-Thang Luong and Dr. [Trieu H. Trinh](#) (Google DeepMind) to build English-Vietnamese machine translation systems that outperform previous SOTA by 4.5% (BLEU).
- Released the first Vietnamese generative models, achieving SOTA in Summarization and Biomedical benchmarks.

May. 2021 – Aug. 2024

*Ho Chi Minh, Vietnam*

### AI Research Engineer

*M\_Service (MoMo) (3rd unicorn in Vietnam)*

- Built Intent-based/ChatGPT-based AI chatbots for Customer Support on MoMo #1 e-wallet platform in Vietnam
- Managed the training pipeline and chatbots on Kubernetes deployed on Google Cloud Platform
- Trained base language models that outperformed open-source Vietnamese models for our internal needs

Aug. 2021 – Aug. 2023

*Ho Chi Minh, Vietnam*

### Undergraduate Research Student

*AISIA Research Lab*

- Explored the efficiency of Transformer-based models for Vietnamese Punctuation Prediction

Sep. 2020 – Feb. 2022

*Ho Chi Minh, Vietnam*

### AI Software Engineering Intern

*Zalo Group - VNG Corporation (1st unicorn in Vietnam)*

- Constructed RDF-triplestore Knowledge Bases for Kiki - the first Vietnamese AI virtual assistant
- Proposed a light and state-of-the-art Vietnamese Spelling Corrector. [Try it!](#)

Apr. 2019 – Aug. 2021

*Ho Chi Minh, Vietnam*

## TECHNICAL SKILLS

---

**Languages:** Python, NodeJS, Typescript, SQL

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

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

## PUBLICATIONS

---

1. Tai Dang, **Hieu Tran**, Sang T. Truong, Long-Hung Pham, Edward A Pham, Jeffrey S. Glenn, Thang Luong. 2026. High-Fidelity Molecular Structure Prediction via Reinforcement Learning. *Under submission to ICML 2026*
2. **Hieu Tran**, Calvin Bao, Marine Carpuat. 2026. Beyond Binary Attributes: Aspect-Aware Activation Steering for Authorship Style Transfer. *Under submission to ACL 2026*
3. **Hieu Tran\***, Phuong-Anh Nguyen-Le\*, Huy Nghiem, Quang-Nhan Nguyen, Wei Ai, Marine Carpuat. 2026. VietMix: A Naturally Occurring Vietnamese-English Code-Mixed Corpus with Iterative Augmentation for Machine Translation. In *EACL 2026*
4. Long Phan\*, Tai Dang\*, **Hieu Tran\***, Trieu H. Trinh\*, Vy Phan, Lam D. Chau, Minh-Thang Luong. 2023. Enriching Biomedical Knowledge for Low-resource Language Through Large-Scale Translation. In *EACL 2023*
5. Chinh Ngo\*, Trieu H. Trinh\*, Long Phan\*, **Hieu Tran\***, Hieu Nguyen, Tai Dang, Minh Nguyen, Minh-Thang Luong. 2022. MTet: Multi-domain Translation for English and Vietnamese. In *ArXiv Preprint*
6. 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*
7. **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  $\sim 20\%$
8. 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*
9. **Hieu Tran**, Cuong Viet Dinh, Long Phan and Son Nguyen. 2021. Hierarchical Transformer Encoders for Vietnamese Spelling Correction. In *Proceedings of IEA/AIE 2021*
10. **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*

## PROJECTS

---

### **Coding Agent on Browser** | *Typescript, APIs, CodeSandbox*

- Built a browser UI for user-submitted coding tasks; an LLM agent follows the *ReAct* framework to plan steps and decide which tools to call
- Executes a chain of tools to operate on the project file system, edit/read/search files and run a sandboxed terminal, to iteratively complete tasks end-to-end

### **Multi-Agent Machine Translation System** | *Python, Transformers, LiteLLM*

- Designed a multi-agent framework that orchestrates specialized LLM agents (*ResearchAgent*, *DraftAgent*, *RefineAgent*, *EvaluationAgent*) to collaboratively perform translation for low-resource languages

### **Deep Research Agent System** | *Python, LLM Agents, RAG*

- Developed a deep research agent that autonomously plans, retrieves, and synthesizes information from the web and internal knowledge sources to produce comprehensive, citation-grounded research reports
- Implemented multi-step reasoning, tool-use orchestration, and iterative refinement to improve factuality, coverage, and reliability of long-horizon research tasks

## TALKS AND ACTIVITIES

---

**The Art of Generative AI**, talk at Google Developer Student Clubs - IU 2024, 100+ audiences

**Unlocking the potential of GenAI for coding**, talk at Google Devfest HCMC 2023, 100+ audiences

**Explore the Power of AI - Career Trends in This Industry**, USAID WISE Online Event 2023, 500+ audiences

**Shaping Your NLP Career Path in 2024**, talk at VietAI Online Event 2023, 150+ audiences

**Challenges and Opportunities When Working In AI Industry (2022)**, Van Lang University, 100+ audiences

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