

Nguyen Song Thien Long
 Publication name: Long S. T. Nguyen

long.nguyencse2023@hcmut.edu.vn
 +84-837026259
 longstnguyen.com
 Long S. T. Nguyen



Research Interests

My research interests lie in **Natural Language Processing (NLP)**, with a primary focus on Information Retrieval (IR), Question Answering (QA), and Large Language Models (LLMs). My work centers on Retrieval-Augmented Generation (RAG), hybrid lexical-semantic retrieval, and knowledge graph-grounded, multi-hop reasoning in both agentic and standalone LLM systems. I am especially interested in making these methods efficient and trustworthy, including in low-resource and multilingual settings.

Employment

- Apr 2026 – Present

AI Researcher.

Saigon AI Hub, VNG Group.

Serving as *Principal Investigator* for a granted research project on structured reasoning and training strategies for LLMs, funded by VNG Group and Vietnam National University Ho Chi Minh City (VNU-HCM).
- Oct 2025 – Present

AI Researcher.

Center for AI Research (CAIR), VinUniversity.

Investigating the training of lexical-dense and hybrid retrieval models and their integration into LLM-based QA systems, mentored by Prof. Luu Anh Tuan (NTU, Singapore) and Dr. Nguyen Tran Cong Duy (CAIR).
- Dec 2023 – Present

Research Assistant.

URA Research Group, Faculty of Computer Science and Engineering, HCMUT.

Conducting research on IR, QA, and LLM-based systems under the supervision of Prof. Quan Thanh Tho, and serving as *Team Leader*, defining research directions and mentoring junior researchers.

Education

- 2023 – 2027

B.Sc. Computer Science (High-Quality Program),

Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Vietnam. CGPA: 3.6/4.0.

Selected Publications
 (full list at <https://longstnguyen.com/publications/>)

- 2026

Long S. T. Nguyen, Dat T. Huynh, Quynh T. N. Vo, Dang Le, Cong-Duy T. Nguyen, Anh Tuan Luu, Anh Nguyen-Duc, and Tho T. Quan.

*Fast and Focused: Interactive Repository-Level Fill-in-the-Middle Code Completion with Hybrid Structural Retrieval.* Conference on Empirical Methods in Natural Language Processing (EMNLP). (CORE A\*)

Long Nguyen, Duc Nguyen, Quan Bui, Dung Phan, Dung Le, Khanh Nguyen, Quynh Vo, Khang Vo, Nam Duong, Anh Dinh, Tri Trinh, Chi Phan, An Nguyen, Thai Nguyen, Dang Le, Vinh Dang, and Tho Quan.

*Catching the First Light of Tomorrow: A Hackathon-Based Framework for Introducing High School Students to AI Agents.* Symposium on Educational Advances in Artificial Intelligence (EAAI), at AAAI. (CORE A\*)

Long S. T. Nguyen and Tho T. Quan.

*Which Works Best for Vietnamese? A Practical Study of Information Retrieval Methods across Domains.* Findings of the Association for Computational Linguistics: EACL. (CORE A)

Hung C. Luu, Long S. T. Nguyen, Trung M. Pham, Hieu M. Pham, and Tho T. Quan.

*HiGraAgent: Dual-Agent Adaptive Reasoning over Hierarchical Knowledge Graph for Open-Domain Multi-hop Question Answering.* Findings of the Association for Computational Linguistics: EACL. (CORE A)

Quoc Phu Dang\*, Phat T. Tran-Truong\*, Duc-Ly Vu, Long S. T. Nguyen, Quynh T. N. Vo, and Tho T. Quan.

*Enhancing Large Language Model Performance for Automatic Zero-Shot Multiple-Choice Question Answering via Single-Token Logit Prompting.* Computers and Education: Artificial Intelligence. (SCIE, Q1, #1 IN AI)
- 2025

Long S. T. Nguyen, Quynh T. N. Vo, Hung C. Luu, and Tho T. Quan.

*When in Doubt, Ask First: A Unified Retrieval Agent-Based System for Ambiguous and Unanswerable Question Answering.* Findings of the Association for Computational Linguistics: IJCNLP-AAACL. (CORE B)

Long S. T. Nguyen, Tran T. B. Le, Huong P. N. Nguyen, Quynh T. N. Vo, Phong H. N. Nguyen, and Tho T. Quan.

*Serving the Underserved: Leveraging BARTBahnar Language Model for Bahnaric-Vietnamese Translation.* Workshop on Language Models for Underserved Communities (LM4UC), at NAACL. (WORKSHOP, CORE A)

Academic Service

- Prog. Cmte.

ACM WSDM 2027 (CORE A); IEEE eScience 2026 (CORE B).
- Conf. Rev.

ACM MM 2026 (CORE A\*).
- Journal Rev.

IEEE Transactions on Neural Networks and Learning Systems (Q1); Journal of Natural Language Processing (Q1); Behaviour & Information Technology (Q1); IEEE Access (Q1); Frontiers in Artificial Intelligence (Q1); Frontiers in Education (Q1); Connection Science (Q2); Int. Journal of Knowledge-Based and Intelligent Engineering Systems (Q3).
- Olympiad

Technical and training support for VOAI 2025 and the preparation of Vietnam’s IOAI 2025 team.

References

- Supervisor

Prof. Quan Thanh Tho. Dean, Faculty of Computer Science and Engineering, HCMUT. qttho@hcmut.edu.vn
- Mentor

Prof. Luu Anh Tuan. College of Computing and Data Science, Nanyang Technological University (NTU), Singapore. anhtuan.luu@ntu.edu.sg