

# Thanh Le-Cong

Last updated: September 18, 2025

[thanhcls1316@gmail.com](mailto:thanhcls1316@gmail.com) | [thanhlecong](https://github.com/thanhlecong) | [Thanh Le-Cong](https://www.linkedin.com/in/thanhlecong) | [thanhlecong](https://www.linkedin.com/in/thanhlecong)

Melbourne, Victoria, Australia

## Research Interests

My research interests lie at the intersection of Software Engineering (SE) and Artificial Intelligence (AI), particularly focusing on **Trustworthy AI for SE**. During my PhD candidature, I have been investigating the trustworthiness of Large Language Models (LLMs) in code synthesis and analysis, systematically examining key dimensions such as reasoning capabilities, robustness, output quality, memory efficiency, privacy, and security. Additionally, I have been also exploring the application of machine learning methods to automated software debugging, focusing on critical tasks such as program repair and bug detection.

Looking ahead, my research vision is to develop **neural-symbolic techniques** that integrates the adaptability and expressiveness of “neural” LLMs with the rigor and precision of “symbolic” program analysis for reliably automating software engineering. I believe that this framework can mitigate the trustworthiness issues of LLMs, enabling the creation of trustworthy autonomous agents for software engineering.

## Education

### • University of Melbourne

*PhD in Engineering and IT*

*February 2023 - Dec 2025 (Expected)*

Melbourne, Australia

- Thesis: Towards Reliable AI-assisted Program Repair
- Advisors: Prof. Toby Murray and Dr. Bach Le
- Funded by Melbourne Research Scholarship
- Awarded Google PhD Fellowship in Software Engineering and Programming Technology (awarded to only two PhD Students worldwide)

### • Hanoi University of Science and Technology

*Bachelor of Engineer in Information Technology*

*August 2016 - August 2021*

Hanoi, Vietnam

- Thesis: Automated Patch Validation for Program Repair
- Advisors: Prof. Quyet-Thang Huynh
- Awarded Best Undergraduate Thesis Award (Top 1 in my cohort)
- Selected to Talented Program (Top 30 in Admission)

## Selected Peer-Review Publications

\* *Note: Underlined authors are students under my mentoring. # denotes equal contributions*

### Journal Papers

[J7] **Thanh Le-Cong**, Dat Nguyen, Bach Le, Toby Murray, "Toward Reliable Evaluation of Neural Program Repair with Natural Robustness Testing", In **TOSEM 2025**: ACM Transaction on Software Engineering and Methodology.

[J6] Abdulaziz Alhefdhi, Hoa Khanh Dam, **Thanh Le-Cong**, Bach Le, Aditya Ghose, "Adversarial Patch Generation for Automated Program Repair", In **SQJ 2025**: Software Quality Journal.

[J5] Yue Liu, **Thanh Le-Cong**, Ratnadira Widayarsi, Chakkrit Tantithamthavorn, Li Li, Bach Le, David Lo, "Refining ChatGPT-Generated Code: Characterizing and Mitigating Code Quality Issues", In **TOSEM 2024**: ACM Transactions on Software Engineering and Methodology. (Journal-first Track at **FSE 2025**).

[J4] Xin Zhou, Bowen Xu, Kisub Kim, DongGyun Han, Hung Nguyen, **Thanh Le-Cong**, Junda He, Bach Le, David Lo, "Leveraging Large Language Models for Automatic Patch Correctness Assessment", In **TSE 2024**: IEEE Transactions on Software Engineering.

[J3] **Thanh Le-Cong**, Duc-Minh Luong, Xuan Bach D. Le, David Lo, Nhat-Hoa Tran, Bui Quang-Huy, "Invalidator: Automated Patch Correctness Assessment via Semantic and Syntactic Reasoning", In **TSE 2023**: IEEE Transactions on Software Engineering. (Journal-first Track at **ICSE 2024**).

[J2] Truong-Giang Nguyen, **Thanh Le-Cong**, Hong Jin Kang, Ratnadira Widayarsi, Chengran Yang, Zhipeng Zhao, Bowen Xu, Jiayuan Zhou, Xin Xia, Ahmed E. Hassan, Bach Le, David Lo, "MiDas: Multi-granularity Model for Vulnerability-fixing Commit Classification", In **TSE 2023**: IEEE Transactions on Software Engineering.

[J1] Thi Thanh Binh Huynh, Dinh Thanh Pham, Ba Trung Tran, **Thanh Le-Cong**, Minh Hai Phong Le, Ananthram Swami, Thu Lam Bui, "A multifactorial optimization paradigm for linkage tree genetic algorithm", In **IS 2020**: Information Sciences.

## **Conference Papers**

[C8] **Thanh Le-Cong**, Bach Le, Toby Murray, "Can LLMs Reason About Program Semantics? A Comprehensive Evaluation of LLMs on Formal Specification Inference", In **ACL 2025**: Annual Meeting of the Association for Computational Linguistics.

[C7] Duong Nguyen, **Thanh Le-Cong**, Triet Huynh Minh Le, M. Ali Babar, Quyet-Thang Huynh, "Toward Realistic Evaluation of Just-In-Time Vulnerability Prediction", In ICSME 2025: International Conference on Software Maintenance and Evolution

[C6] Yen-Trang Dang, **Thanh Le-Cong**, Phuc-Thanh Nguyen, Anh M. T. Bui, Phuong T. Nguyen, Bach Le, "LEGION: Harnessing Pre-trained Language Models for GitHub Topic Recommendations with Distribution-Balance Loss", In **EASE 2024**: International Conference on Evaluation and Assessment in Software Engineering

[C5] Ratnadira Widyasari, Zhipeng Zhao, **Thanh Le-Cong**, Hong Jin Kang, David Lo, "Extreme Multi-Label Learning for GitHub Topic Recommendation: How Far Are We?", In **SANER 2023**: IEEE International Conference on Software Analysis, Evolution, and Reengineering

[C4] Yunbo Lyu<sup>#</sup>, **Thanh Le-Cong**<sup>#</sup>, Hong Jin Kang, Ratnadira Widyasari, Zhipeng Zhao, Bach Le, Ming Li, David Lo, "Chronos: Time-Aware Zero-Shot Identification of Libraries from Vulnerability Reports", In **ICSE 2023**: IEEE/ACM International Conference on Software Engineering.

[C3] **Thanh Le-Cong**, Hong Jin Kang, Truong-Giang Nguyen, Stefanus Agus Haryono, David Lo, Bach Le, Thang Huynh, "AutoPruner: Transformer-based Call Graph Pruning", In **ESEC/FSE 2022**: ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering

[C2] Thanh-Dat Nguyen, **Thanh Le-Cong**, Thanh Le-Cong, Duc-Minh Luong, Van-Hai Duong, Bach Le, David Lo, "FFL: Fine-grained Fault Localization for Student Programs via Syntactic and Semantic Reasoning", In **ICSME 2022**: IEEE International Conference on Software Maintenance and Evolution

[C1] **Thanh Le-Cong**, Xuan Bach D. Le, Quyet Thang Huynh, Phi Le Nguyen, "Better Together for Automated Repair of Web Pages", In **ISSRE 2021**: IEEE International Symposium on Software Reliability Engineering

## **Selected Short/Workshop Papers**

[S4] Duong Nguyen, Duc-Manh Tran, **Thanh Le-Cong**, Triet Huynh Minh Le, M. Ali Babar, Quyet-Thang Huynh, "VulGuard: An Unified Tool for Evaluating Just-In-Time Vulnerability Prediction Models", In **ICSME-Tool 2025**: International Conference on Software Maintenance and Evolution, Tool Demos Track

[S3] Bowen Xu, Thanh-Dat Nguyen, **Thanh Le-Cong**, Thong Hoang, Jiakun Liu, Kisub Kim, Chen Gong, Changan Niu, Chenyu Wang, Bach Le, David Lo, "Are We Ready to Embrace Generative AI for Software Q&A?", In **ASE-NIER 2023**: IEEE/ACM International Conference on Automated Software Engineering, NIER Track

[S2] Thanh-Dat Nguyen<sup>#</sup>, **Thanh Le-Cong**<sup>#</sup>, ThanhVu H. Nguyen, Bach Le, Quyet-Thang Huynh, "Towards the Analysis of Graph Neural Networks", In **ICSE-NIER 2022**: IEEE/ACM International Conference on Software Engineering, NIER Track,

[S1] Truong-Giang Nguyen, **Thanh Le-Cong**, Hong Jin Kang, Bach Le, David Lo, "VulCurator: A Vulnerability Fixing Commit Detector", In **ESEC/FSE-Tool 2022**: ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering, Demos Track

## **Ongoing Papers**

[O6] **Thanh Le-Cong**, Bach Le, Toby Murray, "Memory-Efficient Large Language Models for Program Repair with Semantic-Guided Patch Generation", Submitted to **ICSE 2026**: IEEE/ACM International Conference on Software Engineering, Major Revision

[O5] Huu-Hung Nguyen, Duc-Manh Tran, Yiran Cheng, **Thanh Le-Cong**, Hong Jin Kang, Ratnadira Widyasari, Shar Lwin Khin, Ouh Eng Lieh, Ting Zhang, David Lo, "Mapping NVD Records to Their VFCs: An Exploratory Study", Submitted to **TOSEM**: ACM Transactions on Software Engineering and Methodology, Major Revision

[O4] Hua Yang, Alejandro Velasco, **Thanh Le-Cong**, Md Nazmul Haque, Bowen Xu, Denys Poshyvanyk, "How Do Semantically Equivalent Code Transformations Impact Membership Inference on LLMs for Code?", Submitted to **ICSE 2026**: IEEE/ACM International Conference on Software Engineering, Major Revision

[O3] Anh Ho, **Thanh Le-Cong**, Bach Le, Christine Rizkallah, "From Empirical Evaluation to Context-Aware Enhancement: Repairing Regression Errors with LLMs", Submitted to **TOSEM**: ACM Transactions on Software

Engineering and Methodology, Major Revision


[O2] [Huu-Hung Nguyen](#), [Anh Tuan Nguyen](#), **Thanh Le-Cong**, Yikun Li, Han Wei ANG, Yide Yin, Frank Liauw, Shar Lwin Khin, Ouh Eng Lieh, Ting Zhang, David Lo, "*PatchSeeker: Mapping NVD Records to their Vulnerability-Fixing Commits with LLM Generated Commits and Embeddings*", Submitted to **ICSE**: IEEE/ACM International Conference on Software Engineering

[O1] [Hung Le](#), **Thanh Le-Cong**, Bach Le, Bowen Xu "*Optimizing Spectral Signature in Code Backdoor Detection*", Submitted to **FSE**: IEEE/ACM International Conference on Foundation Software Engineering

## Grants

- Principal Investigator, "Trustworthy Neural Program Repair", AUD 15.000, granted by **Google PhD Fellowship Scheme**.
- Principal Investigator, "Logically Proving Bugs in Randomised Software", GBP 293.660, held in reserve (shortlisted but not granted) by **Royal Society - Newton International Fellowship Scheme**

## Honors and Awards

- **Melbourne Research Scholarship** 2022  
*The University of Melbourne* 
- **Best Thesis Award** 2022  
*School of Communication and Information Technology, HUST*
- **First Prize in the National Mathematical Olympiad for University Students** 2017  
*Vietnam Mathematical Society*
- **Second Prize in the National Mathematical Olympiad for High School students** 2016  
*Vietnam Ministry of Education and Training*

## Research Experiences

- **CISPA Helmholtz Center for Information Security** June 2025 - Now  
*Visiting PhD Student* Stuttgart, Germany
  - Supervisors: Prof. Cristian Cadar (Imperial College London), Prof. Michael Pradel (CISPA)
  - Working on specification inference with LLMs from natural language artifacts.
- **Amazon Web Services** June 2023 - Sep 2023  
*Applied Scientist II (Internship)* Virginia, US
  - Supervisor: Dr. Brandon Paulsen, Dr. Joey Dodds, Prof. Daniel Kroening
  - Developed a repo-level code translation technique using Large Language Models and Automated Reasoning;
  - Received a return offer for next internship.
- **Singapore Management University** Jan 2022 - Jan 2023  
*Research Engineer* Singapore
  - Supervisor: Prof. David Lo
  - Developed AI-enabled techniques for managing vulnerabilities in third-party software;
  - Published research papers at top-tier conferences, e.g., FSE and ICSE.

## Teaching Experiences

- **University of Melbourne** Jan 2024 - Now  
*Academic Tutor*
  - Subject: High Integrity Systems Engineering - Master Course (SWEN90010) - Semester 1/2025
  - Delivered 3 weekly tutorial classes with size of 30 student each class
  - Marked assignments and exams
- **Hanoi University of Science and Technology** Aug 2021 - Now  
*Research Mentor/Lead at [AI4Code Lab](#)* Remote (Part-time)
  - Co-founded this research group with Prof. Quyet Thang Huynh; mentor to 5+ undergraduate research students annually;
  - Designed a training program and organized a weekly seminar series for training undergraduate students;
  - My students have received full scholarships for pursuing PhD from well-known universities such as Singapore Management University, North Carolina State University, University of Melbourne and University of Sydney.

## • North Carolina State University

Sep 2024

Guest Lecturer

- Generative AI for SE - Graduate courses (CS591/791)
- Delivered guest lecture on "Reliability of Code Generation with LLMs"
- Conducted open-ended discussion with PhD and master students on the reliability of LLMs for Code, resulting on two collaboration projects [O1, O4]

## Mentoring Experiences

---

During my PhD, I had the privilege of working with students/junior researchers.

- Anh Ho (PhD Student at The University of Melbourne): working on automated regression bug repair and co-authored a manuscript (under review) at TOSEM [O3].
- Huu Hung Nguyen (PhD Student at Singapore Management University): working on vulnerability analysis and co-authored two manuscript (under review) at TOSEM [O5] and ICSE [O2].
- Yen-Trang Dang (Undergraduate Student at HUST, now PhD Student at The University of Sydney): working on Github topic recommendation and co-authored a paper at EASE [C6].
- Duc-Manh Tran (Undergraduate Student at HUST, now PhD Student at The University of Sydney): working on vulnerability repair and co-authored a paper at ICSME [S4] and a manuscript [O6]
- Duong Nguyen (Undergraduate Student at HUST): working on just-in-time vulnerability prediction and co-authored two papers at ICSME [S4, C7].
- Hung Le (PhD Student at North Carolina State University): working on backdoor detection and co-authored a manuscript (under review) at FSE [O1].
- Duc-Minh Luong (Undergraduate Student at HUST): working on program repair and co-authored a paper at TSE [J3].

## Talks/Presentations

---

- |  |                                  |
|--|----------------------------------|
| • <b>Beyond Accuracy: A Closer Look on the Reliability of Large Language Models for Code</b> | SoftwareLab@CISPA                |
| Sep 2025   | Invited Talk                     |
| • <b>Toward Reliable LLM-assisted Programming</b>  | CREST@The University of Adelaide |
| June 2025  | Invited Talk                     |
| • <b>Repo-level Code Translation with LLMs and Automated Reasoning</b>                       | Amazon Web Services              |
| Sep 2024   | Invited Talk                     |
| • <b>Time-Aware Zero-Shot Identification of Libraries from Vulnerability Reports</b>         | ICSE 2023                        |
| May 2023   | Conference Presentation          |
| • <b>AutoPruner: Transformer-based Call Graph Pruning</b>                                    | ESEC/FSE 2022                    |
| November 2022  | Conference Presentation          |
| • <b>AutoPruner: Transformer-based Call Graph Pruning</b>                                    | University of Stuttgart          |
| December 2022  | Invited Tak                      |
| • <b>Toward Analysis of Graph Neural Network</b>   | Vietnam National University      |
| June 2022  | Invited Tak                      |

## Services

---

### • Journal Reviewers

- IEEE Transactions on Software Engineering (TSE),
- Empirical Software Engineering (EMSE),
- ACM Transactions on Software Engineering and Methodology (TOSEM),
- Information and Software Technology (IST),
- The Journal of Systems & Software (JSS),
- IEEE Transaction on Reliability (TOPS)
- Software Quality Journal (SQJ)
- Automated Software Engineering (ASEJ)

### • Program Committee

- FORGE 2026, Research Track
- SVM 2026, Research Track
- ICSE 2026, Technical Track (Shadow PC)
- ICSE 2025, Artifact Evaluation Track
- MSR 2025, Technical Track (Junior PC)
- EASE 2025, AI Models / Data Track
- SOICT 2023-2025, Software Engineering Track
- KSE 2022-2025, Software Engineering Track

## References

---

1. **Prof. Toby Murray**  
*Director, Defence Science Institute*  
*Professor, University of Melbourne*  
**Email:** [toby.murray@unimelb.edu.au](mailto:toby.murray@unimelb.edu.au)
2. **Dr. Bach Le**  
*ARC DECRA Fellow*  
*Senior Lecturer, University of Melbourne*  
**Email:** [bach.le@unimelb.edu.au](mailto:bach.le@unimelb.edu.au)
3. **Prof. David Lo**  
*IEEE Fellow, ACM Fellow, ASE Fellow*  
*Director, Research Center on Intelligent Software Engineering*  
*Professor, Singapore Management University*  
**Email:** [davidlo@smu.edu.sg](mailto:davidlo@smu.edu.sg)