

HONG JIN KANG

Lecturer (equivalent to Assistant Professor)
University of Sydney

hongjin.kang@sydney.edu.au
<https://kanghj.github.io>

RESEARCH INTERESTS

My research expertise lies in automated software engineering, with a focus on improving developer productivity. I work at the intersection of AI and software engineering, particularly on developing agentic AI methods to support human developers in program analysis and vulnerability management.

SELECTED EDUCATION AND PROFESSIONAL EXPERIENCE

Lecturer The University of Sydney	<i>2025 – current</i>
Postdoctoral Fellow University of California, Los Angeles	<i>2023 – 2024</i>
Ph.D. in Computer Science Singapore Management University	<i>2019 – 2023</i>
Bachelor of Computing (Highest Distinction/First-Class) National University of Singapore	<i>2012 – 2016</i>

SELECTED PUBLICATIONS

Think Like Human Developers: Harnessing Community Knowledge for Structured Code Reasoning

Chengran Yang, Zhensu Sun, **Hong Jin Kang**, Jieke Shi, David Lo
IEEE/ACM International Conference on Software Engineering (ICSE 2026)

Distinguished Paper Award

Let the Trial Begin: A Mock-Court Approach to Vulnerability Detection using LLM-Based Agents.

Ratnadira Widyasari, Martin Weyssow, Ivana Clairine Irsan, Han Wei Ang, Frank Liauw, Eng Lieh Ouh, Lwin Khin Shar, **Hong Jin Kang**, David Lo
IEEE/ACM International Conference on Software Engineering (ICSE 2026)

WhyFlow: Interrogative Debugger for Sensemaking Taint Analysis

Burak Yetiştiren, **Hong Jin Kang**, Miryung Kim
IEEE/ACM International Conference on Software Engineering (ICSE 2026)

TEACHING

Undergraduate Programme

As unit coordinator:

SOFT2201: Software Design and Construction I

As lecturer:

SOFT3202: Software Design and Construction II

Postgraduate Programme:

As unit coordinator:

COMP9201: Software Design and Construction I

As lecturer:

COMP9202: Software Design and Construction II

PEER-REVIEWED CONFERENCE/JOURNAL PUBLICATIONS

Let the Trial Begin: A Mock-Court Approach to Vulnerability Detection using LLM-Based Agents.

Ratnadira Widyasari, Martin Weyssow, Ivana Clairine Irsan, Han Wei Ang, Frank Liauw, Eng Lieh Ouh, Lwin Khin Shar, **Hong Jin Kang**, David Lo

IEEE/ACM International Conference on Software Engineering (ICSE 2026)

Think Like Human Developers: Harnessing Community Knowledge for Structured Code Reasoning

Chengran Yang, Zhensu Sun, **Hong Jin Kang**, Jieke Shi, David Lo

IEEE/ACM International Conference on Software Engineering (ICSE 2026)

Distinguished Paper Award

Back to the Basics: Rethinking Issue-Commit Linking with LLM-Assisted Retrieval.

Huihui Huang, Ratnadira Widyasari, Ting Zhang, Ivana Clairine Irsan, Jieke Shi, Han Wei Ang, Frank Liauw, Eng Lieh Ouh, Lwin Khin Shar, **Hong Jin Kang**, David Lo

IEEE/ACM International Conference on Software Engineering (ICSE 2026)

WhyFlow: Interrogative Debugger for Sensemaking Taint Analysis

Burak Yetiştiren, **Hong Jin Kang**, Miryung Kim

IEEE/ACM International Conference on Software Engineering (ICSE 2026)

Revisiting Vulnerability Patch Identification on Data in the Wild.

Ivana Clairine Irsan, Ratnadira Widyasari, Ting Zhang, Huihui Huang, Ferdian Thung, Yikun Li, Lwin Khin Shar, Eng Lieh Ouh, **Hong Jin Kang**, David Lo.

ACM Transactions on Software Engineering and Methodology (TOSEM) 2026

DuoReduce: Bug Isolation for Multi-Layer Extensible Compilation

Jiyuan Wang, Yuxin Qiu, Ben Limpanukorn, **Hong Jin Kang**, Qian Zhang, Miryung Kim

Foundations on Software Engineering (FSE 25)

Grey-Box Fuzzing in Constrained Ultra-Large Systems: Lessons for SE Community

Jiazhao Yu, Yanlun Tu, Zhanlei Zhang, Tiehua Zhang, Cheng Xu, Weigang Wu, **Hong Jin Kang**, Xi Zheng
Foundations on Software Engineering (FSE 25) Industry Track

Fuzzing MLIR Compilers with Custom Mutation Synthesis
Ben Limpanukorn, Jiyuan Wang, **Hong Jin Kang**, Eric Zitong Zhou, Miryung Kim
IEEE/ACM International Conference on Software Engineering (ICSE 2025)

Scaling Active Learning-based Code Search with Interactive What-If Analysis
Hong Jin Kang, Kevin Wang, Miryung Kim
IEEE/ACM International Conference on Software Engineering (ICSE 2024)

Evaluating SZZ Implementations: An Empirical Study on the Linux Kernel
Yunbo Lyu, **Hong Jin Kang**, Ratnadira Widyasari, Julia Lawall, David Lo
IEEE Transactions on Software Engineering (TSE) 2024

Human-in-the-Loop Synthetic Text Data Inspection with Provenance Tracking
Hong Jin Kang, Fabrice Harel-Canada, Muhammad Ali Gulzar, Violet Peng, Miryung Kim
Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024)

Surveying neuro-symbolic approaches for reliable artificial intelligence of things
Zhen Lu, Imran Afridi, **Hong Jin Kang**, Ivan Ruchkin, Xi Zheng
Journal of Reliable Intelligent Environment 2024

Smaller, Faster, Greener: Compressing Pre-trained Code Models via Surrogate-Assisted Optimization
Jieke Shi, Zhou Yang, **Hong Jin Kang**, Bowen Xu, Junda He, David Lo
IEEE/ACM International Conference on Software Engineering - Software Engineering in Practice (ICSE 2024- SEIP)

Stealthy backdoor attack for code models
Zhou Yang, Bowen Xu, Jie M Zhang, **Hong Jin Kang**, Jieke Shi, Junda He, David Lo
IEEE Transactions on Software Engineering (TSE) 2024
TSE Best Paper Award

Chronos: Time-Aware Zero-Shot Identification of Libraries from Vulnerability Reports
Yunbo Lyu, Thanh Le-Cong, **Hong Jin Kang**, Ratnadira Widyasari, Zhao Zhipeng, Bach Le, Ming Li, David Lo
IEEE/ACM International Conference on Software Engineering (ICSE 2023)

Topic Recommendation for GitHub Repositories: How Far Can Extreme Multi-Label Learning Go?
Ratnadira Widyasari, Zhipeng Zhao, Thanh Le Cong, Hong Jin Kang, David Lo
IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2023)

How to Find Actionable Static Analysis Warnings: A Case Study with FindBugs

Rahul Yedida, **Hong Jin Kang**, Huy Tu, Xueqi Yang, David Lo, Tim Menzies
IEEE Transactions on Software Engineering (TSE) 2023

Multi-Granularity Detector for Vulnerability Fixes

Truong Giang Nguyen, Thanh Le-Cong, **Hong Jin Kang**, Ratnadira Widyasari, Chengran Yang, Zhipeng Zhao, Bowen Xu, Jiayuan Zhou, Xin Xia, Ahmed E Hassan, Xuan-Bach D Le, David Lo

IEEE Transactions on Software Engineering (TSE) 2023

Compressing Pre-trained Models of Code into 3 MB

Jieke Shi, Zhou Yang, Bowen Xu, **Hong Jin Kang**, David Lo

IEEE/ACM International Conference on Automated Software Engineering (ASE 2022)

Detecting False Alarms from Automatic Static Analysis Tools: How Far are We?

Hong Jin Kang, Khai Loong Aw, and David Lo.

IEEE/ACM International Conference on Software Engineering (ICSE 2022)

Test Mimicry to Assess the Exploitability of Library Vulnerabilities.

Hong Jin Kang, Truong Giang Nguyen, Xuan-Bach D. Le, Corina Pasareanu, David Lo.

ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2022).

AutoPruner: Transformer-based Call Graph Pruning

Le-Cong Thanh, **Hong Jin Kang**, Truong Giang Nguyen, Stefanus Agus Haryono, David Lo, Xuan-Bach D. Le, Huynh Quyet Thang.

ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2022)

HERMES: Using Commit-Issue Linking to Detect Vulnerability-Fixing Commits

Truong Giang Nguyen, **Hong Jin Kang**, David Lo, Sharma Abhishek, Santosa Andrew, Sharma Asankhaya, and Ming Yi Ang.

IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2022)

Active Learning of Discriminative Subgraph Patterns for API Misuse Detection

Hong Jin Kang, David Lo.

IEEE Transactions on Software Engineering (TSE 2022)

Automated Identification of Libraries from Vulnerability Data: Can We Do Better?

Stefanus A. Haryono, **Hong Jin Kang**, Abhishek Sharma, Asankhaya Sharma, Andrew Santosa, Ang Ming Yi, David Lo

IEEE/ACM International Conference on Program Comprehension (ICPC 2022)

AndroEvolve: Automated Android API Update with Data Flow Analysis and Variable Denormalization

Stefanus A Haryono, Ferdian Thung, David Lo, Lingxiao Jiang, Julia Lawall, **Hong Jin Kang**, Lucas Serrano, Gilles Muller

Empirical Software Engineering (EMSE 2022)

Adversarial Specification Mining

Hong Jin Kang, David Lo.

ACM Transactions on Software Engineering and Methodology (TOSEM 2021)

IoTBox: Sandbox Mining to Prevent Interaction Threats in IoT Systems

Hong Jin Kang, Sheng Qin Sim, David Lo

IEEE International Conference on Software Testing, Verification and Validation 2021

BiasFinder: Metamorphic Test Generation to Uncover Bias for Sentiment Analysis Systems

Muhammad Hilmi Asyrofi, Zhou Yang, Imam Nur Bani Yusuf, **Hong Jin Kang**, Ferdian Thung, and David Lo

IEEE Transactions on Software Engineering (TSE 2021)

CC2Vec: Distributed representations of code changes,

Thong Hoang, **Hong Jin Kang**, Julia Lawall, David Lo.

ACM/IEEE International Conference on Software Engineering (ICSE 2020)

Towards Generating Transformation Rules without Examples for Android API Replacement

Ferdian Thung, **Hong Jin Kang**, Lingxiao Jiang, David Lo.

IEEE International Conference on Software Maintenance and Evolution (ICSME 2019)

Assessing the Generalizability of code2vec Token Embeddings

Hong Jin Kang, Tegawende F Bissyandé., David Lo.

ACM/IEEE International Conference on Automated Software Engineering (ASE 2019)

Semantic Patches for Java Program Transformation

Hong Jin Kang, Ferdian Thung, Julia Lawall, Gilles Muller, Lingxiao Jiang, David Lo.

European Conference on Object-Oriented Programming (ECOOP 2019)

SHORT/WORKSHOP/DEMO PAPERS

Abductive Reasoning for Neurosymbolic Fault Localization

Minh Tam Le, Xi Zheng, **Hong Jin Kang**

IEEE/ACM International Conference on Software Engineering - New Idea and Emerging Results (ICSE 2026 - NIER)

VulCurator: A Vulnerability-Fixing Commit Detector

by Truong Giang Nguyen, Thanh Le-Cong, **Hong Jin Kang**, and Bach Le, David Lo

ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2022) – Demonstrations

AndroEvolve: Automated Update for Android Deprecated-API Usages

by Stefanus Agus Haryono, Ferdian Thung, David Lo, Lingxiao Jiang, Julia Lawall, **Hong Jin Kang**, Lucas Serrano, Gilles Muller

International Conference on Software Engineering (ICSE 2021) – Demonstrations

BugsInPy: a Database of Existing Bugs in Python Programs to Enable Controlled Testing and Debugging Studies

Ratnadira Widyasari, Sheng Qin Sim, Camellia Lok, Haodi Qi, Jack Phan, Qijin Tay, Constance Tan, Fiona Wee, Jodie Ethelda Tan, Yuheng Yieh, Brian Goh, Ferdian Thung, **Hong Jin Kang**, Thong Hoang, David Li, Eng Lieh Ouh
ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2020) – Tool Demo

Automatic Android Deprecated-API Usage Update by Learning from Single Updated Example

Stefanus Agus Haryono, Ferdian Thung, **Hong Jin Kang**, Lucas Serrano, Gilles Muller, Julia Lawall, David Lo, Lingxiao Jiang.
International Conference on Program Comprehension (ICPC 2020) - Early Research Achievements

A Comparison of Word Embeddings for English and Cross-Lingual Chinese Word Sense Disambiguation

Hong Jin Kang, Tao Chen, Muthu Kumar Chandrasekaran, Min-Yen Kan.
Workshop on NLP Techniques for Educational Applications (NLP-TEA 2016)

SERVICE

JOURNAL REVIEWER:

- IEEE Transactions on Software Engineering (TSE),
- Empirical Software Engineering (EMSE),
- Automated Software Engineering (ASE),
- ACM Transactions on Software Engineering and Methodology (TOSEM),
- Information and Software Technology (IST),
- The Journal of Systems & Software (JSS),
- ACM Transactions on Privacy and Security (TOPS),
- Applied Intelligence (APIN),
- ACM Computing Surveys (CSUR),
- SCICO
- Guest Editor on a Special Issue of Empirical Software Engineering

PROGRAM COMMITTEE:

2026:

- FSE 2026 Industry Track
- FSE 2026 IVR Track
- ASE 2026
- LLMTrust 2026
- Humanise Workshop 2026
- RAIE 2026
- FSE 2026
- Co-Program Chair of InternetWare 2026 New Idea Track

2025:

- ASE 2025
- ICSE 2025

- ESEC/FSE 2025 Student Research Competition
- Humanise Workshop 2025
- RAIE (Responsible AI Engineering) 2025
- ISSTA 2025 Tool Demo
- ESEC/FSE 2025 Industry Track
- MSR 2025 – Data and Tool Showcase
- FORGE 2025 Industry Track
- LLM4Code 2025
- APSEC 2025
- ICSE 2025

2024:

- LLM4Code 2024
- SANER 2024
- ICSE 2024 AE
- SANER ERA 2023
- FSE 2023 IVR
- APSEC 2024
- Co-Program Chair of 4th edition of the International Workshop on Software Engineering and AI for Data Quality (SEA4QD 2024)

SUB-REVIEWER:

ICSME 2020, APSEC 2020, MSR 2020, ASE 2021, ICST 2022, ICSE 2022

MENTORING/SUPERVISION EXPERIENCE

As a lecturer, I had the privilege of supervising the following Honours/Masters students:

- | | |
|-----------------|----------------------------|
| • Le Minh Tam | masters (research pathway) |
| • Ta La | masters (MPhil) |
| • Yuhao Zang | masters (research pathway) |
| • Yue Wang | masters (research pathway) |
| • Steven Ni | honours student |
| • Apurva | honours student |
| • Shawn Li | honours student |
| • Aashna Munjal | honours student |

As a postdoc/PhD student, I had the privilege of working with students/junior researchers.

At UCLA,

- | | |
|--------------------|-----------------|
| • Kevin Wang | masters student |
| • Burak Yetistiren | PhD student |

At SMU,

- | | |
|-----------------------|-----------------------|
| • Aw Khai Loong | undergraduate student |
| • Sim Sheng Qin | undergraduate student |
| • Jonathan Ooi | masters student |
| • Ratnadira Widyasari | PhD student |

- Le Cong Thanh research engineer, started a PhD in Jan 2023
- Lyu Yunbo research engineer, started a PhD in Aug 2023
- Truong Giang Nyugen research engineer
- Stefanus Agus Haryono research engineer

AWARDS

2022 SMU Dean's List
 2020, 2021, 2022 SMU's Presidential Doctoral Fellowship
 Distinguished Reviewer Award at ICSE 25
 TSE Best Paper Award 2024
 Distinguished Paper Award 2026

WORK EXPERIENCE

AUG 2018 – DEC 2018

RESEARCH ENGINEER, SMU

- Developed Coccinelle4J, a port of Coccinelle for the Java programming language. Coccinelle is a widely adopted program matching and transformation tool for C systems software, including the Linux kernel. Over 6000 commits have been in the Linux kernel with the help of Coccinelle in the past 10 years. We apply Coccinelle4J to the automatic migration of deprecated Android API usages.

2018

BACKEND ENGINEER, GRAB

- Worked with cutting-edge tools and managed cloud-based infrastructure.
- Owned a microservice on the critical path of all bookings. Improved testing to identify single points of failure, improving the resilience of the service to ensure uptime and SLAs are met.

2016 –2018

SOFTWARE ENGINEER, WORKS APPLICATIONS CO. LTD

- Worked on enterprise-grade HR Enterprise Resource Management software.
- Full-stack development from the back-end using Spring, management of Cassandra, Elasticsearch, to implementing a new user experience on the front-end with modern frameworks.

AUG 2015 – MAY 2016

STUDENT SOFTWARE ENGINEER, TEAMMATES @ NUS

- Involved in both planning and development of the project. Planned and managed weekly releases.
- Part of the team that introduced static analysis tools and CI to the project.

MAY 2014 – AUG 2014

INTERN, TEAMIE

- Developed several new features, including a feature that summarizes users' activity logs to identify highlights of their history on the platform.

REFERENCES

Prof. David Lo
Full Professor, Singapore Management University
davidlo@smu.edu.sg

Prof. Miryung Kim
Full Professor, University of California, Los Angeles (UCLA)
miryung@cs.ucla.edu

Dr. Julia Lawall
Senior Research Scientist, Inria/LIP6
julia.lawall@inria.fr