Yining Hong

✉ yhong3@andrew.cmu.edu 🔗 hyn0027.github.io in hyn0027 hyn0027

Education

Carnegie Mellon University *Aug. 2024 - May 2029 (Expected)*
Ph.D. in Software Engineering

- **CQPA:** 4.11/4.00
- **Advisors:** Prof. Christian Kästner and Dr. Chris Timperley
- **Research Interest:** Safety/security for software with ML/Agent components; AI for software safety/security

Tsinghua University *Sept. 2020 - July 2024*
B.Eng. in Computer Science and B.Ec. in Economics & Finance

- **CGPA:** 3.91/4.00 , **Rank:** 2/12
- **Thesis (Honored):** A Reinforcement Learning Framework for Training and Testing in Asset Portfolio Optimization
- **Advisors:** Prof. Maosong Sun and Prof. Junliang Xing

Internships

Software Engineer Intern *Beijing, China*
Advanced Micro Devices (AMD) *Jan. 2024 – June 2024*

- Optimized FPGA-based deep learning compiler to improve runtime performance.
- Applied operational research methods to optimize storage unit allocation.

Software Engineer Intern *Beijing, China*
Kuangshi Technology (Megvii) *Sept. 2023 – Dec. 2023*

- Integrated machine learning models into autonomous driving systems.
- Designed and implemented guardrails for object trajectory prediction models.

Student Research Intern *Pittsburgh, PA*
Carnegie Mellon University *June. 2023 – Sept. 2023*

- Conducted research on semantic data slicing for capability testing in language models.
- Co-advised by Prof. Christian Kästner and Prof. Sherry Tongshuang Wu.
- Co-authored a paper accepted at ASE 2024.

Publications

[1] **Towards Verifiably Safe Tool Use for LLM Agents**
Aarya Doshi, **Yining Hong**, Congying Xu, Eunsuk Kang, Alexandros Kapravelos, Christian Kästner.
International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER), 2026.

[2] **From Hazard Identification to Controller Design: Proactive and LLM-Supported Safety Engineering for ML-Powered Systems**
Yining Hong, Christopher S. Timperley, Christian Kästner.
International Conference on AI Engineering (CAIN), 2025.

[3] **What Is Wrong with My Model? Identifying Systematic Problems with Semantic Data Slicing**
Chenyang Yang, **Yining Hong**, Grace Lewis, Tongshuang Wu, Christian Kästner.
International Conference on Automated Software Engineering (ASE), 2024.

[4] **Two Heads Are Better Than One: Exploiting Both Sequence and Graph Models in AMR-To-Text Generation**
Yining Hong, Fanchao Qi, Maosong Sun.

On-going Research

Guaranteeing safety/security for LLM-based Agents

- Agents connect with tools that could be malicious or vulnerable. We explore symbolic mechanisms, including API security practices, information flow analysis, and temporal logic, to provide safety/security guarantees.

Monitoring for risk management for software systems with ML components

- ML components in software introduce new risks, requiring new methods of monitoring. This interview study explores how practitioners monitor subtle risks and how interventions change their mental model.

Projects

Hazard Finder *github.com/hyn0027/Hazard-Finder*

- Developed an LLM-powered tool to assist safety engineers in performing hazard analysis (specifically STPA) for ML-enabled software and AI agentic systems, supporting systematic safety/security risk management.

OpenAI Chat Helper *github.com/hyn0027/OpenAIChatHelper*

- Developed and maintained a Python package for managing asynchronous OpenAI Chat API interactions with strong data typing, prompt templating, markdown formatting, and streamlined workflow support.

RL Portfolio Optimization (Bachelor's Thesis) *github.com/hyn0027/RL-Portfolio-Optimization*

- Developed a unified Python framework for implementing and testing (deep) reinforcement learning methods for financial asset portfolio optimization.

Relational Database Management System *github.com/hyn0027/DataBase-System*

- Developed a complete relational database management system in C++ from the ground up, reusing only ANTLR parser generator.

Skills

Programming Languages: C/C++, Python, JavaScript/TypeScript, Rust, Solidity

Safety & Security: Threat Modeling, STRIDE, STPA, FMEA, HAZOP, FTA

DevOps/MLOps: Docker, Git, MLflow, Grafana, Prometheus, GitHub/GitLab CI/CD, Jenkins, Make/CMake

Web Development: FastAPI, React, Next.js, Vue.js, HTML/CSS, Django

ML: MCP, PyTorch, NumPy, Pandas, Hugging Face

Other: MySQL, LaTeX, Verilog/SystemVerilog, VHDL, Vivado, RISC-V Assembly, Qt, Cocos

Awards and Honers

Outstanding Graduate , Computer Science Department, Tsinghua University *July 2024*

Comprehensive Excellence Scholarship (top 10%) , Tsinghua University *Nov. 2023*

Citadel Securities Scholarship , Citadel Securities *Dec. 2022*

Comprehensive Excellence Scholarship (top 10%) , Tsinghua University *Nov. 2022*

Academic Excellence Scholarship (top 30%) , Tsinghua University *Nov. 2021*

Sports Excellence Scholarship , Tsinghua University *Nov. 2021*

First Prize , National Olympiad in Informatics in Provinces *Nov. 2018 & Nov. 2017*

Teaching & Mentoring

Volunteer Teaching Assistant (Outstanding Level) *Mar. 2022 - July 2024*
Tsinghua University

- Offered drop-in tutoring for computer science and finance courses.

Aarya Doshi (mentoring) *May 2025- Aug. 2025*
Undergraduate student from Georgia Institute of Technology

- Explored safety analysis and control for AI agents.

Abhishek Satpathy (mentoring) *May 2025- Aug. 2025*
Undergraduate student from University of Virginia

- Explored threat modeling for ML-enabled systems.