Zhaoyang Chu

zhaoyang.chu.25@ucl.ac.uk | zhaoyang-chu.github.io

Research Interest

My research interest focuses on the intersection of **software engineering** and **machine learning**, aimed at building next-generation AI systems that make code smarter, more trustworthy, and more accessible.

Education

University College London
Ph.D. in Computer Science — *Sep 2025 – Sep 2028 (anticipated)*

Huazhong University of Science and Technology — GPA: 3.53/4.0
*M.E. in Computer Science and Technology (**Graduated with Honors**)* — *Sep 2022 – Jun 2025*

Huazhong Agricultural University — GPA: **3.93**/4.0
*B.E. in Data Science and Big Data Technology (**Graduated with Honors**)* — *Sep 2018 – Jun 2022*

Research Papers

** indicates equal contribution. † indicates the corresponding author.*

[1] Scrub It Out! Erasing Sensitive Memorization in Code Language Models via Machine Unlearning.
***Zhaoyang Chu**, Yao Wan†, Zhikun Zhang, Di Wang, Zhou Yang, Hongyu Zhang, Pan Zhou, Xuanhua Shi, Hai Jin, David Lo.*
ICSE 2026, **CORE A***.

[2] CODESYNC: Synchronizing Large Language Models with Dynamic Code Evolution at Scale.
Chenlong Wang, **Zhaoyang Chu***, Zhengxiang Cheng*, Xuyi Yang, Kaiyue Qiu, Yao Wan†, Zhou Zhao, Xuanhua Shi, Dongping Chen.*
ICML 2025, **CORE A***.

[3] Graph Neural Networks for Vulnerability Detection: A Counterfactual Explanation.
***Zhaoyang Chu**, Yao Wan†, Qian Li, Yang Wu, Hongyu Zhang, Yulei Sui, Guandong Xu, Hai Jin.*
ISSTA 2024, **CORE A**.

[4] How to Select Pre-Trained Code Models for Reuse? A Learning Perspective.
*Zhangqian Bi, Yao Wan†, **Zhaoyang Chu**, Yufei Hu, Junyi Zhang, Hongyu Zhang, Guandong Xu, Hai Jin.*
SANER 2025, **CORE A**.
IEEE TCSE Distinguished Paper Award.

[5] Can Large Language Models Serve as Evaluators for Code Summarization?
*Yang Wu, Yao Wan†, **Zhaoyang Chu**, Wenting Zhao, Ye Liu, Hongyu Zhang, Xuanhua Shi, Philip S. Yu.*
IEEE Transactions on Software Engineering (TSE), **CORE A***, **Impact Factor 6.5**.

[6] Wait, We Don't Need to "Wait"! Removing Thinking Tokens Improves Reasoning Efficiency.
*Chenlong Wang, Yuanning Feng, Dongping Chen, **Zhaoyang Chu**, Ranjay Krishna, Tianyi Zhou†.*
EMNLP 2025 Findings, **CORE A***.

[7] TESTEVAL: Benchmarking Large Language Models for Test Case Generation.
Wenhan Wang, Chenyuan Yang*, Zhijie Wang*, Yuheng Huang, **Zhaoyang Chu**, Da Song, Lingming Zhang, An Ran Chen, Lei Ma.*
NAACL 2025 Findings, **CORE A**.

[8] Hierarchical Graph Representation Learning for the Prediction of Drug-Target Binding Affinity.
Zhaoyang Chu, Feng Huang*, Haitao Fu, Yuan Quan, Xionghui Zhou, Shichao Liu, Wen Zhang†.*
Information Sciences, **CORE A**, **Impact Factor 8.1**.

Research Experience

Research Intern May 2024 – Jun 2025
University of Illinois Urbana-Champaign, advised by Prof. Lingming Zhang

Test Case Generation (NAACL 2025 Findings)
- *Propose a novel benchmark that evaluates LLMs' capabilities in generating test cases for Python programs.*

Compiler Fuzzing (Ongoing)
- *Develop an LLM-based mutation fuzzing framework for LLVM compilers with integrated domain knowledge.*

Master Research Assistant Sep 2022 – Jun 2025
Huazhong University of Science and Technology, advised by Prof. Yao Wan

Machine Unlearning for code LLMs (ICSE 2026)
- *Develop a privacy-preserving method to erase sensitive information from code LLMs via machine unlearning.*
- *Collaborated with **Prof. David Lo at Singapore Management University**.*

LLM Synchronization with Code Evolution (ICML 2025)
- *Propose a novel benchmark to evaluate LLMs' synchronization with real-time library API updates.*
- *Collaborated with **Prof. Zhou Zhao at Zhejiang University**.*

Counterfactual Vulnerability Detection (ISSTA 2024)
- *Design a counterfactual explainer to uncover the decision mechanisms of GNN-based detection systems.*
- *Collaborated with **Prof. Yulei Sui at University of New South Wales**.*

Code LLM Selection (SANER 2025 Distinguished Paper)
- *Propose learning-based methods for efficiently selecting and reusing pre-trained code LLMs for target software engineering tasks within limited computational budgets.*
- *Collaborated with **Prof. Hongyu Zhang at Chongqing University**.*

LLM-as-a-Judge for Code Summarization (TSE)
- *Develop an LLM-based evaluator to asses the quality of code summaries generated by neural models.*
- *Collaborated with **Prof. Philip S. Yu at University of Illinons at Chicago** .*

Efficient Reasoning for LLMs (EMNLP 2025 Findings)
- *Propose a novel approach that disables explicit self-reflection by suppressing "Wait"-like tokens during inference.*
- *Collaborated with **Prof. Ranjay Krishna at University of Washington**.*

Counterfactual Prompting for Code Generation (ACL 2025 Submission)
- *Design a counterfactual explainer to guide developers in assessing and refining prompt design for code generation.*

Undergraduate Research Assistant Sep 2018 – Jun 2022
Huazhong Agricultural University, advised by Prof. Wen Zhang

Drug-Target Binding Affinity Prediction (Information Sciences)
- *Build a hierarchical GNN to integrate coarse- and fine-grained information from affinity graphs and molecular graphs in a coarse-to-fine manner.*

Honors & Awards

IEEE TCSE Distinguished Paper Award 2025
National Scholarship for Masters (Top 1%) 2024
National Scholarship for Undergraduates (Top 1%) 2020