

Zhihan Guo

PH.D. CANDIDATE · COMPUTER SCIENCE AND ENGINEERING

The University of Hong Kong, Hong Kong SAR

✉ zhguo22@cse.cuhk.edu.hk | 🏠 zhihan-guo.github.io | 🎓 Google Scholar | ☎ 0000-0002-1889-583X

Research Interest

My research focuses on resource-constrained domain-specific federated LLM systems. I study two coupled classes of bottlenecks: data constraints (learning from distributed private data and constructing targeted supervision without gold responses) and computational resource constraints (communication-efficient federated adaptation and decision-centric memory under bounded inference budgets). Across these themes, I develop methods for natural language processing, multilingual and multimodal foundation models, graph-enhanced learning, and long-horizon agents. I am particularly interested in extending this framework to multimodal intent understanding, knowledge-guided issue discovery, and industrial domain-AI applications.

Education

PhD in Computer Science and Engineering

THE CHINESE UNIVERSITY OF HONG KONG

• Advisor: Prof. Irwin King (IEEE, ACM, and AAAI Fellow; Pro-Vice-Chancellor of CUHK)

Hong Kong SAR

2022 - 2027

MSc in Financial Technology

THE UNIVERSITY OF HONG KONG

Hong Kong SAR

2021 - 2022

BSc in Computer Science and Technology

BEIJING INSTITUTE OF TECHNOLOGY

- GPA: 3.8/4.0
- Honors thesis/undergrad research advisor: Prof. Zhendong Niu

Beijing, China

2017 - 2021

Publications

PUBLISHED

A Survey of Audio Reasoning in Audio Foundation Models

Zhihan Guo, Wenqian Cui, Guan-Ting Lin, Daxin Tan, Dingdong Wang, Qiyong Zheng, Jingyao Li, Jing Xiong, Jiaya Jia, Irwin King

The 2026 Conference on Empirical Methods in Natural Language Processing (EMNLP 2026)

[PDF] From General Reward to Targeted Reward: Improving Open-ended Long-context Generation Models

Zhihan Guo, Jiele Wu, Wenqian Cui, Yifei Zhang, Minda Hu, Yufei Wang, Irwin King

The 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP 2025)

[PDF] FedLFC: Towards Efficient Federated Multilingual Modeling with LoRA-based Language Family Clustering

Zhihan Guo, Yifei Zhang, Zhuo Zhang, Zenglin Xu, Irwin King

2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024)

[PDF] FedScholar: Federated Scholarly Clustering with Literature-Aligned Rewards for Academic Report Generation

Zhihan Guo, Yukun Zhang, Zixing Song, Zenglin Xu, Irwin King

ACM Web Conference 2026 (WWW'26), Companion Proceedings

[PDF] FedHLT: Efficient Federated Low-Rank Adaption with Hierarchical Language Tree for Multilingual Modeling

Zhihan Guo, Yifei Zhang, Zhuo Zhang, Zenglin Xu, Irwin King

ACM Web Conference 2024 (WWW'24), Companion Proceedings

[PDF] Remember the Decision, Not the Description: A Rate-Distortion Framework for Agent Memory

Mingxi Zou, **Zhihan Guo**, Irwin King, Zenglin Xu

The Fortieth Conference on Neural Information Processing Systems (NeurIPS 2026)

[PDF] TurnGuide: Enhancing Meaningful Full Duplex Spoken Interactions via Dynamic Turn-Level Text-Speech Interleaving

Wenqian Cui, Lei Zhu, Xiaohui Li, **Zhihan Guo**, Haoli Bai, Lu Hou, Irwin King

Interspeech 2026

IN REVIEW

[PDF] Beyond Retrieval: A Modular Benchmark for Academic Deep Research Agents

[NAACL 2027 under review]

Zhihan Guo, Feiyang Xu, Yifan Li, Muzhi Li, Shuai Zou, Jiele Wu, Han Shi, Haoli Bai, Ho-fung Leung, Irwin King

A Multi-agent AI System Maps the Carbon-Strength Frontier of Carbon-storing Construction Materials

[Nature Communications of AI & Computing under review]

Shuai Zou*, **Zhihan Guo***, Yankai Chen, Jianzhuang Xiao, Chi Sun Poon, Philip S Yu, Irwin King

[PDF] A Survey of Audio Reasoning in Multimodal Foundation Models

[Planned Submission to TPAMI]

Zhihan Guo, Wenqian Cui, Guan-Ting Lin, Daxin Tan, Dingdong Wang, Qiyong Zheng, Jingyao Li, Jing Xiong, Jiaya Jia, Irwin King

Cognitive Alignment in Interactive Environments: A Survey of LLM-Driven Game Agents

[NAACL 2027 under review]

Yanyu CHEN, Dianzhi Yu, Hongzhan Lin, **Zhihan Guo**, Jiahong Liu, Jiyue Jiang, Zheng Wu, Xuhang Zhu, Hao Chen, Jiaming Han, Xianting Huang, Jinhu Qi, Xiao Guo, Yu Li, Ningtao Wang, Xiangyu Yue, Yifei Zhang, Irwin King

Aligning Graph Neural Networks and Large Language Models with Curriculum Learning on Text-Attributed Graphs

[CIKM 2026 under review]

Zixing Song, Yifan Li, **Zhihan Guo**, Prayag Tiwari, Irwin King

[PDF] Hierarchical Molecular Representation Learning via Fragment-Based Self-Supervised Embedding Prediction

[NAACL 2026 under review]

Jiele Wu, Haozhe Ma, **Zhihan Guo**, Thanh Vinh Vo, Tze-Yun Leong

[PDF] What, how, where, and how well? A survey on test-time scaling in large language models

[Planned Submission to TPAMI]

Qiyuan Zhang, Fuyuan Lyu, Zexu Sun, Lei Wang, Weixu Zhang, **Zhihan Guo**, Yufei Wang, Irwin King, Xue Liu, Chen Ma

From grey wastes to green mines: A white carbon cycle for future cities

[Chem Circularity, Cell Press under review; currently under external peer review]

Jianzhuang Xiao, Yuan Gao, Shuai Zou, Taohua Ye, **Zhihan Guo**, Guansheng Han, Jiqing Zhang, Jiajian Yu, Yanming Liu, Dianchao Wang, Takafumi Noguchi, Surendra P. Shah

Teaching Experience

Spring 2024 **CSCI2100 Data Structures**, Teaching Assistant

Spring 2024 **ESTR2102 Data Structures**, Teaching Assistant

Professional Experience

Program Chair: ACM International Conference on Web Search and Data Mining (WSDM) 2026

Conference Reviewer: Association for Computational Linguistics (ACL) 2026

Journal Reviewer: IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI); ACM Computing Surveys; IEEE Transactions on Neural Networks and Learning Systems (TNNLS); IEEE Transactions on Fuzzy Systems (TFS); ACM Transactions on Intelligent Systems and Technology (TIST).

Honors

July 2026 **Outstanding Student Paper Award**, FL@FM, The Web Conference

April 2026 **PhD IMPAC Award**, The Chinese University of Hong Kong

Invited Talks

“Efficient Federated Multilingual Modeling”, WeBank Co., Ltd.

“LLM-based Agent in Material Deep Search and Generation”, Institute of Automation of the Chinese Academy of Sciences.

“Long Text Generation of Large Language Model”, Huawei.

“From Audio Perception to Audio Reasoning”, Tencent.

“Federated LLM System”, Fudan University.

“Toward green in the LLM era”, Dongbei University.