Dingjie James Song

Contact Information

Lehigh University
Bethlehem
Pennsylvania
USA

Email
Home Page
GitHub
Google Scholar

Education

Ph.D. in Computer Science, Lehigh University, Bethlehem, PA, USA **January 2025–Present**

Advisor: Dr. Lichao Sun

Master of Software Engineering, Nanjing University, Nanjing, China **June 2023**

Advisors: Dr. Xinyu Dai, Dr. Jidong Ge

GPA: 4.4/5.0

Bachelor of Software Engineering, Nanjing University, Nanjing, China **June 2021**

Advisors: Dr. Xinyu Dai, Dr. Jidong Ge

GPA: 4.3/5.0

Research Interests

My research is in large language models and natural language processing, currently focused on LLM agents and AI for scientific discovery: building agents that learn and self-evolve in open-ended environments, and AI systems that automate and accelerate the research process end to end. This builds on my earlier work making multimodal and medical LLMs efficient, robust, and trustworthy. Prior to my current focus, I established a strong foundational expertise in task-oriented dialogue systems within the broader field of Natural Language Processing (NLP).

Submitted Publications

28. **Dingjie Song**, Zhiling Yan, Ziqian Zhang, Yuxuan Zhang, Lichao Sun. (2026). *PremiseCheck: Efficient Fact-Checking of Human and AI Peer Reviews.* [arXiv:2610.21046]. *Preprint.*
27. **Dingjie Song***, Zhiling Yan* (equal contribution), Yuxuan Zhang, Lichao Sun. (2026). *LEV: Thinking Fast and Deep with Looped Decision Models.* [arXiv:2610.20951]. *Preprint.*
26. Hongtao Wu*, Boyun Zheng*, **Dingjie Song*** (equal contribution), Yu Jiang, Jianfeng Gao, Lei Xing, Lichao Sun, Yixuan Yuan. (2026). *Towards a Medical AI Scientist.* [arXiv:2603.28589]. *Preprint.*
25. Xianghong Fang, Litao Guo, Hengchao Chen, Yuxuan Zhang, Xiaofan Xia, **Dingjie Song**, Yexin Liu, Hao Wang, Harry Yang, Qiang Sun, Yuan Yuan. (2026). *Distributional Matching for Vector Quantization: A Unified Theoretical and Empirical Framework.* [arXiv:2607.15933]. *Preprint.*
24. Guiyao Tie, Jiawen Shi, **Dingjie Song**, Yixiao Huang, Ziji Sheng, Xueyang Zhou, Daizong Liu, Pan Zhou, Yongchao Chen, Ran Xu, Lifang He, Qingsong Wen, Manling Li, Cong Lu, Shuai Li, Pengtao Xie, Yixuan Yuan, Rui Meng, Lei Xing, Lichao Sun, Caiming Xiong, Philip S. Yu, Jianfeng Gao. (2026). *AutoResearch AI: Towards AI-Powered Research Automation for Scientific Discovery.* [arXiv:2605.23204]. *Preprint.*
23. Mingzhe Zheng, **Dingjie Song**, Guanyu Zhou, Jun You, Jiahao Zhan, Xuran Ma, Xinyuan Song, Ser-Nam Lim, Qifeng Chen, Harry Yang. (2025) *CML-Bench: A Framework for Evaluating and Enhancing LLM-Powered Movie Scripts Generation.* [arXiv:2510.06231]. *Preprint.*
22. Zhejian Yang, Yongchao Chen, Xueyang Zhou, Jiangyue Yan, **Dingjie Song**, Yinuo Liu, Yuting Li, Yu Zhang, Pan Zhou, Hechang Chen, Lichao Sun. (2025) *Agentic Robot: A Brain-Inspired Framework for Vision-Language-Action Models in Embodied Agents.* [arXiv:2505.23450]. *Preprint.*

21. Tao Yu, Yi-Fan Zhang, Chaoyou Fu, Junkang Wu, Jinda Lu, Kun Wang, Xingyu Lu, Yunhang Shen, Guibin Zhang, **Dingjie Song**, Yibo Yan, Tianlong Xu, Qingsong Wen, Zhang Zhang, Yan Huang, Liang Wang, Tieniu Tan. (2025). *Aligning Multimodal LLM with Human Preference: A Survey.* [arXiv:2503.14504]. *Preprint.*
20. Guiyao Tie, Zeli Zhao, **Dingjie Song**, Fuyang Wei, Rong Zhou, Yurou Dai, Wen Yin, Zhejian Yang, Jiangyue Yan, Yao Su, Zhenhan Dai, Yifeng Xie, Yihan Cao, Lichao Sun, Pan Zhou, Lifang He, Hechang Chen, Yu Zhang, Qingsong Wen, Tianming Liu, Neil Zhenqiang Gong, Jiliang Tang, Caiming Xiong, Heng Ji, Philip S. Yu, Jianfeng Gao (2025). *A Survey on Post-training of Large Language Models.* [arXiv:2503.06072]. *Preprint.*
19. Yi-Fan Zhang, Hang Li, **Dingjie Song**, Lichao Sun, Tianlong Xu, Qingsong Wen. (2025). *From Correctness to Comprehension: AI Agents for Personalized Error Diagnosis in Education.* [arXiv:2502.13789]. *Preprint.*
18. Yuhao Du, Shunian Chen, Wenbo Zan, Peizhao Li, Mingxuan Wang, **Dingjie Song**, Bo Li, Yan Hu, Benyou Wang. (2024). *BlenderLLM: Training Large Language Models for Computer-Aided Design with Self-improvement.* [arXiv:2412.14203]. *Preprint.*

Accepted Publications

17. Zhiling Yan*, **Dingjie Song*** (equal contribution), Hanrong Zhang, Wei Liang, Yuxuan Zhang, Yutong Dai, Lifang He, Philip S. Yu, Ran Xu, Xiang Li, Lichao Sun. (2026). *OpenSkill: Open-World Self-Evolution for LLM Agents.* [arXiv:2606.06741].
EMNLP 2026
16. **Dingjie Song**, Hanrong Zhang, Dawei Liu, Yixin Liu, Zongxia Li, Zhengqing Yuan, Siqi Zhang, Henry Peng Zou, Zhiling Yan, Yuxuan Zhang, Yanfang Ye, Philip S. Yu, Lichao Sun. (2026). *Dr. Claw: An AI Scientist Workspace for Vibe Research.* [arXiv:2609.00365].
EMNLP 2026, System Demonstrations Track
15. **Dingjie Song**, Tianlong Xu, Yi-Fan Zhang, Hang Li, Zhiling Yan, Xing Fan, Haoyang Li, Lichao Sun, Qingsong Wen. (2026). *Can MLLMs Read Students' Minds? Unpacking Multimodal Error Analysis in Handwritten Math.* [arXiv:2603.24961].
AIED 2026 Oral
14. Xidong Wang*, **Dingjie Song*** (equal contribution), Shunian Chen, Junying Chen, Zhenyang Cai, Chen Zhang, Lichao Sun, Benyou Wang. (2024). *LongLLaVA: Scaling Multi-modal LLMs to 1000 Images Efficiently via Hybrid Architecture.* [arXiv:2409.02889].
EMNLP Findings 2025
13. **Dingjie Song***, Sicheng Lai* (equal contribution), Mingxuan Wang, Shunian Chen, Lichao Sun, Benyou Wang. (2024). *Both Text and Images Leaked! A Systematic Analysis of Multimodal LLM Data Contamination.* [arXiv:2411.03823].
EMNLP Findings 2025, ICML 2025 DIG-BUG Workshop Oral
12. **Dingjie Song**, Wenjun Wang, Shunian Chen, Xidong Wang, Michael Guan, Benyou Wang. (2024). *Less is More: A Simple yet Effective Token Reduction Method for Efficient Multi-modal LLMs.* [arXiv:2409.10994]
COLING 2025
11. **Dingjie Song**, Shunian Chen, Guiming Hardy Chen, Fei Yu, Xiang Wan, Benyou Wang. (2024). *MileBench: Benchmarking MLLMs in Long Context.* [arXiv:2404.18532]
COLM 2024
10. Xidong Wang*, Guiming Hardy Chen*, **Dingjie Song*** (equal contribution), Zhiyi Zhang, Zhihong Chen, Qingying Xiao, Feng Jiang, Jianquan Li, Xiang Wan, Benyou Wang, Haizhou Li. (2024) *CMB: A Comprehensive Medical Benchmark in Chinese.* [arXiv:2308.08833]
NAACL 2024

9. Pengfei Sun*, **Dingjie Song*** (equal contribution), Yawen Ouyang, Zhen Wu, Xinyu Dai. (2023) *Episode-based Prompt Learning for Any-shot Intent Detection.* Springer
NLPCC 2023 Oral

8. Yuxuan Zhang, Yubo Wang, Yipeng Zhu, Penghui Du, Junwen Miao, Xuan Lu, Zhuofeng Li, Xingwei Qu, Zhengkang Guo, Yuanzhe Shen, **Dingjie Song**, Han Zhou, Tuney Zheng, Xian Wu, Hao Yu, Songcheng Cai, Yi Lu, Yunzhuo Hao, Minyi Lei, Liang Chen, Kai Zou, Huifeng Yin, Wendong Xu, Dongfu Jiang, Ping Nie, Jiaheng Liu, Wenhu Chen, Kelsey R. Allen. (2026). *ClawBench: Can AI Agents Complete Everyday Online Tasks?.* [arXiv:2604.08523].
EMNLP Findings 2026

7. Zhiling Yan, **Dingjie Song**, Zhe Fang, Yisheng Ji, Xiang Li, Quanzheng Li, Lichao Sun. (2026). *LiveMedBench: A Contamination-Free Medical Benchmark for LLMs with Automated Rubric Evaluation.* [arXiv:2602.10367].
KDD 2026, Datasets and Benchmarks Track

6. Zhiling Yan, Sifan Song, **Dingjie Song**, Yiwei Li, Rong Zhou, Weixiang Sun, Zhennong Chen, Sekeun Kim, Hui Ren, Tianming Liu, Quanzheng Li, Xiang Li, Lifang He, Lichao Sun. (2025). *SAMed-2: Selective Memory Enhanced Medical Segment Anything Model.* [arXiv:2507.03698].
MICCAI 2025

5. Zhenyang Cai, Junying Chen, Rongsheng Wang, Weihong Wang, Yonglin Deng, **Dingjie Song**, Yize Chen, Zixu Zhang, Benyou Wang. (2024). *On the Compositional Generalization of Multimodal LLMs for Medical Imaging.* [arXiv:2412.20070].
ACL 2025

4. Wentao Ge*, Shunian Chen*, Guiming Hardy Chen* (equal contribution), Nuo Chen, Junying Chen, Zhihong Chen, Wenya Xie, Shuo Yan, Chenghao Zhu, Ziyue Lin, **Dingjie Song**, Xidong Wang, Anningzhe Gao, Zhiyi Zhang, Jianquan Li, Xiang Wan, Benyou Wang. (2023). *MLLM-Bench: Evaluating Multimodal LLMs with Per-sample Criteria.* [arXiv:2311.13951]
NAACL 2025

3. Junying Chen, Xidong Wang, Anningzhe Gao, Feng Jiang, Shunian Chen, Hongbo Zhang, **Dingjie Song**, Wenya Xie, Chuyi Kong, Jianquan Li, Xiang Wan, Haizhou Li, Benyou Wang. (2023). *HuatuoGPT-II, One-stage Training for Medical Adaption of LLMs.* [arXiv:2311.09774]
COLM 2024

2. Huang Huang*, Fei Yu*, Jianqing Zhu* (equal contribution), Xuening Sun, Hao Cheng, **Dingjie Song**, Zhihong Chen, Abdulmohsen Alharthi, Bang An, Juncai He, Ziche Liu, Zhiyi Zhang, Junying Chen, Jianquan Li, Benyou Wang, Lian Zhang, Ruoyu Sun, Xiang Wan, Haizhou Li, Jinchao Xu. (2024) *AceGPT, Localizing Large Language Models in Arabic.* [arXiv:2309.12053]
NAACL 2024

1. Pengfei Sun, Yawen Ouyang, **Dingjie Song**, Xinyu Dai. (2022) *Self-Supervised Task Augmentation for Few-Shot Intent Detection.* Springer
JCST 2022

Research Experience

Research Intern, Salesforce AI Research, Palo Alto, CA **June 2026–August 2026**
Focus: LLM agents and AI for scientific discovery.

Research Intern, Squirrel AI Learning **October 2024–January 2025**
Mentor: Dr. Tianlong Xu and Dr. Qingsong Wen
Focus: Multimodal LLM and Math LLM.

Research Assistant, The Chinese University of Hong Kong, Shenzhen **July 2023–September 2024**
Mentor: Dr. Benyou Wang
Focus: Multimodal LLM, Medical LLM, and Multilingual LLM.

NLP Research Intern, International Digital Economy Academy (IDEA) **July 2022–May 2023**
Mentor: Dr. Jiaxing Zhang
Focus: In-context learning and Code LLM.

Graduate Student, NLP Group, Nanjing University **August 2021–June 2022**
Advisor: Dr. Xinyu Dai
Focus: Low-resource spoken language understanding.

Teaching Experience

Teaching Assistant, Lehigh University, Bethlehem, PA **Fall 2026**
CSE 326/426 Fundamentals of Machine Learning
CSE 325/425 Natural Language Processing

Talks

Dr. Claw: One AI Workspace for the Whole Research Process **November 2026**
Agentic AI Meetup, Lehigh University. Talk and live demo.

Computing Allocations

PI, NSF ACCESS Discover Allocation (CIS251418) December 2025–December 2026
Multi-Agent System for Automated Information Collection and Idea Generation
750,000 ACCESS Credits (CloudBank research cloud, NCSA Delta GPU/CPU).

Co-PI, NSF ACCESS Explore Allocation (CIS261470) August 2026–August 2027
Stress-Testing Scientific Gatekeeping: Adversarial Evidence Corruption and Verifiable Human–AI Peer Review
200,000 ACCESS Credits (NCSA Delta GPU, DeltaAI). PI: Dr. Lichao Sun.

Honors and Awards

Outstanding Graduate Student, Nanjing University **2023**

Yingcai Scholarship, Nanjing University **2022**

Renmin Scholarship (People's Scholarship), Nanjing University **2018–2021**

Third Runner's Up in 15th Citi Cup Financial Innovation Application Competition, Citigroup **2019**

Second Runner's Up in 2019 "Chain to Future" University Blockchain Technology Application Competition, CCF **2019**

Outstanding Student Leader of the Communist Youth League, Nanjing University **2018–2019**

Skills

Programming: Python, TypeScript/JavaScript, SQL, Bash.
LLM Training & Inference: PyTorch, Hugging Face (Transformers, Accelerate, PEFT), DeepSpeed, FSDP2, verl, vLLM.
LLM Agents & Evaluation: Claude Agent SDK, Codex SDK, Model Context Protocol (MCP), LiteLLM, Harbor.
Engineering: Docker, Node.js, React, SQLite, uv, Git, Linux/ssh, LaTeX.
Natural languages: English, Mandarin, Cantonese.

Professional Service

Conference reviewer: ACL Rolling Review (2024–2026), COLM (2025–2026), EMNLP (2024–2026, incl. 2026 Demo Track), ICLR (2027).
Session chair: EMNLP 2026 (Virtual Session).