

Zhuochun Li

Mobile: +1 (878)-999-6935

Website: <https://zhuochunli.github.io/>

Email: glenn.seven.li@gmail.com

LinkedIn: <https://www.linkedin.com/in/zhuochunli>

Education

University of Pittsburgh

Ph.D. in Information Science

Master of Science in Information Science, GPA: 3.9/4.0

School of Computing and Information (SCI)

Pittsburgh, USA

Aug 2023 – Present (Expected 2027)

Jan 2021 – Dec 2022

Xi'an University of Technology

Bachelor of Engineering in Computer Science

Xi'an, China

Sep 2016 – Jul 2020

Profile

Highly motivated Ph.D. student in Information Science with research interests in **NLP** and **Machine Learning**.

Proficient in Python, Machine Learning, and advanced NLP methods. Research focuses on **LLM Reasoning**,

Knowledge Distillation, and **Post-training**. Completed **Applied Scientist Intern** at **Amazon** and **Microsoft**.

Published research in top venues such as **SIGIR**, **EMNLP**, **ACL**, and **ICLR**.

Selected Publications [[Google Scholar](#)]

Li, Zhuochun, et al. "One Success Isn't Reliability: Thinkingbox, a Sandbox and Benchmark for Agents in Stateful Business Workflows." arXiv preprint arXiv:2608.19741 (2026). Submitted to ICLR 2027

Li, Zhuochun, Yuelyu Ji, Yiming Zeng, and Daqing He. "SPEAR: Distilling Domain-Adaptive Reasoning Skeletons via Sequential Symbolic Alignment in Reinforcement Learning." arXiv preprint arXiv:2608.26550 (2026).

Li, Zhuochun, Zhang Yong, Ming Li. "Rethinking LLM-as-a-Judge: Representation-as-a-Judge with Small Language Models via Semantic Capacity Asymmetry." **ICLR 2026**

César Guerra-Solano, **Zhuochun Li**, Xiang Lorraine Li. "Think Globally, Group Locally: Evaluating LLMs Using Multi-Lingual Word Grouping Games." In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 20155-20176. 2025.

Li, Zhuochun, Yuelyu Ji, Rui Meng, Daqing He. "Learning from Committee: Reasoning Distillation from a Mixture of Teachers with Peer-Review." In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 4190–4205, Vienna, Austria. Association for Computational Linguistics.

Ji, Yuelyu, **Zhuochun Li**, Rui Meng, Daqing He. "Reason-to-Rank: Distilling Direct and Comparative Reasoning from Large Language Models for Document Reranking." In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 2320-2329. 2025.

Ji, Yuelyu, **Zhuochun Li**, Rui Meng, Sonish Sivarajkumar, Yanshan Wang, Zeshui Yu, Hui Ji, Yushui Han, Hanyu Zeng, and Daqing He. "RAG-RLRC-LaySum at BioLaySumm: Integrating Retrieval-Augmented Generation and Readability Control for Layman Summarization of Biomedical Texts." In *Proceedings of the 23rd Workshop on Biomedical Natural Language Processing: ACL 2024*, pp. 810–817, Bangkok, Thailand. Association for Computational Linguistics.

Li, Zhuochun, Khushboo Thaker, and Daqing He. "SiaKey: A Method for Improving Few-shot Learning with Clinical Domain Information." In *2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, pp. 1-4. IEEE, 2023.

Li, Zhuochun, Wang, Zhixiao, Wenyaoyan, Min Huang, Qinyuan Fan, and Xin Wang. "Domestic Violence Crisis Recognition Method based on Bi-LSTM+ Attention." In *2022 8th Annual International Conference on Network and Information Systems for Computers (ICNISC)*, pp. 569-575. IEEE, 2022.

Academic Experience

PhD Preliminary Exam Project, University of Pittsburgh

Mar 2024 – Mar 2025

Enhance the reasoning ability of student LM via knowledge distillation from multiple LLMs. Pittsburgh, U.S.

- Introduced a novel Fault-Aware Distillation via Peer-Review (FAIR) approach that enables student LLMs to better acquire reasoning skills from multiple teacher LLMs: ChatGPT, Gemini, and Mistral.

- Helped student LM learn not only from the gold-standard rationale but also from feedback on their own mistakes via instruction tuning through a simulated peer-review process between teacher LLMs.
- Demonstrated the effectiveness of our method across comprehensive experiments and analysis on mathematical(GSM8K, SVAMP), commonsense(StrategyQA), and logical(LogiQA) reasoning tasks.

Research Assistant, iRIS Lab at University of Pittsburgh

May 2022 – May 2023

Improve Few-shot Learning with Clinical Domain Information.

Pittsburgh, U.S.

- Conducted the project of Ovarian Cancer Forum, which cooperated with the School of Nursing.
- Studied few-shot learning methods in the application of text classification and recommendation systems.
- Achieved average accuracy of over 60% by 10-shot training on Siamese Networks via triple loss function, which is comparable to the BERT performance using the whole dataset on this task.

Computer Vision - 16-720A, Carnegie Mellon University

Jan 2022 - Apr 2022

Complete the Computer Vision course with grade A taught by Professor David Held of CMU *Pittsburgh, U.S.*

- Learned technologies about Spatial Pyramid Matching, Planar Homographies and Lucas-Kanade Tracking.
- Contributed to tasks about 3D Reconstruction, Neural Networks for Recognition and Photometric Stereo.
- Comprehended the cutting-edge deep learning models such as GAN, VAE and Transformers.

Undergraduate Thesis

Jan 2020 - Jul 2020

Implementation of Domestic Violence Crisis Recognition Method Based on Deep Learning

Xi'an, China

- Related paper has been published in the 2022 8th Annual International Conference on Network and Information Systems for Computers (ICNISC).
- Collected 1654 posts related to Domestic violence and built 50-dimensional word vectors via Word2Vec.
- Constructed CNN, RNN, LSTM, Bi-LSTM+self-Attention neural network models to accomplish text categorization task, Bi-LSTM+self-Attention model had the best performance, with an accuracy rate of 90.22% and recall rate of 93.98%.

Work Experience

Applied Scientist Intern, Microsoft

May 2026 - Jul 2026

Research on training and evaluation of multi-turn long-horizon agents in Copilot Studio

Redmond, U.S.

- Integrated the agent sandbox, Thinkingbox, with Verl frameworks to improve agent RL.
- Open-sourced the business agent benchmark: Thinkingbox-bench, and established leaderboards on LLMs.

Applied Scientist Intern, Amazon

Sep 2025 - Dec 2025

Research with evaluation and optimization of LLM agents

San Diego, U.S.

- Improved natural language optimization of LLMs based on GEPA framework and AWS Bedrock platform.
- Implemented prompt learning rate and token usage controls to optimize training efficiency and cost.

Research Intern, Ping An Technology

May 2025 - Aug 2025

Research Internship about LLM Probing on reference-free evaluation tasks

Shenzhen, China

- Published **ICLR 2026**: “Rethinking LLM-as-a-Judge: Representation-as-a-Judge with Small Language Models via Semantic Capacity Asymmetry”.
- Probed small LMs’ layer representations evaluation scores across 5 aspects on {question, answer} pairs.
- Built effective probing classifiers to filter data and improve fine-tuning performance on LLMs.

Machine Learning Engineer Intern, MEDA AI

May 2022 - Aug 2022

Internship with Startup Half Moon Tech to work on the Meda Metaverse project

Remote, U.S.

- Improved the text-to-speech (TTS) based on the Tacotron model and built server to train over 10M steps.
- Assisted in constructing 3D character model and got average accuracy over 85% on facial attributes classification task on dataset CelebA.
- Developed API for integrating with backend services and maintained Linux servers.

Software Engineering Intern, Pactera Technology

May 2021 - Aug 2021

Design and Development of an Intelligent Customer Service System

Wuhan, China

- Conducted field surveys over 100 customers, assisted designing database E-R model containing 34 tables.
- Contributed more than 20 web page interface implementations for different service requirements of clients.
- Developed online semantic analysis system using Baidu voice recognition API with 80% code coverage.

Skills

Programming Language/Platform: Python (main), C, C++, JAVA, SQL, Linux, GIT, AWS, Bedrock, Azure

Tools: Pytorch, Transformers, DeepSpeed, FSDP, Trl, Verl, Huggingface, Numpy, Pandas, Scikit-learn, Spacy.

Teaching

Fall 2026: TA for CS 1675 Intro to Machine Learning

Spring 2026: TA for CS 1671 Human Language Technologies

Spring 2025: TA for INFSCI 2440 Artificial Intelligence

Fall 2024: TA for INFSCI 0201 Intermediate Programming Python (LAB LEADER)

Fall 2024: TA for INFSCI 0410 Human Centered Systems

Spring 2024: TA for INFSCI 2440 Artificial Intelligence

Fall 2023: TA for INFSCI 2140 Information Storage & Retrieval

Fall 2023: TA for INFSCI 2410 Intro to Neural Networks

Professional Service

Reviewer, ICLR 2027

Reviewer, ARR August 2026

Reviewer, NeurIPS 2026

Reviewer, ACL 2026

Reviewer, ASIS&T 2026

Reviewer, ICHI 2026

Awards

Catherine Ofiesh and Gerald Orner Award (2026) - PhD academic excellence and top-tier research publication

Outstanding Undergraduate Thesis Award (2020)

Third Prize XUT Scholarship (2018)