Jason / Xiaochuan Li

✉ lixiaochuan20@gmail.com · ⌂ https://xiaochuanli.com/

Education

Carnegie Mellon University, Language Technology Institute   Jun. 2024 - Aug. 2024
Visiting Student

The University of Hong Kong , Department of Computer Science   Jan. 2024 - Jun. 2024
Exchange Student with full scholarship

Tsinghua University, School of Software   2021 - 2025(Expected)
Bachelor's degree in Computer Science and Artificial Intelligence
With the highest honor in THU, 10 out of 3000

Publications

(* stands for equal contribution; Listed in reverse chronological order)

Conference Publications

[1] **(preprint)** Montessori-Instruct: Generate Influential Training Data Tailored for Student Learning
Xiaochuan Li, Zichun Yu, Chenyan Xiong
arXiv [PDF] [Arxiv]

[2] **(NIPS 2024 D&B Track)** OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments
Tianbao Xie, Danyang Zhang, Jixuan Chen, Xiaochuan Li, Siheng Zhao, Ruisheng Cao, Toh Jing Hua, Zhoujun Cheng, Dongchan Shin, Fangyu Lei, Yitao Liu, Yiheng Xu, Shuyan Zhou, Silvio Savarese, Caiming Xiong, Victor Zhong, Tao Yu
The Thirty-Eighth Annual Conference on Neural Information Processing Systems [PDF] [Arxiv]

[3] **(ACL 2024 Findings)** Evaluating Robustness of Generative Search Engine on Adversarial Factoid Questions
Xiaochuan Li*, Xuming Hu*, Junzhe Chen, Yinghui Li, Yangning Li, Xiaoguang Li, Yasheng Wang, Qun Liu, Lijie Wen, Philip S. Yu, Zhijiang Guo
The 62nd Annual Meeting of the Association for Computational Linguistics [PDF] [OpenReview]

[4] **(ICLR 2024 Spotlight)** Towards Understanding Factual Knowledge of Large Language Models
Xiaochuan Li*, Xuming Hu*, Junzhe Chen*, Yufei Guo, Lijie Wen, Philip S. Yu, Zhijiang Guo
The Twelfth International Conference on Learning Representations [PDF] [OpenReview]

Journal Articles

[1] SongC: A Compiler for Hybrid Near-Memory and In-Memory Many-Core Architecture
Junfeng Lin, Huanyu Qu, Songchen Ma, Xinglong Ji, Hongyi Li, Xiaochuan Li, Chenhang Song, Weihao Zhang
IEEE Transactions on Computers, [TCSI-2023-01-0048] (IEEE TC) [PDF]

Research Experience

CX's Group, LTI, Carnegie Mellon University   Jun. 2024 - Present
Research Assistant to Professor Chenyan Xiong
Topic: **Data-Centric Language Models**
• Lead the project on synthesizing higher-quality instruction-following data and extend to pre-training.

XLANG lab, The University of Hong Kong   Jan. 2024 - Jun. 2024
Research assistant to Prof. Tao Yu
Topic: **Agent Evaluation**

- Working on evaluating the capabilities and limitations of digital computer agents in real-world, open-ended tasks.

THUBPM, Tsinghua University Dec. 2022 - Dec. 2023

Research assistant to Prof. Philip S. Yu and Prof. Lijie Wen

Topic: **Credibility and Hallucination Phenomena of LLMs**

- Leading the project on discovering factual errors in LLMs' reasoning using Adversarial Attacks.

WORK EXPERIENCE

Qwen Team, Alibaba Nov. 2024 - Present

Research Intern; Advisor: Binyuan Hui

Topic: **Language Agents as World Models**

SELECTED AWARDS AND HONORS

Scholarship

- The Special-Class Scholarship, Tsinghua University Nov. 2024
  (The highest honor, 10 out of 3000)
- The Fung Scholarship, Fung Foundation Jan. 2024
- Comprehensive Excellence Scholarship, Tsinghua University Oct. 2024/2023/2022
  (Top 1% in both Academics and Research)
- Academic Excellence Scholarship, Tsinghua University Sep. 2021
  (Top 5% in Academics)

Awards

- First Prize in Software Development Race, Tsinghua University Nov. 2023
- Tsinghua University Undergraduate Academic Promotion Fund Apr. 2023
- Outstanding student leaders, Tsinghua University Oct. 2022
- Chairman of the 21st Student Union, School of Software, Tsinghua University Jan. 2023 - Jan. 2024

ADDITIONAL INFORMATION

Conference Reviewing

- EMNLP 2024; NAACL 2024; ICLR 2025;

Skills

- C++; Python; PyTorch; LaTex; Linux; Front & Back-end development
- Chinese (Native); English (Fluent); French (Basic)