Zhaoxuan Tan

Address: Room 355B, Fitzpatrick Hall of Engineering, Notre Dame, IN 46556, United States
Email: ztan3@nd.edu, tanzx9@gmail.com
Homepage: zhaoxuan.info
Last updated: May 22, 2025

Research Interests

My research focuses on leveraging user data, including user-generated content and behavior data, to personalize and enhance large language models, as well as to detect suspicious user behavior.

Education

University of Notre Dame, Notre Dame, IN, United States 2023.08 - present
Ph.D. in Computer Science and Engineering
Advisor: Prof. Meng Jiang

Xi'an Jiaotong University, Xi'an, Shaanxi, China 2019.08 - 2023.07
B.E. in Computer Science and Technology
Advisor: Prof. Minnan Luo

Industrial Experience

Student Researcher @ Core ML, Google, Remote 2025.10 - 2025.12
Project: Efficient Large User Models for Personalization Tasks.
Host: Dr. Ao Liu, Dr. Yan Zhu

Applied Scientist Intern @ Rufus, Amazon, Palo Alto, CA 2025.06 - 2025.10
Project: Personalized Recommendation Reasoning with User-Centric Rationales
Host: Dr. Ruijie Wang

Research Intern @ Office of Applied Research, Microsoft, Remote 2025.03 - 2025.06
Project: Making chain-of-thought reasoning more steerable and adaptable.
Host: Dr. Pei Zhou, Dr. Mengting Wan

Applied Scientist Intern @ Rufus, Amazon, Palo Alto, CA 2024.05 - 2024.10
Project: Aligning large language models with implicit preferences from user-generated content.
Host: Dr. Zheng Li, Dr. Tianyi Liu

Publications (* indicates equal contribution)

Aligning Large Language Models with Implicit Preferences from User-Generated Content.
Zhaoxuan Tan, Zheng Li, Tianyi Liu, Haodong Wang, Hyokun Yun, Ming Zeng, Pei Chen, Zhihan Zhang, Yifan Gao, Ruijie Wang, Priyanka Nigam, Bing Yin, Meng Jiang.
In *Proceedings of ACL 2025*.

Enhancing Mathematical Reasoning in LLMs by Stepwise Correction.
Zhenyu Wu, Qingkai Zeng, Zhihan Zhang, Zhaoxuan Tan, Chao Shen, Meng Jiang.
In *Proceedings of ACL 2025*.

Modality-Aware Neuron Pruning for Unlearning in Multimodal Large Language Models
Zheyuan Liu, Guangyao Dou, Xiangchi Yuan, Chunhui Zhang, Zhaoxuan Tan, Meng Jiang.
In *Proceedings of ACL 2025*.

CodeTaxo: Enhancing Taxonomy Expansion with Limited Examples via Code Language Prompts.
Qingkai Zeng, Yuyang Bai, Zhaoxuan Tan, Zhenyu Wu, Shangbin Feng, Meng Jiang.

In *Proceedings of ACL 2025-Findings.*

IHEval: Evaluating Language Models on Following the Instruction Hierarchy.

Zhihan Zhang, Shiyang Li, Zixuan Zhang, Xin Liu, Haoming Jiang, Xianfeng Tang, Yifan Gao, Zheng Li, Haodong Wang, Zhaoxuan Tan, Yichuan Li, Qingyu Yin, Bing Yin, Meng Jiang.

In *Proceedings of NAACL 2025.*

Protecting Privacy in Multimodal Large Language Models with MLLMU-Bench.

Zheyuan Liu, Guangyao Dou, Mengzhao Jia, Zhaoxuan Tan, Qingkai Zeng, Yongle Yuan, Meng Jiang.

In *Proceedings of NAACL 2025.*

Can Large Language Models Understand Preferences in Personalized Recommendation?

Zhaoxuan Tan, Zinan Zeng, Qingkai Zeng, Zhenyu Wu, Zheyuan Liu, Fengran Mo, Meng Jiang.

arXiv preprint 2025.

Personalized Pieces: Efficient Personalized Large Language Models through Collaborative Efforts.

Zhaoxuan Tan, Zheyuan Liu, Meng Jiang.

In *Proceedings of EMNLP 2024.*

Democratizing Large Language Models via Personalized Parameter-Efficient Fine-tuning.

Zhaoxuan Tan, Qingkai Zeng, Yijun Tian, Zheyuan Liu, Bing Yin, Meng Jiang.

In *Proceedings of EMNLP 2024.*

Can LLM Graph Reasoning Generalize beyond Pattern Memorization?

Yizhuo Zhang*, Heng Wang*, Shangbin Feng*, Zhaoxuan Tan, Xiaochuang Han, Tianxing He, Yulia Tsvetkov.

In *Proceedings of EMNLP 2024-Findings.*

Machine Unlearning in Generative AI: A Survey.

Zheyuan Liu, Guangyao Dou, Zhaoxuan Tan, Yijun Tian, Meng Jiang.

arXiv preprint 2024.

Empirical Guidelines for Deploying LLMs onto Resource-constrained Edge Devices.

Ruiyang Qin, Dancheng Liu, Zheyu Yan, Zhaoxuan Tan, Zixuan Pan, Zhenge Jia, Meng Jiang, Ahmed Abbasi, Jinjun Xiong, Yiyu Shi.

ACM Transactions on Design Automation of Electronic Systems, 2025.

Large Language Models Can Self-Correct with Minimal Effort .

Zhenyu Wu, Qingkai Zeng, Zhihan Zhang, Zhaoxuan Tan, Chao Shen, Meng Jiang.

In *Proceedings of EMNLP 2024.*

Chain-of-Layer: Iteratively Prompting Large Language Models for Taxonomy Induction from Limited Examples.

Qingkai Zeng*, Yuyang Bai*, Zhaoxuan Tan, Shangbin Feng, Zhenwen Liang, Zhihan Zhang, Meng Jiang.

In *Proceedings of CIKM 2024.*

DELL: Generating Reactions and Explanations for LLM-Based Misinformation Detection.

Herun Wan*, Shangbin Feng*, Zhaoxuan Tan, Heng Wang, Yulia Tsvetkov, Minnan Luo.

In *Proceedings of ACL-Findings 2024.*

What Does the Bot Say? Opportunities and Risks of Large Language Models in Social Media Bot Detection.

Shangbin Feng, Herun Wan, Ningnan Wang, Zhaoxuan Tan, Minnan Luo, Yulia Tsvetkov.

In *Proceedings of ACL 2024.*

Towards Safer Large Language Models through Machine Unlearning.

Zheyuan Liu, Guangyao Dou, Zhaoxuan Tan, Yijun Tian, Meng Jiang.
In *Proceedings of ACL-Findings 2024.*

Knowledge Crosswords: Geometric Reasoning over Structured Knowledge with Large Language Models.
Wenxuan Ding*, Shangbin Feng*, Yuhan Liu, Zhaoxuan Tan, Vidhisha Balachandran, Tianxing He, Yulia Tsvetkov. In *Proceedings of ACL-Findings 2024.*

KGQUIZ: Evaluating the Generalization of Encoded Knowledge in Large Language Models.
Yuyang Bai*, Shangbin Feng*, Vidhisha Balachandran, Zhaoxuan Tan, Shiqi Lou, Tianxing He, Yulia Tsvetkov.
In *Proceedings of The Web Conference (WWW) 2024 (oral).*

User Modeling in the Era of Large Language Models: Current Research and Future Directions.
Zhaoxuan Tan, Meng Jiang.
In *Proceedings of IEEE Data Engineering Bulletin 2023.*

LMBot: Distilling Graph Knowledge into Language Model for Graph-less Deployment in Twitter Bot Detection.
Zijian Cai, Zhaoxuan Tan, Zhenyu Lei, Zifeng Zhu, Hongrui Wang, Qinghua Zheng, Minnan Luo.
In *Proceedings of WSDM 2024.*

BotPercent: Estimating Twitter Bot Populations from Groups to Crowds.
Zhaoxuan Tan*, Shangbin Feng*, Melanie Sclar, Herun Wan, Minnan Luo, Yejin Choi, Yulia Tsvetkov
In *Proceedings of EMNLP-Findings 2023.*

Detecting Spoilers in Movie Reviews with External Movie Knowledge and User Networks.
Heng Wang, Wenqian Zhang, Yuyang Bai, Zhaoxuan Tan, Shangbin Feng, Qinghua Zheng, Minnan Luo.
In *Proceedings of EMNLP 2023.*

Can Language Models Solve Graph Problems in Natural Language?
Heng Wang, Shangbin Feng, Tianxing He, Zhaoxuan Tan, Xiaochuang Han, Yulia Tsvetkov.
In *Proceedings of NeurIPS 2023 (spotlight).*

HOFA: Twitter Bot Detection with Homophily-Oriented Augmentation and Frequency Adaptive Attention.
Sen Ye, Zhaoxuan Tan, Zhenyu Lei, Ruijie He, Hongrui Wang, Qinghua Zheng, Minnan Luo.
arXiv preprint 2023.

KALM: Knowledge-Aware Integration of Local, Document, and Global Contexts for Long Document Understanding.
Shangbin Feng, Zhaoxuan Tan, Wenqian Zhang, Zhenyu Lei, Yulia Tsvetkov.
In *Proceedings of ACL 2023.*

BotMoE: Twitter Bot Detection with Community-Aware Mixtures of Modal-Specific Experts.
Yuhan Liu, Zhaoxuan Tan, Heng Wang, Shangbin Feng, Qinghua Zheng, Minnan Luo.
In *Proceedings of SIGIR 2023.*

KRACL: Contrastive Learning with Graph Context Modeling for Sparse Knowledge Graph Completion.
Zhaoxuan Tan, Zilong Chen, Shangbin Feng, Qingyue Zhang, Qinghua Zheng, Jundong Li, Minnan Luo.
In Proceedings of The Web Conference (WWW) 2023.

TwiBot-22: Towards Graph-Based Twitter Bot Detection.
Shangbin Feng*, Zhaoxuan Tan*, Herun Wan*, Ningnan Wang*, Zilong Chen*, Binchi Zhang*, Qinghua

Zheng, Wenqian Zhang, Zhenyu Lei, Shujie Yang, Xinshun Feng, Qingyue Zhang, Hongrui Wang, Yuhan Liu, Yuyang Bai, Heng Wang, Zijian Cai, Yanbo Wang, Lijing Zheng, Zihan Ma, Jundong Li, Minnan Luo.
In *Proceedings of NeurIPS 2022, Datasets and Benchmarks Track*.

PAR: Political Actor Representation Learning with Social Context and Expert Knowledge.
Shangbin Feng, Zhaoxuan Tan, Zilong Chen, Peisheng Yu, Qinghua Zheng, Xiaojun Chang, Minnan Luo.
In *Proceedings of EMNLP 2022*.

Heterogeneity-Aware Twitter Bot Detection with Relational Graph Transformers.
Shangbin Feng, Zhaoxuan Tan, Rui Li, Minnan Luo.
In *Proceedings of AAAI 2022*.

Honors and Awards

Top Reviewer, NeurIPS 2024	2024
Excellent Bachelor Thesis (rank 1/172), XJTU	2023
AAAI Student Scholarship	2022
National Second Prize, CUMCM	2021
Dean's List, XJTU	2020 - 2022

Services

Area Chair for ACL Rolling Review	from 2025 Feb
Reviewer for TMLR	from 2025
Reviewer for ICML	from 2025
Reviewer for AISTATS	from 2025
Reviewer for COLM	from 2024
Reviewer for KDD	from 2024
Reviewer for ACL Rolling Review	from 2023 Dec
Reviewer for The Web Conference	2024, 2025
Reviewer for ICLR	from 2024
Reviewer for TKDE	from 2023
Reviewer for TNNLS	from 2023
Reviewer for ICWSM	from 2023
Reviewer for NeurIPS	from 2023
Virtual Volunteer for EMNLP	2022, 2023
Reviewer for NeurIPS Datasets and Benchmarks Track	2022
Reviewer for Learning on Graphs Conference	from 2022
Reviewer for Student Research Workshop @ ACL	2025
Reviewer for AGI Workshop @ ICLR	2024
Reviewer for WiNLP Workshop @ EMNLP	2024
Reviewer for KnowledgeNLP Workshop @ ACL	2024
Reviewer for Transactions on Networking	2023
Reviewer for GCLR Workshop @ AAAI	2024
Reviewer for Temporal Graph Learning Workshop @ NeurIPS	2023
Director of the LUD lab	2022 - 2023

Mentorship

Heng Wang $\rightarrow$ first author in 1*NeurIPS 2023 (spotlight), 1*EMNLP, PhD student @ UIUC

Yuhan Liu → first author in 1*SIGIR 2023, PhD student @ NYU

Yuyang Bai → first author in 1*WWW 2024, PhD student @ TAMU

Sen Ye → first author in 1*WSDM 2024 submission, PhD student @ Peking University

Zijian Cai → first author in 1*WSDM, MS student @ The Institute of Automation, CAS

Teaching

Teaching Assistant of CSE 30124 Introduction to Artificial Intelligence, University of Notre Dame 2023 Fall

Teaching Assistant of CSE 40113 Design/Analysis of Algorithm, University of Notre Dame 2024 Spring

Skills

- Programming Skills: Python, PyTorch, MATLAB, C/C++, bash, HTML/CSS, LaTeX
- Language Skills: Mandarin (native), English (TOEFL 107: R 29, L 29, S 22, W 27), Cantonese (native)

Funding

OpenAI Researcher Access Program

Source of Support: OpenAI

Total Award Amount: $10,000 for OpenAI API credits

Period of Performance: Feb 2025 to Feb 2026