YIGE YUAN

State Key Laboratory of AI Safety
Institute of Computing Technology, Chinese Academy of Sciences
No.6 Kexueyuan South Road Zhongguancun, Haidian District, Beijing, China

yuanyige20z@ict.ac.cn
github.com/yuanyige
Google Scholar

RESEARCH INTEREST

My research goal is to build **Trustworthy AI** that performs reliably and ethically for social good and human well-being. To achieve this, I worked on generalization, alignment and agentic system across the domains of graph, vision, and language.

- **Machine Learning Generalization**: To make models generalize stably under domain shifts, noises, or perturbations.
- **Large Language Model Alignment**: To align large language models with human values and safety requirements.
- **Agent System Reasoning and Planning**: To enhance agent's logical reasoning and strategic planning for complex tasks.

EDUCATION

Institute of Computing Technology, Chinese Academy of Sciences Sep 2020 - current
Ph.D. in Computer Software and Theory (Advisor: Prof. Xueqi Cheng & A.P. Bingbing Xu)

Xidian University, School of Cyberspace Security Sep 2016 - Jun 2020
B.S. in Information Security (Experimental Class, GPA: 3.8/4.0)

INTERNSHIP

Tongyi Lab, Alibaba Group Feb 2025 - current
Research Internship in Large Language Models and Multi-Agent Systems
Project: Reinforcement Learning for AI Agent Reasoning and Planning

PUBLICATIONS

C: conference, J: journal, W: workshop, P: preprint / * equal contribution

Machine Learning Generalization

In this topic, I focus on enhancing generalization under multiple learning paradigms: **(1) Test-Time Adaptation**, where I leverage energy-based modeling to handle distribution shifts and enable continual learning on the fly [C1,P1]. **(2) Physics-Informed Learning**, where I introduce Partial Differential Equations (PDEs) to impose function-level constraints beyond purely data-driven approaches [C2], and **(3) Self-Supervised Learning**, where I address overfitting and out-of-distribution issues in both contrastive learning [J1,C4,C5] and generative tasks such as GAN training [C3] and multi-modal generation [P2].

[C1] TEA: Test-time Energy Adaptation
Yige Yuan, Bingbing Xu, Liang Hou, Fei Sun, Huawei Shen, Xueqi Cheng
IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2024, Main Track, CCF-A

[C2] PDE+: Enhancing Generalization via PDE with Adaptive Distributional Diffusion
Yige Yuan, Bingbing Xu, Bo Lin, Liang Hou, Fei Sun, Huawei Shen, Xueqi Cheng
AAAI Conference on Artificial Intelligence (**AAAI**), 2024, Main Track, CCF-A

[J1] Towards Generalizable Graph Contrastive Learning: An Information Theory Perspective
Yige Yuan, Bingbing Xu, Huawei Shen, Qi Cao, Keting Cen, Wen Zheng, Xueqi Cheng
Neural Networks (**NN**), Volume 172, CCF-B, Q1, IF=8.4

[P1] MITA: Bridging the Gap between Model and Data for Test-Time Adaptation
Yige Yuan, Xu Bingbing, Liang Hou, Fei Sun, Huawei Shen, Xueqi Cheng
IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**), UnderReview

[C3] Augmentation-Aware Self-Supervision for Data-Efficient GAN Training
Liang Hou, Qi Cao, **Yige Yuan**, Songtao Zhao, Chongyang Ma, Siyuan Pan, et al.
Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2023, Main Track, CCF-A

[C4] InfoNCE is a Free Lunch for Semantically guided Graph Contrastive Learning
Zixu Wang, Bingbing Xu, **Yige Yuan**, Huawei Shen and Xueqi Cheng
International ACM Conference on Research and Development in Information Retrieval (**SIGIR**), 2025, Full Paper, CCF-A

[C5] Negative as Positive: Enhancing Out-of-distribution Generalization for Graph Contrastive Learning
Zixu Wang, Bingbing Xu, **Yige Yuan**, Huawei Shen and Xueqi Cheng
International ACM Conference on Research and Development in Information Retrieval (**SIGIR**), 2024, Short Paper, CCF-A

[C6] History Driven Sampling for Scalable Graph Neural Networks
Yang Li, Bingbing Xu, Fei Sun, Qi Cao, **Yige Yuan**, and Huawei Shen
International Conference on Database Systems for Advanced Applications (**DASFAA**), 2024, Reseach Track, CCF-B

[C7] MIGE: A Unified Framework for Multimodal Instruction-Based Image Generation and Editing
Xueyun Tian, Wei Li, Bingbing Xu, **Yige Yuan**, Yuanzhuo Wang, Huawei Shen
ACM International Conference on Multimedia (**ACMMM**), 2025, CCF-A

Large Language Model Alignment

In this topic, I focus on Reinforcement Learning from Human Feedback (RLHF) and Direct Preference Optimization (DPO), aiming to address the following key challenges: **(1) Training instability and computational cost**: I eliminate the need for hyperparameters and reference models by proposing a minimalist objective for preference optimization, which also stabilizes training by avoiding gradient dominance from offline reject samples [C9]. I enable alignment at inference time, achieving an optimal RLHF policy without costly training, which also avoids shallow alignment issues and ensures strong safety guarantees against attacks [C8]. **(2) Limitations of Bradley-Terry (BT) modeling**: I introduce a calibration term for point-wise preference modeling [C11], reinterpret RLHF from an imitation learning perspective, which replaces BT modeling with density ratio estimation [C10, C12]. Above methods alleviate the margin-only bias of BT's pairwise nature and mitigate the issue of likelihood displacement. Moreover, I explore partial reward modeling to ensure consistency over response trajectories [C13].

[C8] Inference-time Alignment in Continuous Space
Yige Yuan*, Teng Xiao*, Yunfan Li, Bingbing Xu, Shuchang Tao, Yunqi Qiu, Huawei Shen, Xueqi Cheng
Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2025, Main Track, CCF-A

[C9] SimPER: A Minimalist Approach to Preference Alignment without Hyperparameters
Teng Xiao*, **Yige Yuan***, Zhengyu Chen, Mingxiao Li, Shangsong Liang, Zhaochun Ren, Vasant G Honavar
International Conference on Learning Representations (**ICLR**), 2025, Main Conference

[C10] On a Connection Between Imitation Learning and RLHF
Teng Xiao, **Yige Yuan**, Mingxiao Li, Zhengyu Chen, Vasant G Honavar
International Conference on Learning Representations (**ICLR**), 2025, Main Conference

[C11] Calibrated Preference Optimization for Direct Language Model Alignment
Teng Xiao, **Yige Yuan**, Huaisheng Zhu, Mingxiao Li, Vasant G Honavar
Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2024, Main Track, CCF-A

[C12] How to Leverage Demonstration Data in Alignment for Large Language Model? A Self-Imitation Learning Perspective
Teng Xiao, Mingxiao Li, **Yige Yuan**, Huaisheng Zhu, Chao Cui, Vasant G Honavar
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2024, Main, CCF-B

[C13] Score Consistency Meets Preference Alignment: Dual-Consistency for Partial Reward Modeling
Bin Xie, Bingbing Xu, **Yige Yuan**, Shengmao Zhu, Huawei Shen
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2025, Main, CCF-A

[P2] MPG: Multi-Personality Generation of Large Language Models at Decoding-time
Rongxin Chen, Li Yunfan, **Yige Yuan**, Bingbing Xu, Huawei Shen
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025, UnderReview

Agent System Reasoning and Planning

In this topic, I focus on scaling Chain-of-Thought reasoning and encouraging cognitive behaviors such as reflection and correction within agent systems. Specifically, my work covers: **(1) Reinforcement Learning with Verifiable Rewards (RLVR)**, where I find that reasoning ability can be effectively incentivized by simple and efficient Supervised Fine-Tuning (SFT) from very weak supervision, achieving performance comparable to or even surpassing much more costly RLVR [P3]. **(2) Agentic Modeling for Social Simulation**, including knowledge-driven LLM agent framework for user modeling [C16], fine-grained agentic self-correction & reflection [C14] and a survey about LLM-based agent empowered social sciences [P4]. I will be co-organizing a workshop on LLM agents for social simulation at CIKM 2025 [C15].

[P3] Incentivizing Strong Reasoning from Weak Supervision
Yige Yuan*, Teng Xiao*, Shuchang Tao, Xue Wang, Jinyang Gao, Bolin Ding, Bingbing Xu
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2026, UnderReview

[C14] Fact-Level Calibration and Correction for Long-Form Generations
Yige Yuan, Xu Bingbing, Hexiang Tan, Fei Sun, Teng Xiao, Wei Li, Huawei Shen, Xueqi Cheng
International ACM Conference on Research and Development in Information Retrieval (**SIGIR**), 2025, Short Paper, CCF-A

[C15] The 1st Workshop on LLM Agents for Social Simulation
Yige Yuan, Junkai Zhou, Bingbing Xu, Liang Pang, Du Su, An Zhang, Teng Xiao, Fengli Xu, Zhaochun Ren, Xu Chen
International Conference on Information and Knowledge Management (**CIKM**), 2025, CCF-B

[C16] Unveiling the Potential of LLMs in Simulated Society: A Knowledge-Driven LLM Agent Framework for User Modeling
Shengmao Zhu, Bingbing Xu, **Yige Yuan**, Bin Xie, Yunfan Li, Huawei Shen
ACM Web Conference (**WWW**), 2025, Companion Proceedings, CCF-A

[P4] LLM-based Agent Empowered Social Sciences: From a Social Simulation Perspective
Zixu Wang, Bin Xie, Bingbing Xu, Shengmao Zhu, **Yige Yuan**, Liang Pang, Du Su, Long Yang, et al.
ACM Computing Surveys (**CSUR**), UnderReview

FUTURE DIRECTION

Agentic Reinforcement Learning. During my internship at Alibaba, I contributed to the development and deployment of the BeyondAgent Project, where I was primarily responsible for fine-grained reward modeling to enhance agent exploration. This experience highlighted the pivotal role of entropy in reinforcement learning: by strategically balancing exploration and exploitation through entropy, we can address several fundamental challenges that I aim to explore in my future research: (1) How can we effectively activate reasoning, reflection, and other cognitive behaviors across turns in multi-turn tool use/function calling for agent learning? (2) How can we develop agents with general reasoning abilities beyond task- or domain-specificity? (3) How can agents achieve long-term self-evolution with minimal or no external feedback?

Trustworthy Agent System. As AI agents enter high-stakes and open-ended environments, ensuring their trustworthiness is essential. My Ph.D. research focuses on building trustworthy AI, improving robustness under distributional shifts and adversarial attacks, and aligning large language models with human values. A trustworthy agent must go beyond generating accurate outputs to making safe, ethical, and transparent decisions. Looking ahead, I aim to study Agent Alignment and Trustworthy Tool Use. As agents gain the ability to interact with external tools and influence the real world, preventing misuse, manipulation, and misalignment with human intent becomes increasingly important.

Agentic Modeling for Social Simulation. Agent-based modeling provides a powerful framework to simulate and understand complex social dynamics, particularly in settings where human-AI interactions shape collective behavior. My research experience with LLM-based agent user simulation has motivated me to further explore the potential of agentic systems in social simulation, which can be leveraged for discovering social patterns, interpreting social phenomena, validating social theories, and forecasting policy outcomes. By developing cognitively and socially grounded agents, I aim to bridge AI modeling and computational social science, enabling more realistic and insightful simulations of societal processes.

HONORS & AWARDS

President Award, *Chinese Academy of Sciences*	2025
First Place, *AgentSociety Challenge @ WWW 2025*	2025
National Scholarship, *Ministry of Education of the People's Republic of China*	2024
First-Class Scholarship, *University of Chinese Academy of Sciences*	2024
President Award, *Institute of Computing Technology, Chinese Academy of Sciences*	2023
First-Class Scholarship, *University of Chinese Academy of Sciences*	2022
Outstanding Student Award, *University of Chinese Academy of Sciences*	2022
First Prize, *The 12th National College Students Information Security Contest*	2019
First Prize, *15th National Science and Technology Academic Competition of Challenge Cup*	2017

INVITED TALKS

NICE Webinar, On Energy Perspectives of Alignment: From Training to Inference	Sep 2025
WiseModel Talk, On a Connection Between Imitation Learning and RLHF	April 2025
NICE Webinar, On a Connection Between Imitation Learning and RLHF	March 2025
AITime Youth PhD Talk, On a Connection Between Imitation Learning and RLHF	March 2025
LOGS Webinar, Partial Differential Equation-Driven Generalizable Neural Networks	Mar 2024
AITime Webinar, TEA: Test-time Energy Adaptation	April 2024
WizSci Webinar, PDE+: Enhancing Generalization via PDE with Adaptive Distributional Diffusion	Jan 2024

ACADEMIC SERVICES

- Conference Reviewer: NeurIPS (2024, 2025), ICML 2025, ICLR (2025, 2026), AISTATS (2025, 2026), KDD 2025, WWW (2025, 2026), ACMMM 2025, AAAI (2025, 2026), IJCAI 2025, ACL 2025, EMNLP (2024, 2025), COLING 2025, ACL Rolling Review, MIDL 2025, IJCNN 2025
- Journal Reviewer: IEEE Transactions on Knowledge and Data Engineering (TKDE), Applied Intelligence (APIN), CAAI Transactions on Intelligence Technology, IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)