Mengyue Yang (She/Her)

United Kingdom

School of Engineering Mathematics and Technology, University of Bristol

mengyue.yang@bristol.ac.uk yangmy6750@gmail.com
Google Scholar: kJJkqdcAAAAJ Personal Website: ymy4323460.github.io

I am a Lecturer in AI at the University of Bristol in United Kingdom with emphasis on **causality**, and **reinforcement learning** – from theory to applications. My work spans (1) **causal representation learning**, (2) **causality and LLMs**: reasoning, planning and explanation, (3) **causal modelling in open-ended worlds**: causal foundation world models and (4) **general decision-making** problems: reinforcement learning and LLMs.

Professional Experience

University of Bristol, UK *Oct 2024 – Present*
School of Engineering Mathematics and Technology, MaVi Group
Lecturer (=U.S. Assistant Professor) in AI
Research interests: Causality and reinforcement learning.

Education

University College London, UK *Sep 2020 – Aug 2025*
PhD in Computer Science Supervisor: Prof. Jun Wang *PhD Thesis Submitted in Sep 2024*
Research Interests: Causal Representation Learning, Reinforcement Learning
First-year viva examiner: Prof. Ricardo Silva. Thesis examiners: Dr. Yingzhen Li and Prof. Hao Ni

University of Chinese Academy of Sciences, China *Sep 2017 – Jul 2020*
MSc in Computer Application Technology
Research Interests: Causal Inference, Online Learning and Reinforcement Learning.

Internships

KAUST, Visiting PhD Student *Aug 2024 – Sep 2024*
Supervised by Prof. Jürgen Schmidhuber. Causal World Models.

MBZUAI, Visiting PhD Student *Mar 2024 – Jul 2024*
Supervised by Prof. Kun Zhang. Causal representation learning.

TikTok (ByteDance Research), London, Research Intern *Feb 2022 – Jul 2022*
Supervised by Dr. Jean-Francois Ton and Dr. Hang Li. Causal representation learning, fairness.

Microsoft STCA, Beijing, Software Engineer Summer Intern *Jul 2019 – Oct 2019*
Ads Data & AI Platform Team. Supervised by Dr. Hao Wu and Dr. Keli Gui. Large-scale clusters, Microsoft Azure, distributional machine learning.

Didi AI Labs, Beijing, Research Intern *Sep 2018 – Jul 2019*
Reinforcement Learning Team, supervised by Dr. Zhiwei Qin and Dr. Qingyang Li. Online learning.

Selected Publications (Full list please check Google Scholar)

Publications in 2025:

- **Causal Sufficiency and Necessity Improves Chain-of-Thought Reasoning**. NeurIPS 2025.
  Xiangning Yu, Zhuohan Wang, Linyi Yang, Haoxuan Li, Anjie Liu, Xiao Xue, Jun Wang, **Mengyue Yang**.
- **Unveiling Extraneous Sampling Bias with Data Missing-Not-At-Random**. NeurIPS 2025.
  Haocheng Yang, Chunyuan Zheng, Haoxuan Li, **Mengyue Yang**.
- **A Principle of Pre-Strategy Intervention for Multi-Agent Reinforcement Learning**. NeurIPS 2025.
  Anjie Liu, Jianhong Wang, Samuel Kaski, Jun Wang, **Mengyue Yang**.
- **Curious Causality-Seeking Agents Learn Meta Causal World**. NeurIPS 2025.
  Zhiyu Zhao, Haoxuan Li, Haifeng Zhang, Jun Wang, Francesco Faccio, Jürgen Schmidhuber, **Mengyue Yang**.
- **MF-LLM: Simulating Population Decision Dynamics via a Mean-Field Large Language Model Framework**. NeurIPS 2025.

Qirui Mi, **Mengyue Yang**, Xiangning Yu, Zhiyu Zhao, Cheng Deng, Bo An, Haifeng Zhang, Xu Chen, Jun Wang.

- **Decentralized Dynamic Cooperation of Personalized Models for Federated Continual Learning**. NeurIPS 2025.
  Danni Yang, Zhikang Chen, Sen Cui, **Mengyue Yang**, Ding Li, Abudukelimu Wuerkaixi, Haoxuan Li, Jinke Ren, Mingming Gong.
- **Fine-Grained Interpretation of Political Opinions in Large Language Models** . AAAI 2026 (Artificial Intelligence for Social Impact Track).
  Jingyu Hu, **Mengyue Yang**, Mengnan Du, Weiru Liu.
- **Large Language Models are Demonstration Pre-Selectors for Themselves**. ICML 2025.
  Jiarui Jin, Yuwei Wu, Xiaoting He, Haoxuan Li, Weinan Zhang, Yiming Yang, Yong Yu, Jun Wang, **Mengyue Yang**.
- **When Can Proxies Improve the Sample Complexity of Preference Learning?** ICML 2025.
  Yuchen Zhu, Daniel Augusto de Souza, Zhengyan Shi, **Mengyue Yang**, Pasquale Minervini, Alexander D'Amour, Matt J. Kusner.
- **Causal Representation Learning from Multimodal Biological Observations**. ICLR 2025.
  Xue Yan, Yan Song, Xidong Feng, **Mengyue Yang**, Haifeng Zhang, Haitham Bou Ammar, Jun Wang.
- **Efficient Reinforcement Learning with Large Language Model Priors**. ICLR 2025.
  Yuewen Sun, Lingjing Kong, Guangyi Chen, Loka Li, Gongxu Luo, Zijian Li, Yixuan Zhang, Yujia Zheng, **Mengyue Yang**, Petar Stojanov, Eran Segal, Eric P. Xing, Kun Zhang.
- **Learning Macroeconomic Policies through Dynamic Stackelberg Mean-Field Games**. ECAI 2025.
  Qirui Mi, Zhiyu Zhao, Chengdong Ma, Siyu Xia, Yan Song, **Mengyue Yang**, Jun Wang, Haifeng Zhang.
- **Mean Field Correlated Imitation Learning**. AAMAS 2025.
  Zhiyu Zhao, Chengdong Ma, Qirui Mi, Ning Yang, Xue Yan, **Mengyue Yang**, Haifeng Zhang, Jun Wang, Yaodong Yang.

Publications before 2025:

- **InfoRank: Unbiased Learning-to-Rank via Conditional Mutual Information Minimization**. WWW 2024.
  Jiarui Jin, Zexue He, **Mengyue Yang**, Weinan Zhang, Yong Yu, Jun Wang, Julian McAuley.
- **Invariant Learning via Probability of Sufficient and Necessary Causes**. NeurIPS 2023 Spotlight.
  Mengyue Yang, Zhen Fang, Yonggang Zhang, Yali Du, Jean-Francois Ton, Jianhong Wang, Jun Wang.
- **Rectifying Unfairness in Recommendation Feedback Loops**. SIGIR 2023.
  Mengyue Yang, Jun Wang, Jean-Francois Ton.
- **Specify Robust Causal Representation from Mixed Observations**. KDD 2023.
  Mengyue Yang, Xinyu Cai, Furui Liu, Xu Chen, Zhitang Chen, Jianye Hao, Jun Wang.
- **Lending Interaction Wings to Recommender Systems with Plug-and-Play Conversational Agents**. NeurIPS 2023.
  Jiarui Jin, Xianyu Chen, Fanghua Ye, **Mengyue Yang**, Yue Feng, Weinan Zhang, Yong Yu, Jun Wang.
- **ChessGPT: Bridging Policy Learning and Language Modeling**. NeurIPS 2023.
  Xidong Feng, Yicheng Luo, Ziyan Wang, Hongrui Tang, **Mengyue Yang**, Kun Shao, David Mguni, Yali Du, Jun Wang.
- **Replace Scoring with Arrangement: A Contextual Set-to-Arrangement Framework for Learning-to-Rank**. CIKM 2023.
  Jiarui Jin, Xianyu Chen, Weinan Zhang, **Mengyue Yang**, Yang Wang, Yali Du, Yong Yu, Jun Wang.
- **Debiased Recommendation with User Feature Balancing**. ACM TOIS, 2022.
  Mengyue Yang, Guohao Cai, Furui Liu, Zhenhua Dong, Xiuqiang He, Jianye Hao, Jun Wang, Xu Chen.
- **Top-N Recommendation with Counterfactual User Preference Simulation**. CIKM 2021.
  Mengyue Yang, Quanyu Dai, Zhenhua Dong, Xu Chen, Xiuqiang He, Jun Wang.
- **CausalVAE: Disentangled Representation Learning via Neural Structural Causal Models**. CVPR 2021.
  Mengyue Yang, Furui Liu, Zhitang Chen, Jianye Hao, Jun Wang.
- **Deconfounding Representation Learning Based on User Interactions in Recommendation Systems**. PAKDD 2021.
  Junruo Gao, **Mengyue Yang**, Yuyang Liu, Jun Li.
- **Hierarchical Adaptive Contextual Bandits for Resource Constraint based Recommendation**. WWW 2020.
  Mengyue Yang, Qingyang Li, Zhiwei Qin, Jieping Ye.

Talks and Tutorials

Conference/Workshop/Symposium Invited Talks, Panels and Tutorials:

- AAAI 2026 Workshop on LLM-based Multi-Agent Systems: Panel discussion.
- AAAI 2026 New Faculty Highlight Program: Causal Foundation World Models.
- DAI 2025 Workshop on LLM-based Multi-Agent Systems: Invited keynote talk and panel discussion.
- ICML 2025 Workshop on Multi-Agent Systems: Panel discussion.
- ICML 2025 Workshop on Scaling Up Intervention Models Invited keynote talk: Scaling up Causal Learning in Open-Ended World
- IJTCS-FAW 2025 invited forum talk: From Games to Graphs: Causal Mechanisms for Steering Multi-Agent Outcomes.
- NILAB Workshop on AI for Bioscience: Learning Meta-Causal Worlds with Curious Agents.
- DAI 2024 Tutorial: Causality and Large Models.
- ACML 2024 Tutorial: Causality and Large Models.
- ICDM 2024 Tutorial: Causality and Large Models.
- Rising Star in AI symposium (KAUST): Causal Representation Learning.
- PCIC 2024: Invited talk on Essential Causal Representation Learning.

2025 Talks:

- Scaling-up Causal Learning: Shanghai Jiao Tong University and OMNI talk.
- Large Language Models and Reinforcement Learning: Fudan University.

2024 Talks:

- Causal Agent & Foundation Models: Jizhi Causality Reading Party (5th season), Renmin University of China, Peking University.
- Essential Causal Representation Learning via Probability of Sufficient and Necessary Causes:
  PCIC 2024, University of Manchester, University of Bristol, Imperial College London, University College London, AI Time, and DataFun Summit.

2022–2023 Talks:

- Causal Disentanglement Representation and Models. Paper Weekly, University College London, Beijing Institute of Technology, Jizhi Causality Reading Party, and Northeastern University.

Honours & Awards

- AAAI 2026 New Faculty Highlight.
- Rising Star in AI 2024 (KAUST).
- Third Prize, 2015 National University Students Computer Design Competition.
- First Prize, 2014 Youth Science Popularization Innovation Competition.
- Second Prize, 2011 Chinese Physics Olympiad (provincial level).

Academic Services

- Co-organizer, NeurIPS 2025 Workshop: *Embodied World Models*.
- Co-organizer, Causal AI workshop with IHES, France.
- Co-organizer and Guest Editor, Journal of Machine Learning Special Issue: *Build Trust in Foundation Models: Interpretability, Safety, and Robustness*.
- Organizer and Program Chair, ICLR 2025 Workshop: *World Models: Understanding, Modelling and Scaling* (More than 1,500 participants).
- Co-organizer, AAAI 2025 Workshop: *Artificial Intelligence with Causal Techniques*.
- Co-organizer, ICDM 2024 Workshop: *Causal Representation Learning*.
- Co-organizer, NeurIPS 2024 Workshop: *Causality and Large Model*.
- Co-organizer, NeurIPS 2023 Competition: *Causal Structure Learning from Event Sequences and Prior Knowledge*.
- Reviewer: TNNLS, TPAMI, KDD, NeurIPS, ICML, ICLR, AAAI, AISTATS, SDM, CVPR.

Teaching

- University of Bristol: Lecturer (2025-2026) SEMT30008 Advanced Methods in Artificial Intelligence (UG 3rd year).
- University of Bristol: Lecturer (2025-2026) SEMT10005 Principles of Artificial Intelligence (UG 1st year)).
- University of Bristol: Lecturer (2024-2026) SEMTM0016 Artificial Intelligence for Robotics (PG).
- University College London: Teaching Assistant (2021-2024) COMP0124 Multi-Agent Artificial Intelligence (PG).