

# Yifei He | Curriculum Vitae

✉ yifeihe3@illinois.edu • 🌐 yifei-he.github.io

I am interested in making foundation models and agents learn from multiple sources and tasks in an efficient, robust and scalable manner. In particular, I work on i) Multimodal agent for computer use. ii) Multi-domain learning for foundation models (model merging, multilingual LLMs). iii) LLM data (reward modeling) and inference efficiency (MoE). Previously, I have worked on multi-objective optimization, domain adaptation and multimodal learning.

## Education

### University of Illinois Urbana-Champaign (UIUC)

*Ph.D. in Computer Science*

**Urbana, IL, USA**

*Aug 2021 - May 2026 (Expected)*

### University of Michigan (UM)

*B.S.E. in Data Science, minor in Mathematics*  
Summa Cum Laude

**Ann Arbor, MI, USA**

*Aug 2019 - Apr 2021*

### Shanghai Jiao Tong University (SJTU)

*B.S.E. in Electrical and Computer Engineering*

**Shanghai, China**

*Sept 2017 - Aug 2021*

## Industry Experience

### Microsoft

*Applied Scientist Intern, Turing*

**Redmond, WA, USA**

*May 2025 - Aug 2025*

● Improved reasoning capabilities of computer-use agent.

*Aug 2024 - Feb 2025*

*Research Intern, GenAI*

● Improved efficiency of Mixture-of-Experts (MoE) models.

*Applied Scientist Intern, Turing*

*May 2024 - Aug 2024*

● Developed scaling laws for multilingual language models.

### Amazon

*Applied Scientist Intern, Search Science and AI*

**Seattle, WA, USA**

*May 2023 - Aug 2023*

● Improved large-scale multi-task tuning of foundation models.

● Developed a vision-language retrieval foundation model with instruction tuning.

## Publications (\* denotes equal contribution)

### [1] Scalable Data Synthesis for Computer Use Agents with Step-Level Filteringing.

Yifei He\*, Pranit Chawla\*, Yaser Souris, Subhrojit Som, Xia Song

*Under review.*

### [2] MergeBench: A Benchmark for Merging Domain-Specialized LLMs.

Yifei He, Siqi Zeng, Yuzheng Hu, Rui Yang, Tong Zhang, Han Zhao

*Neural Information Processing Systems Datasets and Benchmarks Track. (NeurIPS 2025)*

### [3] Efficiently Editing Mixture-of-Experts Models with Compressed Experts.

Yifei He, Yang Liu, Chen Liang, Hany Awadalla.

*Conference on Empirical Methods in Natural Language Processing 2025. (EMNLP 2025 Findings)*

### [4] Scaling Laws for Multilingual Language Models.

Yifei He, Alon Benhaim, Barun Patra, Praneetha Vaddamanu, Sanchit Ahuja, Parul Chopra, Vishrav Chaudhary, Han Zhao, Xia Song.

*Meeting of the Association for Computational Linguistics. (ACL 2025 Findings)*

- [5] **Efficient Model Editing with Task Vector Bases: A Theoretical Framework and Scalable Approach.**  
 Siqi Zeng, **Yifei He**, Weiqiu You, Yifan Hao, Yao-Hung Hubert Tsai, Makoto Yamada, Han Zhao.  
*Under review.*
- [6] **Towards Understanding the Fragility of Multilingual LLMs against Fine-Tuning Attacks.**  
 Samuele Poppi, Zheng-Xin Yong, **Yifei He**, Bobbie Chern, Han Zhao, Aobo Yang, Jianfeng Chi.  
*The Nations of the Americas Chapter of the Association for Computational Linguistics 2025. (NAACL 2025 Findings)*
- [7] **Localize-and-Stitch: Efficient Model Merging via Sparse Task Arithmetic.**  
**Yifei He**, Yuzheng Hu, Yong Lin, Tong Zhang, Han Zhao.  
*Transactions of Machine Learning Research. (TMLR, J2C Certificate)*
- [8] **Semi-Supervised Reward Modeling via Iterative Self-Training.**  
**Yifei He\***, Haoxiang Wang\*, Ziyan Jiang, Alexandros Papangelis, Han Zhao.  
*Conference on Empirical Methods in Natural Language Processing 2024. (EMNLP 2024 Findings)*
- [9] **Robust Multi-Task Learning with Excess Risks.**  
**Yifei He**, Shiji Zhou, Guojun Zhang, Hyokun Yun, Yi Xu, Belinda Zeng, Trishul Chilimbi, Han Zhao.  
*International Conference on Machine Learning. (ICML 2024)*
- [10] **Gradual Domain Adaptation: Theory and Algorithms.**  
**Yifei He\***, Haoxiang Wang\*, Bo Li, Han Zhao.  
*Journal of Machine Learning Research. (JMLR)*
- [11] **Efficient Modality Selection in Multimodal Learning.**  
**Yifei He\***, Runxiang Cheng\*, Gargi Balasubramaniam\*, Yao-Hung Hubert Tsai, Han Zhao.  
*Journal of Machine Learning Research. (JMLR)*  
 (Extended version of publication [12].)
- [12] **Greedy Modality Selection via Approximate Submodular Maximization.**  
 Runxiang Cheng\*, Gargi Balasubramaniam\*, **Yifei He\***, Yao-Hung Hubert Tsai, Han Zhao.  
*Conference on Uncertainty in Artificial Intelligence. (UAI 2022)*
- [13] **Conformer-RL: A Deep Reinforcement Learning Library for Conformer Generation.**  
 Runxuan Jiang, Tarun Gogineni, Joshua Kammeraad, **Yifei He**, Ambuj Tewari, Paul Zimmerman.  
*Journal of Computational Chemistry. (JCC)*
- [14] **A Hierarchical Approach to Multi-Event Survival Analysis.**  
 Donna Tjandra, **Yifei He**, Jenna Wiens.  
*AAAI Conference on Artificial Intelligence. (AAAI 2021)*

## Professional Service

---

**Reviewer:** UAI, NeurIPS, ICLR, AISTATS, ICML, TMLR, ACL

## Teaching Experience

---

Teaching assistant at UIUC

- CS 598 Foundations of Data Science 2025 Fall
- CS 357 Numerical Methods I 2022 Fall, 2022 Spring
- CS 441 Applied Machine Learning 2021 Fall

Teaching assistant at UM

- EECS 445 Intro to Machine Learning 2020 Fall

## Skills

---

**Framework:** TRL, PyTorch, DeepSpeed, TensorFlow, Keras, Gym