Ziyu Yao

Department of Computer Science
George Mason University
Nguyen Engineering Building 4415
Fairfax, VA 22030 USA

 ziyuyao.org ziyuyao@gmu.edu ZiyuYao GScholar

(Last update: March 7, 2025)

Research Interests

My research interests are mainly in three folds: (1) Machine Intelligence: advancing LLMs and LLM Agents in knowledge grounding, reasoning, and planning (exemplar tasks include math reasoning, code generation, and motion planning); (2) Responsible and Trustworthy Human-Machine Interface: studying human-AI interaction and LLM interpretability; (3) Interdisciplinary Applications: bring the impact of NLP/AI to critical domains such as Software Engineering and Education.

Education

2015–2021	**Ph.D.**, The Ohio State University (OSU), Computer Science & Engineering. (Thesis: On Advancing Natural Language Interfaces: Data Collection, Model Development, and User Interaction. Advisor: Prof. Huan Sun)
2011–2015	**B.E.**, Beijing University of Posts and Telecommunications (BUPT), Communication Engineering.

Professional Experience

2021–present	**Assistant Professor**, George Mason University (GMU). Affiliation: C4I & Cyber Center, Center for Advancing Human-Machine Partnership, and Institute for Digital InnovAtion (IDIA) at GMU
2021 Summer	**Research Intern**, Microsoft Corporation.
2020 Summer	**Research Intern**, Carnegie Mellon University.
2018 Summer	**Research Intern**, Microsoft Corporation.
2017 Summer	**Research Intern**, Fujitsu Lab of America.
2014–2015	**Research Intern**, Tsinghua University, Beijing.

Awards and Honors

2024	**Outstanding PhD Student Award (to Murong Yue)**, GMU CS.
2024	**Recognition for Service on Hangs-On Generative AI Tutorial**, GMU College of Engineering and Computing.
2023	**Top Reviewer**, NeurIPS.
2023	**Accelerating Foundation Models Research Program**, Microsoft Research.
2023	**Editor Commendation**, Journal of Business and Psychology.
2021	**Best Paper Award**, The IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
2021	**Graduate Student Research Award**, OSU, CSE.
2020	**Presidential Fellowship**, OSU.
2020	**Rising Star in EECS**, UC Berkeley.
2020	**Graduate Students Poster Award (Honorable Mention)**, CSE, OSU.
2019	**NSF Student Travel Award**, The Web Conference (WWW).
2019	**Graduate Students Poster Award (Honorable Mention)**, CSE, OSU.

2018	**Travel Grant for *Graduate Cohort for Women*, San Francisco**, CRA-W.
2015	**Beijing Outstanding Graduate Award**, Beijing, China.
2012, 2014	**First-class Scholarship (top 3%)**, BUPT, China.
2013	**Second-class Scholarship (top 5%)**, BUPT, China.

Publications

(<u>My Advisee</u>; †Equal Contribution).

PhD Dissertation.

Ziyu Yao, "On Advancing Natural Language Interfaces: Data Collection, Model Development, and User Interaction", The Ohio State University, **PhD Dissertation, 2021**. [Dissertation]

Preprints.

5   Mohamed Aghzal, Erion Plaku, Gregory J Stein, and Ziyu Yao, "A Survey on Large Language Models for Automated Planning", **ArXiv, 2024**. [Paper]

4   <u>Daking Rai</u>, Rydia R Weiland, Kayla Margaret Gabriella Herrera, Tyler H Shaw, and **Ziyu Yao**, "Understanding the Effect of Algorithm Transparency of Model Explanations in Text-to-SQL Semantic Parsing", **ArXiv, 2024**. [Paper]

3   <u>Daking Rai</u>, Yilun Zhou, Shi Feng, Abulhair Saparov, and **Ziyu Yao**, "A Practical Review of Mechanistic Interpretability for Transformer-Based Language Models", **ArXiv, 2024**. [Paper][GitHub Awesome-Mech-Interp-LM][X/Twitter]

2   <u>Hao Yan</u>, Thomas D. Latoza, and **Ziyu Yao**, "IntelliExplain: Enhancing Conversational Code Generation for Non-Professional Programmers", **ArXiv, 2024**. [Paper][Website]

1   Qineng Wang, Chen Qian, Xiaochang Li, **Ziyu Yao**, and Huajie Shao, "Lens: A Foundation Model for Network Traffic", **ArXiv, 2024**. [Paper]

Accepted to Refereed Journals.

2   Philseok Lee, Shea Fyffe, Mina Son, Zihao Jia, and **Ziyu Yao**, "A paradigm shift from "human writing" to "machine generation" in personality test development: An application of state-of-the-art natural language processing", **Journal of Business and Psychology (2023) 38:163–190, IF 6.604 (Editor Commendation, one of the 13 selected papers out of over 1,000 submissions in 2022)**. [Paper]

1   **Ziyu Yao**, Shouzhong Tu, Minlie Huang, Xiaoyan Zhu. "A Semi-Supervised Method for Filtering Chinese Spam Tweets". **Journal of Chinese Information Processing 2016**, 30(5): 176-186. [Paper in Chinese]

Accepted to Refereed Conference Proceedings.

23   <u>Mohamed Aghzal</u>, Xiang Yue, Erion Plaku, and **Ziyu Yao**, "Evaluating Vision-Language Models as Evaluators in Path Planning", To appear at **CVPR 2026**. [Paper]

22   <u>Murong Yue</u>, Wenlin Yao, Haitao Mi, Dian Yu, **Ziyu Yao**, and Dong Yu, "DOTS: Learning to Reason Dynamically in LLMs via Optimal Reasoning Trajectories Search", In *The Thirteenth International Conference on Learning Representations*. **(ICLR 2025)** [Paper]

21 Yuqing Zhou, Ruixiang Tang, **Ziyu Yao**, and Ziwei Zhu, "Navigating the Shortcut Maze: A Comprehensive Analysis of Shortcut Learning in Text Classification by Language Models", In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 2586–2614, Miami, Florida, USA. Association for Computational Linguistics. **(EMNLP Findings 2024)** [Paper][Code]

20 Mohamed Aghzal, Erion Plaku, and **Ziyu Yao**, "Look Further Ahead: Testing the Limits of GPT-4 in Path Planning." In *2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)*, pp. 1020-1027. IEEE, 2024. **(IEEE CASE 2024)** [Paper][Code]

19 Daking Rai and **Ziyu Yao**, "An Investigation of Neuron Activation as a Unified Lens to Explain Chain-of-Thought Eliciting Arithmetic Reasoning of LLMs", In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7174–7193, Bangkok, Thailand. Association for Computational Linguistics. **(ACL 2024)** [Paper][Code]

18 Saurabh Srivastava, Chengyue Huang, Weiguo Fan, and **Ziyu Yao**, "Instance Needs More Care: Rewriting Prompts for Instances Yields Better Zero-Shot Performance", In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 6211–6232, Bangkok, Thailand. Association for Computational Linguistics. **(ACL Findings 2024)** [Paper][Code]

17 Murong Yue, Jie Zhao, Min Zhang, and Liang Du, **Ziyu Yao**, "Large Language Model Cascades with Mixture of Thoughts Representations for Cost-efficient Reasoning", In *The Twelfth International Conference on Learning Representations*. **(ICLR 2024)** **Featured in Hugging Face Daily Papers on Oct 5, 2023** [Paper][Code]

16 Saurabh Srivastava, Gaurav Singh, Shou Matsumoto, and Ali Raz, Paulo Costa, Joshua Poore, **Ziyu Yao**, "MAILEX: Email Event and Argument Extraction". In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 12964–12987, Singapore. Association for Computational Linguistics. **(EMNLP 2023, acceptance rate 23.3%)**. [Paper][Dataset&Code]

15 Binfeng Xu, Xukun Liu, Hua Shen, and Zeyu Han, Yuhan Li, Murong Yue, Zhiyuan Peng, Yuchen Liu, **Ziyu Yao**, Dongkuan Xu, "Gentopia: A Collaborative Platform for Tool-Augmented LLMs". In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 237–245, Singapore. Association for Computational Linguistics. **(EMNLP 2023 Demo-Track)** [Paper][Code][Project Website]

14 Hao Yan, Saurabh Srivastava, Yintao Tai, Sida I. Wang, Scott Wen-tau Yih, **Ziyu Yao**, "Learning to Simulate Natural Language Feedback for Interactive Semantic Parsing". In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3149–3170, Toronto, Canada. Association for Computational Linguistics. **(ACL 2023, acceptance rate: 20.73%)**. [Paper][Code]

13 Daking Rai, Bailin Wang, Yilun Zhou, **Ziyu Yao**, "Improving Generalization in Language Model-based Text-to-SQL Semantic Parsing: Two Simple Semantic Boundary-based Techniques". In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 150–160, Toronto, Canada. Association for Computational Linguistics. **(ACL 2023, acceptance rate: 20.73%)**. [Paper][Code]

12 Daking Rai, Yilun Zhou, Bailin Wang, **Ziyu Yao**, "Explaining Large Language Model-based Neural Semantic Parsers" (Student Abstract). In *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(13), 16308-16309. https://doi.org/10.1609/aaai.v37i13.27014 **(AAAI 2023 Student Abstract)**.[Paper]

11 Tianbao Xie[†], Chen Henry Wu[†], Peng Shi, Ruiqi Zhong, Torsten Scholak, Michihiro Yasunaga, Chien-Sheng Wu, Ming Zhong, Pengcheng Yin, Sida Wang, Victor Zhong, Bailin Wang, Chengzu Li, Connor Boyle, Ansong Ni, **Ziyu Yao**, Dragomir Radev, Caiming Xiong, Lingpeng Kong, Rui Zhang, Noah A. Smith, Luke Zettlemoyer, Tao Yu, "UnifiedSKG: Unifying and Multi-Tasking Structured Knowledge Grounding with Text-to-Text Language Models". In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 602–631, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. (**EMNLP 2022, acceptance rate: 22.1%**). [Paper, Code, Project Website]

10 Xiang Yue, **Ziyu Yao**, Huan Sun, "Synthetic Question Value Estimation for Domain Adaptation of Question Answering". In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1340–1351, Dublin, Ireland. Association for Computational Linguistics. (**ACL 2022, acceptance rate: 20.75%**). [Paper, Code]

9   Xiang Yue, Xinliang Frederick Zhang, **Ziyu Yao**, Simon Lin, Huan Sun, "CliniQG4QA: Generating Diverse Questions for Domain Adaptation of Clinical Question Answering". *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Houston, TX, USA, 2021, pp. 580-587, doi: 10.1109/BIBM52615.2021.9669300. (**IEEE BIBM 2021, Best Paper Award**). [Paper, Code]

8   **Ziyu Yao**, Frank Xu, Pengcheng Yin, Huan Sun, Graham Neubig, "Learning Structural Edits via Incremental Tree Transformations". In *The Ninth International Conference on Learning Representations* (**ICLR 2021, acceptance rate: 28.7%**). [Paper, Code]

7   **Ziyu Yao**, Yiqi Tang, Scott Wen-tau Yih, Huan Sun, Yu Su, "An Imitation Game for Learning Semantic Parsers from User Interaction", In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6883–6902, Online. Association for Computational Linguistics. (**EMNLP 2020, acceptance rate: 22.4%**). [Paper, Code]

6   **Ziyu Yao**, Yu Su, Huan Sun, Scott Wen-tau Yih, "Model-based Interactive Semantic Parsing: A Unified Formulation and A Text-to-SQL Case Study". In *2019 Conference on Empirical Methods in Natural Language Processing* (**EMNLP 2019, acceptance rate: 23.8%**). [Paper, Code]

5   Boyuan Pan, Hao Li, **Ziyu Yao**, Deng Cai, Huan Sun, "Reinforced Dynamic Reasoning for Conversational Question Generation". In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (**ACL 2019, acceptance rate: 25.7%**). [Paper, Code]

4   **Ziyu Yao**, Jayavardhan Reddy Peddamail, Huan Sun, "CoaCor: Code Annotation for Code Retrieval with Reinforcement Learning", The Web Conference 2019 (**WWW 2019, acceptance rate: 18%**). [Paper, Code]

3   **Ziyu Yao**, Xiujun Li, Jianfeng Gao, Brian Sadler, Huan Sun, "Interactive Semantic Parsing for If-Then Recipes via Hierarchical Reinforcement Learning", The AAAI Conference on Artificial Intelligence 2019 (**AAAI 2019, acceptance rate: 16.2%; SPOTLIGHT**). [Paper, Code]

2   **Ziyu Yao**, Daniel S. Weld, Wei-Peng Chen, Huan Sun, "StaQC: A Systematically Mined Question-Code Dataset from Stack Overflow", The Web Conference 2018 (**WWW 2018, acceptance rate: 14.8%**). [Paper, Code]

1   Li Zhao, Minlie Huang, **Ziyu Yao**, Rongwei Su, Yingying Jiang, Xiaoyan Zhu. "Semi-Supervised Multinomial Naive Bayes for Text Classification by Leveraging Word-Level Statistical Constraint", The AAAI Conference on Artificial Intelligence 2016 (**AAAI 2016, acceptance rate: 25.8%**). [Paper]

Accepted to Refereed Workshop Proceedings.

5 <u>Murong Yue</u>, Wenhan Lyu, <u>Wijdane Mifdal</u>, Yixuan Zhang, Jennifer Suh, and **Ziyu Yao**, "MathVC: An LLM-Simulated Multi-Character Virtual Classroom for Mathematics Education", **AAAI AI4Edu Workshop, 2025 (Invited for presentation at Wolfram Research LLM Agents Colloquium)**. [Paper][Website]

4 <u>Sam Miller</u>, <u>Daking Rai</u>, and **Ziyu Yao**, "Mechanistic Understanding of Language Models in Syntactic Code Completion", **AAAI Know4FM Workshop, 2025**. [Paper]

3 <u>Mohamed Aghzal</u>, Erion Plaku, and **Ziyu Yao**, "Can Large Language Models be Good Path Planners? A Benchmark and Investigation on Spatial-temporal Reasoning", **ICLR Workshop on LLM Agents, 2024**. [Paper][Code]

2 Peizhao Li, Xuchao Zhang, **Ziyu Yao**, Wei Cheng, Haifeng Chen, Hongfu Liu, "Code Editing from Few Exemplars by Adaptive Multi-Extent Composition", *The Deep Learning for Code Workshop at The Tenth International Conference on Learning Representations* (**ICLR 2022 DL4Code**). [Paper]

1 Jayavardhan Reddy Peddamail, **Ziyu Yao**, Zhen Wang, Huan Sun, "A Comprehensive Study of StaQC for Deep Code Summarization", **SIGKDD 2018 Deep Learning Day**. [Paper]

Talks & Presentations

2024 **Mechanistic Understanding of LLMs in Reasoning and Code Generation**, <u>Invited talk</u> at University of Cambridge [Publicity].

2024 **MathVC: An LLM-Simulated Multi-Character Virtual Classroom for Mathematics Education**, <u>Invited talk</u> at Wolfram Research, Online Colloquium for LLM Agents for Modeling Group Dynamics [Colloquium Website].

2024 **The Progress, Opportunities, and Challenges of Semantic Parsing**, <u>Invited talk</u> at CTO Data Science Speaker Series at Bloomberg. Host: Daniel Preotiuc-Pietro [Twitter/X Cover].

2024 **Towards Understanding and Modeling the Interaction with Large Language Models**, <u>Invited talk</u> at New York University. Host: Abulhair Saparov.

2024 **Interactive Semantic Parsing**, <u>Invited talk</u> at Microsoft Research (PROSE team). Host: José Cambronero, Sumit Gulwani.

2024 **Building Natural Language Interfaces in the Age of Large Language Models**, <u>Invited talk</u> at Indiana University, Indianapolis. Host: Ming Jiang.

2024 **Towards Enhancing the Utilization of Large Language Models for Humans**, <u>Invited talk</u> at Virginia Tech. Host: Chang-Tien Lu.

2024 **Towards Enhancing the Utilization of Large Language Models for Humans**, <u>Invited talk</u> at University of Arizona. Host: Steven Bethard.

2023 **Building Natural Language Interfaces in the Age of Large Language Models**, <u>Invited talk</u> at Tsinghua University. Host: Minlie Huang.

2023 **Gentopia: A Collaborative Platform for Tool-Augmented LLMs**, <u>Demo Presentation</u> at EMNLP.

2023 **Large Language Model Cascades with Mixture of Thoughts Representations for Cost-efficient Reasoning**, <u>Invited talk</u> at Microsoft Semantic Machines. Host: Hao Fang.

2023 **Building Natural Language Interfaces in the Age of Large Language Models**, <u>Invited talk</u> at Google NLP Reading Group. Host: Pengcheng Yin.

2023 **Building Natural Language Interfaces in the Age of Large Language Models**, <u>Invited talk</u> at Apple Inc. Host: Yunyao Li.

2023 **Information Acquisition in the Age of Language Models**, Invited talk at GMU C4I & Cyber Center.

2022 **Challenges in Constructing Human-centric Natural Language Interfaces**, Invited talk at the GMU Statistics Department. Host: Nicholas Rios.

2022 **On Building Natural Language Interfaces for "Humans"**, Invited talk at ServiceNow Research, Canada (virtually). Host: Dzmitry Bahdanau.

2022 **Building Interactive and Adaptive Question Answering Systems**, Invited talk at CLSP (Center for Language and Speech Processing) Seminar, Johns Hopkins University. Host: Kevin Duh, Ben Van Durme.

2022 **Building Interactive and Adaptive Question Answering Systems**, Invited talk at Center for Advancing Human-Machine Partnership, George Mason University..

2021 **Building Interactive Natural Language Interfaces**, Invited talk at Brandeis University (virtually). Host: Hongfu Liu.

2021 **Building Interactive Natural Language Interfaces**, OSU AI Club (virtually).

2020 **An Imitation Game for Learning Semantic Parsers from User Interaction**, EMNLP (virtually).

2020 **On Building an Interactive Natural Language Interface: An intelligent agent that can communicate, collaborate and learn**, Invited talk at Microsoft Research AI Breakthroughs Workshop (virtually). Invited by: Ahmed Hassan Awadallah, Alex Polozov.

2020 **Building Natural Language Interfaces to Programs**, Invited talk at VMware, Beijing (virtually). Host: Jin He, Kevin Song.

2020 **Towards Building Interactive and Collaborative Natural Language Interfaces**, Invited talk at NeuLab, Carnegie Mellon University (virtually). Host: Graham Neubig.

2020 **Learning Semantic Parsers from User Interaction**, CSE AI Seminar, OSU.

2019 **Towards Building Interactive and Collaborative Natural Language Interfaces**, Invited talk at Chair of Management Information System, ETH Zurich, Host: Stefan Feuerriegel.

2019 **CoaCor: Code Annotation for Code Retrieval with Reinforcement Learning**, WWW, San Francisco, USA.

2019 **Interactive Semantic Parsing for If-Then Recipes via Hierarchical Reinforcement Learning**, AAAI, Hawaii, USA.

2018 **Interactive Semantic Parsing for If-Then Recipes via Hierarchical Reinforcement Learning**, CSE AI Seminar, OSU.

2018 **StaQC: A Systematically Mined Question-Code Dataset from Stack Overflow**, WWW, Lyon, France.

Research Funding

8 **"Collaborative Research: Fostering Mathematical Modeling Competencies through Collaborative Learning in a Large Language Model (LLM) Simulated Virtual Classroom"**, National Science Foundation, $899,920 (CMU: $578,047, my part: $300,000), 09/2024-08/2027. Role: PI (lead institution).

7 **"Securing the Supply Chain of LLMs as Software"**, Virginia Commonwealth Cyber Initiative, $49,997 (GMU:$49,997, my part: ~$40,000) without overhead, 01/2024-01/2025. Role: PI.

6 **"A Foundation Model-Simulated Virtual Classroom for PreK-12 STEM Education"**, Microsoft Research, Accelerating Foundation Models Research Program, $20,000 Azure credits, 10/2023-01/2024. Role: PI.

5 **"Collaborative Research: SHF: Medium: Toward Understandability and Interpretability for Neural Language Models of Source Code"**, National Science Foundation, $1,074,380 (GMU:$745,197, my part: $391,609), 10/2023-09/2027. Role: Co-PI. (PI: Kevin Moran at GMU → UCF)

4 **"ARLIS Support for IARPA's STG AI-ML Solutions Program"**, Intelligence Advanced Research Projects Activity/UMD ARLIS, $100,000 (GMU:$100,000, my part: ~$30,000), 10/2022-03/2024. Role: Co-PI. (PI: Ali Raz at GMU)

3 **"NSF Convergence Accelerator Track G: Secure Texting over Non-cooperative Networks and Anti-jamming Enhancement in 5G"**, National Science Foundation, 09/2022-07/2023. Role: Senior Personnel. (PI: Kai Zeng at GMU)

2 **"Securing Interactions Between Humans and Machines"**, Virginia Commonwealth Cyber Initiative, $59,971 (GMU:$59,971, my part: ~$30,000) without overhead, 12/2022-12/2023. Role: PI.

1 **"AI Engineering Seedlings"**, US Department of Defense/UMD ARLIS, $100,000 (GMU:$100,000, my part: ~$30,000), 07/2021–07/2022. Role: Co-PI. (PI: Ali Raz at GMU)

Teaching Experience

At GMU:
*denotes newly developed courses

2024 Fall	**CS478: Natural Language Processing**.
2024 Spring	**CS678: Advanced Natural Language Processing** [Website].
2023 Fall	**CS678: Advanced Natural Language Processing** [Website].
2023 Spring	**CS478*: Natural Language Processing** [Website].
2022 Fall	**CS678*: Advanced Natural Language Processing** [Website].
2022 Spring	**CS499: Natural Language Processing** [Website].
2021 Fall	**CS695: Special Topics on Natural Language Processing** [Website].

At OSU:

2020 Fall	**CSE Panel: Graduate School Experience and Expectation**, CSE, OSU. Panelist.
2020 Fall	**Foundations of Speech and Language Processing**, CSE, OSU. Guest lecturer.
2018 Fall	**Course Panel: Data Management in the Cloud**, CSE, OSU. Guest panelist.
2017 Spring	**Introduction to Computer Programming in C++**, CSE, OSU. Student instructor.

Mentoring Experience

Mentored PhD Students.
Hao Yan (Spring'22-)
Daking Rai (Spring'22-)
Saurabh Srivastava (Spring'22-)
Murong Yue (Fall'22-; Outstanding PhD Student at GMU CS, 2024)
Mohamed Aghzal (Spring'23-)

Mentored/Collaborated Master Students.
Sweta Pati (Summer'24-)
Samuel Miller (Summer'24-)
Gaurav Singh (Spring'22)
Janit Bidhan (Fall'22)

Mentored/Collaborated Undergraduate Students.

Sai Adith Senthil Kumar (2024 Spring-, funded by VA CCI)

Saipavan Perepa (2024 Spring-, funded by VA CCI)

Wijdane Mifdal (Fall'23-Spring'24 through OSCAR FW)

Mariana Ritchie (Spring'22)

Brian Meike (Spring'22)

Soumithri Gadepalli (Spring'22)

Yintao Tai (Fall'21, remote intern from Univ. of Edinburgh)

Mentored/Collaborated High School Students.

Angela Xiang (Summer'24, Montgomery Blair High School, through ASSIP)

Sarah Li (Summer'24, McLean High School, through ASSIP

Mihir Kulshreshtha (Summer'23, Thomas Jefferson High School for Science and Technology, through ASSIP)

Anika Giri (Summer'23, Chantilly High School)

(Co-)Mentored/Collaborated Students at OSU.

Jayavardhan Reddy Peddamail (Master Student; now at Google)

Varun Ranganathan Dhanvanth (Master Student; now at Facebook)

Heming Sun (Undergraduate Student)

Tianshu Zhang (PhD Student)

Leslie Zhou (Master Student)

School/Department Services

Department Committees.

SWE Master Recruiting & TT Faculty Recruiting Committee (Fall'23 - Spring'24)

PhD Recruiting Committee (Fall'22 - Spring'23)

Faculty Search Committee (Fall'21 - Spring'22)

Comprehensive Exam Committees.

Joshua Otten (Fall'24)

Nathaniel Krasner (Spring'24)

Murong Yue (Spring'24)

Saurabh Srivastava (Spring'24)

Daking Rai (Fall'23)

Hao Yan (Fall'23)

Syeda Sabrina Akter (Fall'23)

Milind Agarwal (Fall'23)

Fahim Faisal (Fall'22)

Jonathan Mbuya (Fall'22)

Joost Bottenbley (Fall'22)

Md Mahfuz Ibn Alam (Spring'22)

Abhisekh Rana (Spring'22)

Proposal Defense Committees.

Abhishek Paudel (Fall'24)

Milind Agarwal (Fall'24)

Jonathan Mbuya (Summer'24)

Fahim Faisal (Spring'24)

Kai North (Spring'24)

Anowarul Kabir (Spring'24)

Md Mahfuz Ibn Alam (Spring'23)

Thesis Defense Committees.

Kai North (Summer'24)

Md Mahfuz Ibn Alam (Summer'24)

Professional Activities

Outreach.

Faculty Mentor, CRA Undergraduate Research to PhD (UR2PhD)	2024
Invited Guest Lecturer at GMU School of Education	Spring 2024
Invited Guest Lecturer at GMU Department of Health Administration and Policy	Spring 2024
Invited Speaker at Morgan State University AI Camp	2023
Invited Speaker at "Girls Who Code" Chapter at Chantilly High School	2023
Panelist at NSF-funded Webinar about "Securely Operating Through 5G Infrastructure"	2023
Faculty Judge at the Science and Engineering Fair at Thomas Jefferson High School for Science and Technology	2023
Faculty Mentor at SIGKDD Undergraduate Consortium	2022

Panel Reviewer.

NSF: National Science Foundation	2023 –

Editor Services.

ACM SIGAI: Newsletter co-editor	2022 –

Organizing Committees.

NAACL'24: 2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Diversity & Inclusion Chair)	2024
MASC-SLL: 10th Mid-Atlantic Student Colloquium on Speech, Language and Learning (GMU's first-ever lead of this annual student event; we achieved a record number of 66 submissions and 130 attendees from ~30 organizations.)	2023
SUKI at NAACL'22: Workshop on Structured and Unstructured Knowledge Integration	2022
NLP4Prog at ACL'21: The First Workshop on NLP for Programming	2021

Area Chairs or Senior Program Committees.

ARR: ACL Rolling Review (This is an open review platform for papers submitted to ACL, EMNLP, NAACL, etc., starting from 2023)	2023 –
ACL: Annual Meeting of the Association for Computational Linguistics	2023 –
EMNLP: Conference on Empirical Methods in Natural Language Processing	2022 –
AAAI: The AAAI Conference on Artificial Intelligence	2023,2024

Session Chair.

AAAI: The AAAI Conference on Artificial Intelligence	2023

Program Committees/Reviewers (Conference).

ARR: ACL Rolling Review	2021 –
EACL SRW: EACL Student Research Workshop	2023
ICLR: International Conference on Learning Representations	2024
NeurIPS: Annual Conference on Neural Information Processing Systems	2023
NAACL: Annual Conference of the North American Chapter of the Association for Computational Linguistics	2021
EMNLP: Conference on Empirical Methods in Natural Language Processing 2018, 2020,	2021
MSR: Mining Software Repositories Conference	2021
AACL: Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics	2020
AAAI: The AAAI Conference on Artificial Intelligence	2020
NLPCC: Conference on Natural Language Processing and Chinese Computing	2020
CoNLL: Conference on Computational Natural Language Learning	2019

Secondary/External Reviewer:

ACL: Annual Meeting of the Association for Computational Linguistics	2018, 2020
KDD: SIGKDD Conference on Knowledge Discovery and Data Mining	2017, 2019, 2020
NAACL: Annual Conference of the North American Chapter of the Association for Computational Linguistics	2019
ICDM: International Conference on Data Mining	2016, 2018
WWW: International World Wide Web Conference	2017
CIKM: International Conference on Information and Knowledge Management	2016

Program Committees/Reviewers (Workshop).

SRW at NAACL 2025: NAACL Student Research Workshop	2025
TRL at NeurIPS 2024: 3th Table Representation Learning Workshop	2024
NLP4PI at EMNLP 2024: Workshop on NLP for Positive Impact	2024
InterNLP at NeurIPS 2022: Workshop on Interactive Learning for Natural Language Processing	2022
BigData-IT at IEEE BigData 2022: Workshop on Knowledge Discovery and Data Mining on IT Operations	2022
HCI+NLP at NAACL 2022: Bridging Human-Computer Interaction and Natural Language Processing	2022
DL4Code at ICLR 2022: Deep Learning for Code	2022
IntEx-SemPar at EMNLP 2020: Interactive and Executable Semantic Parsing Workshop	2020
NLI at ACL 2020: Natural Language Interface Workshop	2020
SRW at ACL 2019: Student Research Workshop	2019

Reviewers (Journal).

Computational Linguistics: Computational Linguistics, MIT Press	2024 –
TACL: Transactions of the Association for Computational Linguistics	2024 –
TKDD: ACM Transactions on Knowledge Discovery from Data	2022 –
TKDE: IEEE Transactions on Knowledge and Data Engineering	2022 –