Siyang Liu

Email: lsiyang@umich.edu, liusyang641@gmail.com Mobile: +1-734-8347522
Google Scholar: scholar.google.com/citations?user=2OjUAPUAAAAJ

Education

- **University of Michigan** Ann Arbor, US
  Ph.D. student - Computer Engineering; *September 2022 - present*
  Courses: *NLP, Advance AI, HCI*

- **Tsinghua University** Beijing, China
  Master of Science - Computer Science and Technology; *September 2018 - June 2021*
  Courses: *Learning from Data, Big Data Analytics, Computational Photography, Computer Vision*

- **Northeastern University** Shenyang, China
  Bachelor of Management - Accounting; *September 2012 - June 2016*
  Courses: *Introduction to Computer Programming(C Programming Language), Data Structure, Computer Organization Principle, Psychology, Probability Theory and Mathematical Statistics*

Experience

(1-year full-time engineer experience; 3 internships; 2 lab experiences;)

- **Language and Information Technologies Group** Ann Arbor
  Graduate Student Research Assistant (Full-time) Advisor: Rada Mihalcea *September 2022 - Present*
- **Educational Testing Service** Princeton
  Ida Lawrence Research Summer Internship. Mentors: Jessica Andrews Todd and Yang Jiang *Summer 2025*
- **Tencent America** Bellevue
  Research Intern at TEG AI lab. Advisors: Dian Yu and Xiaoyang Wang *Summer 2024*
- **Kuaishou Inc.** Beijing
  Algorithm Engineer at Multi Modality Understanding Group (Full-time) *July 2021 - July 2022*
- **Beijing National Research Center for Information Science and Technology** Beijing
  Master Student (Full-time) Advisors: Minlie Huang and Yong Jiang *October 2019 - June 2021*

Publications

(7 first-authored papers; citations $\sim$ 880 until 2025 July)

- **Patient-Centered RAG for Oncology Visit Aid Following the Ottawa Decision Guide**: Siyang Liu, Lawrence Chin-I An, Rada Mihalcea. EMNLP System Demonstrations (Under Review), 2025.
- **Free Lunch for User Experience: Crowdsourcing Agents for Scalable User Studies**: Siyang Liu, Sahand Sabour, Xiaoyang Wang, Rada Mihalcea. CSCW (Under Review), 2025.
- **Eeyore: Realistic Depression Simulation via Expert-in-the-loop Supervised and Preference Optimization**: Siyang Liu, Bianca Bie, Wenda Li, Laura Biester, Andrew Lee, James Pennebaker, Rada Mihalcea. ACL Findings, 2025.
- **Human Decision-making is Susceptible to AI-driven Manipulation**: Sahand Sabour, June M Liu, Siyang Liu, Chris Z Yao, Shiyao Cui, Xuanming Zhang, Wen Zhang, Yaru Cao, Advait Bhat, Jian Guan, Wei Wu, Rada Mihalcea, Tim Althoff, Tatia Lee, Minlie Huang. Nature Communications (Under Review), 2024.
- **The Generation Gap: Exploring Age Bias in the Value Systems of Large Language Models**: Siyang Liu, Trisha Maturi, Siqi Shen, Rada Mihalcea. Main@EMNLP, 2024.
- **EmoBench: Evaluating the Emotional Intelligence of Large Language Models**: Sahand Sabour, Siyang Liu, et al. The 62nd Annual Meeting of the Association for Computational Linguistics (Main@ACL), 2024.
- **Task-Adaptive Tokenization: Enhancing Long-Form Text Generation Efficacy in Mental Health and Beyond**: Siyang Liu, Naihao Deng, Sahand Sabour, Minlie Huang, Rada Mihalcea. Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (Main@EMNLP), 2023.
- **Rethinking and Refining the Distinct Metric**: Siyang Liu*, Sahand Sabour*, Yinhe Zheng, Pei Ke, Xiaoyan Zhu, Minlie Huang. The Joint Conference of the 60th Annual Meeting of the Association for Computational Linguistics (Main@ACL), 2022.
- **Towards Emotional Support Dialogue Systems**: Siyang Liu*, Chujie Zheng*, Orianna Demasi, et al. The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Main@ACL-IJCNLP), 2021.
- **PsyQA: A Chinese Dataset for Generating Long Counseling Text for Mental Health Support**: Hao Sun, Zhenru Lin, Chujie Zheng, Siyang Liu and Minlie Huang. The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Findings (Findings@ACL-IJCNLP'2021),2021.)
- **Sentilare: Linguistic Knowledge Enhanced Language Representation for Sentiment Analysis**: Pei Ke, Haozhe Ji, Siyang Liu, et al. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020, 6975–6988.

Projects

(skill keywords: simulation, agents, RAG, post-training, alignment; domains: healthcare, education, and game)

- **Web Agent Framework Supporting Real-time UI Interaction and Role-Execution Hierarchy**: Contributing an open-sourced web agent tool to facilitate researchers in this field (Work in progress)
- **Simulating Participants: Experiencing Digital Assessment with Web Agent in Role**: (Internship Project) Developed an architecture with modular student role and web task performer that helps usability-test or generate student experiences with ETS's digital assessment systems (July '2025)
- **Towards Tutor-style GUI Bot: Helping Less Digitally Experienced Generation Integrate into the App World**: transfer web **agents** from automation to tutoring to foster an all-user friendly digital world (Work in progress)
- **Cancer AI Decision Coach with Personalized Retrieval-Augmented Generation**: Integrating Ottawa Decision Guide with **Retrieval-Augmented Generation (RAG)** to assist patients in preparing for outpatient visit and recommend tailored questions based on their disease-related background. (July '25)
- **Agentic H-CI: Crowdsourcing Agents to Experience and Evaluate**: (Internship Project) Exploring how AI agents can facilitate user experience designers in product iteration. We **simulated** diverse user backgrounds to serve as proxies for collecting extensive experience data on interactions with NPCs in games (May '25)
- **Eeyore: Realistic Depression Simulation via Supervised and Preference Optimization**: (leverage **RLHF** along with a tailored **alignment** mechanism, designed with input from psychologists, to shift the model's focus from an alignment with generic human values to an emphasis on a more precise representation of depressive traits. (Feberuary '25)
- **The Generation Gap: Exploring the Age Bias in the Value Systems of Large Language Models**: investigate the **alignment** in the values of LLMs with different age groups. (June '24)
- **Task-Adaptive Tokenization: Enhancing Long-Form Text Generation Efficacy in Mental Health and Beyond** : Achieve better generation efficacy by applying task-adaptive tokenizer for **post-training** (August '22)
- **Graph Neural Networks for Recommendation Systems in Industry (Graph models, Recommendation Systems)**: Large-scaled, heterogeneous U2I Graph model for live recommendation. Tech: Python, TensorFlow, SQL (November '21)
- **Improving an Evaluation Method for Diverse Response Generation (Dialog Systems, Diverse Generation)**: Highlight the shortcomings of a popular metric and refine it. (August '21)
- **Building Emotional Support Dialog Systems (Dialog Systems, Language Generation, Psycholinguistics)**: Proposed an emotional support framework and built an emotional support dialog system. (December '2020)
- **PsyQa: Generating Counseling Texts for Mental Health (Text Generation, Question Answering, Psycholinguistics)**: Scratch data and annotate texts according to discourse structure. (September '2020)
- **Pre-trained Representation Learning Enhanced with Sentiment Knowledge (Representation learning, Sentiment Analysis)**: Training Sentilare under sentence-level and Aspect-level sentiment classification tasks (November '2019)

Activities and Honors

- **Peer Review**: Journal, Nature Communications, ISSN: 2041-1723
- **Actively Peer-reviewing for the Association of Computational Linguistics**: completing in total of 40 reviews (Present)
- **University of Michigan AI Network Coordinator**: one-year host (2025)
- **NLP@Michigan Day**: event organizer (March 2025)
- **Building Bridges in CSE PhD Studies**: giving research talks & promote the graduate program on behalf of CSE department (April, 2024)
- **NLP Reading Group**: acting as a host and coordinator for students who would like to read and present the most recent papers in the field of natural language processing every week at University of Michigan (Summer, 2023)
- **Qingfeng Front-loaded Research Scholarships**: being rewarded for inter-discipline research between material chemistry and machine learning (May, 2020)