

Yueqi Song (She/Her/Hers)

EDUCATION

PhD, Carnegie Mellon University

Aug 2025 - 2028/2029 (Expected Graduation)

Language and Information Technologies

Advisor: [Professor Graham Neubig](#)

BS/MS, Carnegie Mellon University

Aug 2020 - Aug 2025

Dual Degree in Computer Science, Stats and Machine Learning (Honors)

Research Advisor: [Professor Graham Neubig](#)

Academic Advisor: [Professor Zeliha Dilsun Kaynar](#)

PUBLICATIONS

[14] Agent Data Protocol: Unifying Datasets for Diverse, Effective Fine-tuning of LLM Agents

Yueqi Song, Ketan Ramaneti, Zaid Sheikh, Ziru Chen, Boyu Gou, Tianbao Xie, Yiheng Xu, Danyang Zhang, Apurva Gandhi, Fan Yang, Joseph Liu, Tianyue Ou, Zhihao Yuan, Frank Xu, Shuyan Zhou, Xingyao Wang, Xiang Yue, Tao Yu, Huan Sun, Yu Su, Graham Neubig

Under Review; Link: <https://arxiv.org/abs/2510.24702>

[13] Grounding Multilingual Multimodal LLMs With Cultural Knowledge

Jean de Dieu Nyandwi, **Yueqi Song**, Simran Khanuja, Graham Neubig

EMNLP 2025 (Oral); Link: <https://arxiv.org/abs/2508.07414>

[12] Synthetic Socratic Debates: Examining Persona Effects on Moral Decision and Persuasion Dynamics

Jiarui Liu, **Yueqi Song***, Yunze Xiao*, Mingqian Zheng*, Lindia Tjuatja, Jana Schaich Borg, Mona Diab, Maarten Sap

EMNLP 2025; Link: <https://arxiv.org/abs/2506.12657>

[11] FieldWorkArena: Agentic AI Benchmark for Real Field Work Tasks

Atsunori Moteki, Shoichi Masui, Fan Yang, **Yueqi Song**, Yonatan Bisk, Graham Neubig, Ikuo Kusajima, Yasuto Watanabe, Hiroyuki Ishida, Jun Takahashi, Shan Jiang

Under Review; Link: <https://arxiv.org/abs/2505.19662>

[10] VisualPuzzles: Decoupling Multimodal Reasoning Evaluation from Domain Knowledge

Yueqi Song*, Tianyue Ou*, Yibo Kong†, Zecheng Li†, Graham Neubig, Xiang Yue

Under Review; Link: <https://arxiv.org/abs/2504.10342>

[9] **SkillWeaver: Web Agents can Self-Improve by Discovering and Honing Skills**

Boyuan Zheng*, Michael Y Fatemi*, Xiaolong Jin*, Zora Zhiruo Wang, Apurva Gandhi, **Yueqi Song**, Yu Gu, Jayanth Srinivasa, Gaowen Liu, Graham Neubig, Yu Su

Under Review; Link: <https://arxiv.org/abs/2504.07079>

[8] **Beyond browsing: Api-based web agents**

Yueqi Song, Frank Xu, Shuyan Zhou, Graham Neubig

ACL 2025 (Findings); Link: <https://arxiv.org/abs/2410.16464>

[7] **Crowdsource, Crawl, or Generate? Creating SEA-VL, a Multicultural Vision-Language Dataset for Southeast Asia**

Samuel Cahyawijaya, Holy Lovenia, Joel Ruben Antony Moniz, Tack Hwa Wong, Mohammad Rifqi Farhansyah, Thant Thiri Maung, Frederikus Hudi, David Anugraha, Muhammad Ravi Shulthan Habibi, Muhammad Reza Qorib, Amit Agarwal, Joseph Marvin Imperial, Hitesh Laxmichand Patel, Vicky Feliren, Bahrul Ilmi Nasution, Manuel Antonio Rufino, Genta Indra Winata, Rian Adam Rajagede, Carlos Rafael Catalan, Mohamed Fazli Imam, Priyaranjan Pattnayak, Salsabila Zahirah Pranida, Kevin Pratama, Yeshil Bangera, Adisai Na-Thalang, Patricia Nicole Monderin, **Yueqi Song**, Christian Simon, Lynnette Hui Xian Ng, et al.

ACL 2025; Link: <https://arxiv.org/abs/2503.07920>

[6] **Pangea: A fully open multilingual multimodal llm for 39 languages**

Xiang Yue*, **Yueqi Song***, Akari Asai, Seungone Kim, Jean de Dieu Nyandwi, Simran Khanuja, Anjali Kantharuban, Lintang Sutawika, Sathyanarayanan Ramamoorthy, Graham Neubig

ICLR 2025; Link: <https://arxiv.org/abs/2410.16153>

[5] **OpenHands: An Open Platform for AI Software Developers as Generalist Agents**

Xingyao Wang, Boxuan Li, Yufan Song, Frank F Xu, Xiangru Tang, Mingchen Zhuge, Jiayi Pan, **Yueqi Song**, Bowen Li, Jaskirat Singh, Hoang H Tran, Fuqiang Li, Ren Ma, Mingzhang Zheng, Bill Qian, Yanjun Shao, Niklas Muennighoff, Yizhe Zhang, Binyuan Hui, Junyang Lin, Robert Brennan, Hao Peng, Heng Ji, Graham Neubig

ICLR 2025; Link: <https://arxiv.org/abs/2407.16741>

[4] **What Is Missing in Multilingual Visual Reasoning and How to Fix It**

Yueqi Song, Simran Khanuja, Graham Neubig

NAACL 2025 (Findings); Link: <https://arxiv.org/abs/2403.01404>

[3] **An image speaks a thousand words, but can everyone listen? On image transcreation for cultural relevance (*Best Paper Award*)**

Simran Khanuja, Sathyanarayanan Ramamoorthy, **Yueqi Song**, Graham Neubig

EMNLP 2024; Link: <https://aclanthology.org/2024.emnlp-main.573/>

[2] **GlobalBench: A Benchmark for Global Progress in Natural Language Processing**

Yueqi Song, Catherine Cui, Simran Khanuja, Pengfei Liu, Fahim Faisal, Alissa Ostapenko, Genta Indra Winata, Alham Fikri Aji, Samuel Cahyawijaya, Yulia Tsvetkov, Antonios Anastasopoulos, Graham Neubig

EMNLP 2023; Link: <https://aclanthology.org/2023.emnlp-main.875/>

[1] **The Slowdown of the Chinese Economy**

Yueqi Song*, Yiru Wang*, Zhushan Xiang*

AJIBM, 2019; Link: <https://www.scirp.org/journal/paperinformation?paperid=94337>

TEACHING

Lead TA for CMU 15-213/14-513/15-513 (Computer Systems)

Pittsburgh, PA

Jan 2022 - May 2022

- Taught recitations that review coursework and prepare students for labs, assignments, and exams.
- Held office hours and answered questions from more than 200 students.
- Designed, wrote, and proctored exams, labs, and assignments.
- Received ratings of 5/5 from all students.

COMMUNITY SERVICES

Paper Review

ACL, ACL Student Research Workshop, ACL AfricaNLP Workshop

EMNLP

ICLR

NAACL Student Research Workshop

Neurips TS4H Workshop

WORK EXPERIENCE

Backend Software Engineering Intern at TurboTax, [Intuit](#)

San Diego, CA

May 2023 - Aug 2023

- Developed TurboTax Taxbot on top of GPT-3.5 and GPT-4 through prompt engineering, enabling users to conveniently file their taxes in a conversational manner instead of with traditional tax forms.
- Implemented an advanced question answering mechanism on Turbotax, integrating GPT-3.5 and GPT-4, empowering users to ask questions related to tax filings when using traditional tax forms.
- Automated big tests on tax data, enabling tax analysts to easily locate their interest.

Software Development Intern at [PreVeil, Inc.](#) (Frontend)

Boston, MA

May 2022 - Aug 2022

- Used Golang to add file status and modification time features to PreVeil's secure file sharing solution, reducing time complexity of PreVeil's file status checking from $O(n)$ to $O(1)$.
- Implemented prefix tree structure in syncing, reducing time complexity from $O(n)$ to $O(\log(n))$.
- Performed testing of software features and fixed functional bugs associated with file lock status.

HONORS

- Dean's List (High Honors): 2020, 2021, 2022, 2023
- CMU Undergraduate University Honors
- CMU School of Computer Science College Honors
- Best Paper Award: EMNLP 2024