

Siyan (Sylvia) Li

PHD STUDENT · COLUMBIA UNIVERSITY

✉ siyan.li@columbia.edu | 🏠 siyan-sylvia-li.com | 📷 [siyan-sylvia-li](https://www.instagram.com/siyan-sylvia-li) | 🐦 [Sylvia_Sparkle](https://twitter.com/Sylvia_Sparkle)

Education

Columbia University

New York, NY

PHD IN COMPUTER SCIENCE

2023 - 2028

- GPA: 3.91 / 4.0
- Advisors: Prof. Zhou Yu, Prof. Julia Hirschberg
- **Research Areas: Interactive Dialogue Systems in Education and Privacy**

Stanford University

Stanford, CA

MASTER'S IN COMPUTER SCIENCE

2021 - 2023

- GPA: 4.013 / 4.0
- Advisors: Prof. Chris Manning, Prof. Chris Potts
- **Research Areas: Natural Language Processing, Speech, Dialogue Systems, Linguistics**

Georgia Institute of Technology

Atlanta, GA

BS IN COMPUTER SCIENCE

2017 - 2020

- GPA: 3.94 / 4.0
- Threads: People and Intelligence, Dean's List 2017-2020
- **Research Areas: Natural Language Processing, Story Generation, Natural Language Generation, Alignment, Computer Vision, Robotics**

Professional Experience

- Jun. 2023 - **Research Engineering Intern**, RAIA HEALTH (Early stage startup)
Aug. 2023
- Jun. 2022 - **Research Intern**, Affective Computing Group, UNIVERSITY OF SOUTHERN CALIFORNIA
Sept. 2022
- Jan. 2021 - **Remote Research Assistant**, Robotics, Evolution, and Art Lab, IT UNIVERSITY OF COPENHAGEN
Aug. 2021
- May 2020 - **Software Engineering Intern**, AMAZON ROBOTICS
Aug. 2020
- May 2019 - **Research Intern**, Intelligent Control Lab, CARNEGIE MELLON UNIVERSITY
Aug. 2019

Publications

- [1] **Li Siyan**, Vethavikashini Chithra Raghuram, Omar Khattab, Julia Hirschberg, and Zhou Yu. PAPIL-LON: PrivAcy Preservation from Internet-based and Local Language MOdel ENsembles. In *arXiv preprint arXiv:2410.17127*, 2024. URL: <https://arxiv.org/pdf/2410.17127>.
- [2] **Li Siyan**, Teresa Shao, Zhou Yu, and Julia Hirschberg. EDEN: Empathetic Dialogues for English learning. In *Findings of EMNLP 2024*, 2024. URL: <https://arxiv.org/abs/2406.17982>.
- [3] **Li Siyan**, Teresa Shao, Zhou Yu, and Julia Hirschberg. Using Adaptive Empathetic Responses for Teaching English. In *Proceedings of the 19th Workshop on Innovative Use of NLP for Building Educational Applications*, Mexico City, June 2024. Arxiv. URL: <https://arxiv.org/pdf/2404.13764>.
- [4] **Siyan Li**. Color me intrigued: quantifying usage of colors in fiction. In *Creative AI Across Modalities, AAAI 2023*, Washington, DC, February 2023. Arxiv. URL: <https://arxiv.org/pdf/2301.03559>.

- [5] **Siyan Li**, Riley Carlson, and Christopher Potts. Systematicity in GPT-3's Interpretation of Novel English Noun Compounds. In *Findings of EMNLP 2022*, 2022. URL: <https://arxiv.org/abs/2210.09492>.
- [6] Eric Mitchell, Joseph J. Noh, **Siyan Li**, William S. Armstrong, Ananth Agarwal, Patrick Liu, Chelsea Finn, and Christopher D. Manning. Enhancing Self-Consistency and Performance of Pretrained Language Models with NLI. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, 2022. URL: <https://ericmitchell.ai/concord.pdf>.
- [7] Xiangyu Peng, **Siyan Li**, Sarah Wiegrefe, and Mark Riedl. Inferring the Reader: Guiding Automated Story Generation with Commonsense Reasoning. In *Findings of EMNLP 2022*, 2022. URL: <https://arxiv.org/abs/2105.01311>.
- [8] **Siyan Li**, Ashwin Paranjape, and Christopher Manning. When can I Speak? Predicting initiation points for spoken dialogue agents. In *Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 217–224, Edinburgh, UK, September 2022. Association for Computational Linguistics. URL: <https://aclanthology.org/2022.sigdial-1.22>.
- [9] Amal Alabdulkarim, **Siyan Li**, and Xiangyu Peng. Automatic story generation: Challenges and attempts. In *Proceedings of the Third Workshop on Narrative Understanding*, pages 72–83, Virtual, June 2021. Association for Computational Linguistics. URL: <https://aclanthology.org/2021.nuse-1.8>, doi:10.18653/v1/2021.nuse-1.8.
- [10] Xiangyu Peng, **Siyan Li**, Sarah Wiegrefe, and Mark Riedl. Inferring the reader: Guiding automated story generation with commonsense reasoning. In *Proceedings of the Third Workshop on Narrative Understanding*, Virtual, June 2021. Arxiv. URL: <https://arxiv.org/pdf/2105.01311.pdf>.
- [11] Shyam Sudhakaran, Djordje Grbic, **Siyan Li**, Adam Katona, Elias Najarro, Claire Glanois, and Sebastian Risi. Growing 3D Artefacts and Functional Machines with Neural Cellular Automata. In *Proceedings of the Artificial Life Conference 2021*, volume abs/2103.08737, 2021. URL: <https://arxiv.org/abs/2103.08737>, arXiv: 2103.08737.
- [12] Xiangyu Peng, **Siyan Li**, Spencer Frazier, and Mark Riedl. Reducing non-normative text generation from language models. In *Proceedings of the 13th International Conference on Natural Language Generation*, pages 374–383, Dublin, Ireland, December 2020. Association for Computational Linguistics. URL: <https://aclanthology.org/2020.inlg-1.43>.
- [13] Abulikemu Abuduweili, **Siyan Li**, and Changliu Liu. Adaptable Human Intention and Trajectory Prediction for Human-Robot Collaboration. In *Proceedings of AI-HRI 2019*, volume abs/1909.05089, 2019. URL: <https://arxiv.org/pdf/1909.05089.pdf>.

Awards, Fellowships, & Grants

- 2025 **INSPIRE Award**, The International Network for Simulation-based Pediatric Innovation, Research, & Education (INSPIRE)
- 2023 **Mudd Fellow**, School of Engineering and Applied Sciences, Columbia University
- 2023 **1st Place, Scientific Innovation Prize**, Amazon Alexa Prize Social Bot Challenge

Invited Talks

- March 2025 **Invited Speaker**, 2025 Columbia University U.S. China Education Forum

Teaching Experience

- Sept. 2024 - Dec. 2024 **COMS 4705W Natural Language Processing**, Head Teaching Assistant
- Mar. 2023 - Jun. 2023 **CS 224U Natural Language Understanding**, Teaching Assistant
- Jan. 2023 - Mar. 2023 **CS 224N Natural Language Processing with Deep Learning**, Teaching Assistant
- Sept. 2021 - Dec. 2021 **CS 103 Mathematical Foundations of Computing**, Teaching Assistant
- Jan. 2017 - Aug. 2020 **CS 3600 Introduction to AI**, Teaching Assistant, Head Teaching Assistant

Mentoring

- 2025 **Yilin Ye**, Alumna, Columbia University
- 2025 **Omar Yahia**, BS Student, UC Irvine
- 2024 - 2025 **Riya Raj**, MS Student, Columbia University
- 2024 - 2025 **Zackary Rackauckas**, MS Student, Columbia University
- 2024 - 2025 **Vethavikashini Chithra Raghuram**, MS Student, Columbia University
- 2024 **Brittney Lilly**, BS Student, Columbia University
- 2024 **Teresa Shao**, BS Student, Columbia University
- 2022 **Riley D. Carlson**, BS Student, Stanford University

Research Experience

Columbia NLP, Columbia University

New York, NY

ADVISORS: PROF. ZHOU YU, PROF. JULIA HIRSCHBERG

Aug. 2023 - Present

- Building empathetic spoken chatbot systems for teaching English
- Improving English-teaching chatbot's user retention capability

Stanford NLP Group, Stanford University

Stanford, CA

ADVISORS: PROF. CHRIS MANNING, PROF. CHRIS POTTS

Aug. 2021 - Jun. 2023

- Training deep neural networks to predict dialogue agent utterance initiation points in a continuous manner
- Participating in the Amazon Alexa Prize competition
- Probing linguistic knowledge of GPT-3 in the field of noun compounds

Affective Computing Group, University of Southern California

Los Angeles, CA

ADVISOR: PROF. JONATHAN GRATCH

Jun. 2022 - Sept. 2022

- Examining conversational behaviors in human face-to-face multi-issue negotiations
- Conducting statistical analyses between facial expressions during negotiations and various subjective and objective metrics
- Using visual, audio, and lexical features to predict appropriate backchannel opportunities

Robotics, Evolution, and Art Lab, IT University of Copenhagen

Remote

ADVISOR: PROF. SEBASTIAN RISI

Jan. 2021 - Aug. 2021

- Designing web interfaces for interactive evolution of Minecraft structures through Neural Cellular Automata

Entertainment Intelligence Lab, Georgia Tech

Atlanta, GA

ADVISOR: PROF. MARK RIEDL

Aug. 2019 - Dec. 2020

- Utilizing tools such as COMET and VerbNet to enhance the quality of automatic story generation
- Training language models to produce less toxic content
- Gaining experience in conducting online human subject studies

Skills

TECHNICAL

- Natural Language Processing, Speech Processing, Machine Learning, Agile Methodologies

PROGRAMMING LANGUAGES

- Python, R, Java, C/C++

FRAMEWORKS AND TOOLS

- Pytorch, scikit-learn, Tensorflow, nltk, Flask, ROS

LANGUAGES

- English, Mandarin, Japanese, Spanish, French

Professional Activities

ORGANIZER

- Sept.2024 - present **HCI + NLP Workshop Co-Organizer**, Sept.2024 - present
- 2023 - present **Columbia NLP Seminar Co-Organizer**, Fall 2023 - present
- 2022 - 2023 **Stanford NLP Seminar Co-Organizer**, Fall 2022 - Spring 2023
- 2021 - 2022 **Stanford NLP Lunch Organizer**, Fall 2021 - Spring 2022

SERVICE AND OUTREACH

- Nov. 2024 - **Computer Science PhD Representative**, Columbia University
- Oct. 2026
- Feb. 2024 **Reviewer**, ACL 2024
- Dec. 2022 - **Reviewer, Program Committee**, ACL 2023
- Apr. 2023
 - Reviewer for the Dialogue and Interactive Systems Track.
- Dec. 2022 **Reviewer**, AAAI 2023
- Reviewer for the Creative AI Across Modalities Workshop.
- Jun. 2022 - **Reviewer**, EMNLP 2022
- Oct. 2022
 - Reviewer for the Ethics track and the Dialogue and Interactive Systems track.
- Dec. 2020 - **Volunteer**, Women in ML (WiML)
- Dec. 2022
 - Social media outreach and website design.

Media Coverage

Is It My Turn Yet? Teaching a Voice Assistant When to Speak. [Stanford HAI](#).