Sayan Ghosh (ghoshsay@usc.edu)

EDUCATION	**University of Southern California**, Los Angeles, CA	2022-

- Ph.D Candidate, advised by Xiang Ren

University of Michigan, Ann Arbor, MI — 2017-2021

- B.S.E. in Computer Science Engineering (*summa cum laude*)
- B.S. in Honors Mathematics (*with distinction*)

RESEARCH AND INDUSTRY EXPERIENCE

- **ASAPP** — 2021-2022
  Research Engineer: NLP Team, focusing on dialogue summarization and natural language understanding

- **ASAPP** — 2021
  Research Intern: NLP Team

- **University of Michigan Department of Computer Science** — 2019-2021
  Research Assistant, with Joyce Chai

- **University of Michigan School of Information** — 2019-2021
  Research Assistant, with David Jurgens

TEACHING EXPERIENCE AND SERVICE

- **University of Southern California Department of Computer Science** — 2024-
  Teaching Assistant for CSCI 499 (Language Models), CSCI 544 (Applied NLP), and CSCI 567 (Machine Learning)

- **University of Michigan Department of Computer Science** — 2020-2021
  Instructional Assistant for EECS 376 (an introductory course on Theoretical Computer Science)

- **University of Michigan Department of Mathematics** — 2018-2020
  Math Lab Tutor

PREPRINTS AND UNDER REVIEW

- Sayan Ghosh, Shahzaib Saqib Warraich, Dhruv Tarsadiya, Gregory Yauney, and Swabha Swayamdipta. Sample, Align, Synthesize: Graph-Based Response Synthesis with Con-Grs, 2025a. URL https://arxiv.org/abs/2510.03527

WORKSHOP PUBLICATIONS

- Sayan Ghosh, Shahzaib Saqib Warraich, Dhruv Tarsadiya, Gregory Yauney, and Swabha Swayamdipta. Sample, Align, Synthesize: Graph-Based Response Synthesis with Con-Grs. In *The 1st Workshop on Test-time Scaling and Reasoning Models*, 2025b. URL https://openreview.net/forum?id=OmpqFNzUwa

- Sayan Ghosh, Dylan Baker, David Jurgens, and Vinodkumar Prabhakaran. Detecting Cross-Geographic Biases in Toxicity Modeling on Social Media. In *Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021)*, pages 313–328, Online, November 2021. Association for Computational Linguistics. URL https://aclanthology.org/2021.wnut-1.35 (**Best paper**)

PEER-REVIEWED CONFERENCE PUBLICATIONS

- Abrar Anwar, Rohan Gupta, Zain Merchant, Sayan Ghosh, Willie Neiswanger, and Jesse Thomason. Efficient Evaluation of Multi-Task Robot Policies With Active Experiment Selection. In *9th Annual Conference on Robot Learning*, 2025. URL https://openreview.net/forum?id=pJ5FONkM9N

- Sayan Ghosh, Tejas Srinivasan, and Swabha Swayamdipta. Compare without Despair: Reliable Preference Evaluation with Generation Separability. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 12787–12805, Miami, Florida, USA, November 2024. Association for Computational Linguistics. URL https://aclanthology.org/2024.findings-emnlp.747

- Julia Mendelsohn, Sayan Ghosh, David Jurgens, and Ceren Budak. Bridging Nations: Quantifying the Role of Multilinguals in Communication on Social Media. *Proceedings of the International AAAI Conference on Web and Social Media*, 17(1):626–637, Jun. 2023. doi:10.1609/icwsm.v17i1.22174. URL https://ojs.aaai.org/index.php/ICWSM/article/view/22174 (**Best Methodology**)

- Yuwei Bao, Sayan Ghosh, and Joyce Chai. Learning to Mediate Disparities Towards Pragmatic Communication. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2829–2842, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi:10.18653/v1/2022.acl-long.202. URL https://aclanthology.org/2022.acl-long.202

- Innocent Ndubuisi-Obi*, Sayan Ghosh*, and David Jurgens. Wetin dey with these comments? Modeling Sociolinguistic Factors Affecting Code-switching Behavior in Nigerian Online Discussions. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 6204–6214, Florence, Italy, July 2019. Association for Computational Linguistics. doi:10.18653/v1/P19-1625. URL https://aclanthology.org/P19-1625

HONORS AND AWARDS

- Best Methodology Award - ICWSM 2023   2023

- Best Paper Award - Workshop on Noisy User-Generated Text (W-NUT) at EMNLP 2021   2021

- CRA Outstanding Undergraduate Researcher Award - Honorable Mention   2020

- University of Michigan Dean's List: Every semester of study

TALKS

- Google Data Ethics Team   Feb. 2021

MENTORING

- (Masters) Shahzaib Saqib Warrach   2024 -

- (Masters) Dhruv Tarsadiya   2025 -

- (Undergraduate) Naina Panjwani   2025 -

- (Undergraduate) Marios Tsotras   2025 -

- (Undergraduate) Aryan Gulati   2024 -

- (Undergraduate) Catherine He   2024 -

- (Undergraduate) Risha Surana (USC CURVE Program)   2023-2024

REVIEWING

- ARR, ACL, EMNLP, NAACL, ICLR

LANGUAGES SPOKEN

English (native), Bengali (native), French, Hindi