8601 34th Avenue
College Park, MD-20740
+1(631)652-5197
rupak@umd.edu
styx97.github.io

Rupak Sarkar

Education

2021-present **Ph.D. Student**, *Department of Computer Science, University of Maryland, College Park*, Maryland, USA.
Advisor: Prof. Philip Resnik

2016-2020 **B. Tech**, *Kalyani Government Engineering College*, Kalyani, India.
Discipline: Computer Science and Engineering, CGPA: 8.98/10.00

Experience

Research Scientist Intern at Microsoft Research + OfficeAI, *Redmond*, Washington, USA.

May 2025 - Aug 2025 Worked on semantic understanding of noisy user interactions on Microsoft Office applications to enable breakthrough insights and features. *Manuscript under preparation*

Research Scientist Intern at Microsoft Research + Bing Search, *Redmond*, Washington, USA.

June 2024 - Aug 2024 Worked on understanding the information needs of users in human-AI conversations, and how we can make conversational agents better serve a diverse range of human queries. *Manuscript under Review*

Research Scientist Intern at Adobe, *College Park/Hybrid*, Maryland, USA.

June 2022 - Aug 2022 Designed and Implemented a framework for identifying related concepts in an educational document and fetching intuitive explanations of concepts from the web. Worked with Prof. Ani Nenkova and Dr. Varun Manjunatha.

Research Intern at Carnegie Mellon University, *Pittsburgh*, Pennsylvania, USA (remote).

Sep 2020 - Dec 2020 Course Research Engineer for *11-865/11-665: Tracking Political Sentiments Using Machine Learning* with instructors Dr. Ashiqur R. Khudabukhsh, Prof. Mark Kamlet and Prof. Tom Mitchell.

Publications

2025 *"Understanding Common Ground Misalignment in Goal-Oriented Dialog: A Case-Study with Ubuntu Chat Logs"*, by Rupak Sarkar, Neha Srikanth, Taylor Hudson, Rachel Rudinger, Claire M. Bonial and Philip Resnik. [***ACL 2025 Main (Long Paper)***]

2025 *"Conversational User-AI Intervention: A Study on Prompt Rewriting for Improved LLM Response Generation"*, by Rupak Sarkar, Bahareh Sarrafzadeh, Nirupama Chandrasekaran, Nagu Rangan, Philip Resnik, Longqi Yang, Sujay Kumar Jauhar. [*Under Review at ARR*]

2025 *"Measuring Scalar Constructs in Social Science with LLMs"*, by Hauke Licht*, Rupak Sarkar*, Patrick Y. Wu, Pranav Goel, Niklas Stoehr, Elliott Ash, Alexander Miserlis Hoyle. [***EMNLP 2025 Main (Long Paper)***]

2024 *"PairScale: Analyzing Attitude Change in Online Communities"*, by Rupak Sarkar, Patrick Y Wu, Kristina Miler, Alexander Hoyle* and Philip Resnik*. [***NAACL 2025 Findings (long paper)***]

2024 *"Pregnant Questions: The Importance of Pragmatic Awareness in Maternal Health Question Answering"*, by Neha Srikanth*, Rupak Sarkar*, Rachel Rudinger and Jordan Boyd-Graber. [***NAACL 2024 Main (long paper)***]

2023 *"Making the Implicit Explicit: Implicit Content as a First Class Citizen in NLP"*, by Alexander M. Hoyle*, Rupak Sarkar*, Pranav Goel and Philip Resnik. [***EMNLP 2023 Main (long paper)***]

2022 *"Are Neural Topic Models Broken?"*, by Alexander M. Holye, Pranav Goel, Rupak Sarkar, and Philip Resnik. [***Findings of EMNLP 2022 (long paper)***]

2022 *"Fringe News Networks: Dynamics of US News Viewership following the 2020 Presidential Election"*, by Ashiqur R. KhudaBukhsh*, Rupak Sarkar*, Mark S. Kamlet, and Tom M. Mitchell. [***14th ACM Web Science Conference 2022, 269-278***]

2021 *"Are Chess Discussions Racist? An Adversarial Hate Speech Data Set"*, by Rupak Sarkar and Ashiqur R. KhudaBukhsh. [***AAAI 2021, Best AAAI-21 Student Abstract 3-Minute Presentation*** *(Student Abstract)*]

2021 *"We Don't Speak the Same Language: Interpreting Polarization Through Machine Translation"*, by Ashiqur R. KhudaBukhsh*, Rupak Sarkar*, Mark S. Kamlet, and Tom M. Mitchell. [***AAAI 2021****, (long paper) acceptance rate : 21%*]

2020 *"Social Media Attributions in the Context of Water Crisis"*, by Rupak Sarkar*, Sayantan Mahinder*, Hirak Sarkar, and Ashiqur R. KhudaBukhsh. [***EMNLP 2020 Main****, (long paper) acceptance rate : 22.4%*]

Selected Media Coverage

NYTimes *"For Trump's Backers in Congress,'Devil Terms' Help Rally Voters"*, *The New York Times, October, 2022*

CMU Press *"The Science of Political Polarisation"*, *CMU Science Magazine (cover story), August, 2021*

WIRED *"Why a YouTube Chat About Chess Got Flagged for Hate Speech"*, *WIRED, March, 2021*

CMU Press *"AI May Mistake Chess Discussions as Racist Talk"*, *CMU Press Release, February, 2021*

Independent *"Newsmax and OAN: How are the ultra-conservative cable channels coping without Trump in the White House?"*, *Independent, February, 2021*

WIRED *"The Left and the Right Speak Different Languages — Literally"*, *WIRED, October, 2020*

CMU Press *"Even our language is polarized"*, *CMU Press Release, October, 2020*

Futurity *"Mask vs. Muzzle: Even Words Are Now Polarized."*, *Futurity, October, 2020*

Skills and Interests

Tools Python (PyTorch), Huggingface, R, SQL, C++

Research Interests Conversational AI, NLP, Understanding Human Intent from Text

Awards

2021 Was awarded the **Dean's Fellowship - PhD, CS** by University of Maryland, College Park.

2021 Won the **Best AAAI-21 Student Abstract 3-Minute Presentation** award for *"Are Chess Discussions Racist? An Adversarial Hate Speech Data Set"* at AAAI-2021 out of 400+ submissions.

2019 Won the **Smart India Hackathon 2019 Software Edition** with a prize money of *75,000 INR* as part of the team representing our college. Our problem statement (AM1) was focused on building an Augmented Reality App for the popular e-furniture company, PepperFry.