

Profile

Ph.D. student in Computer and Information Science specializing in computational social science, with a focus on understanding patterns of news production in America and how individuals consume it, spanning television, online, and social media platforms. Expertise in machine learning, natural language processing, network science, and causal inference. Experienced in interdisciplinary research analyzing large-scale longitudinal data using large language models, statistical modeling, and social network analysis.

Education

PhD in Computer and Information Science, University of Pennsylvania GPA: 3.94/4.0 Advisors: Duncan Watts, Aaron Clauset	2022-2027
Master of Science in Computer Science, University of Colorado Boulder GPA: 4.0/4.0	2019-2022
B.Tech in Computer Science and Engg., Heritage Institute of Technology, India GPA: 9.36/10	2015-2019

Internships

Applied Scientist Intern, Microsoft, Redmond WA Team: Copilot Notebooks Developed an LLM-as-a-judge evaluation framework for Copilot Notebooks, validated against inter-rater reliability metrics, surfacing strengths and failure modes and benchmarking against Google's NotebookLM.	May-August 2026
Machine Learning Intern, Meta (Facebook), Menlo Park CA Team: Profile Relevance Developed end-to-end models for ranking Facebook friend-lists across four view types, beating production AUC by 18.3% with A/B engagement gains; launched model to production for 3 billion Facebook users.	May-August 2025

Publications and Ongoing Research

Asymmetric Framing Strategies Drive Partisanship in Television News [PDF] (under review) Upasana Dutta, Homa Hosseinmardi, Amir Ghasemian, Aaron Clauset, Duncan Watts Leveraged student-teacher model setup for training neural networks on LLM labels for detecting partisanship and framing strategies, showing media bias in 400,000+ TV news episodes (2013-2025).	2026
Framing the Fringe: Dynamics of Ingroup and Outgroup Narratives in Fringe Telegram Channels Upasana Dutta, Yasmine Hour Published in the Proceedings of the 18th ACM Web Science Conference 2026 [PDF] [Best Paper Award] Conducted multi-channel analysis of 300,000 Telegram messages to quantify the prevalence of ingroup and outgroup narratives and statistically comparing its propagation in direct vs. forwarded messages.	2026
Exploring Political Attitudes under Selective Media Consumption with Large Language Models [Slides] Upasana Dutta, Josh Nguyen, Samar Haider, Bryan Li, Neil Fasching, Duncan Watts Simulated American news consumption with LLMs grounded in real web-browsing histories, and ran a randomized controlled trial estimating how partisan media environments shape political attitudes.	2026
Algorithmic systems, human agency and the future of platform research [PDF] H. Hosseinmardi, Upasana Dutta, David Rothschild, Duncan Watts To appear in Nature Computational Science 2026 Reviewed the evidence on algorithmic recommendation and user self-selection in online content diets, finding the case against algorithmic curation inconclusive and calling for large-scale empirical research.	2026
LLM-Assisted Reranking to Operationalize Nuanced Objectives in Recommender Systems [PDF] (under review) A. Ghasemian, H. Hosseinmardi, U. Dutta, J. Allen, D. Rothschild, D. Watts Reranked YouTube recommendations via zero-shot LLM prompting on real browsing histories; prompt-level regularization curbed extreme content and broadened ideological exposure.	2026
News-like Information Ecosystem on YouTube [Abstract] A. Ghasemian, H. Hosseinmardi, U. Dutta, J. Allen, D. Rothschild, D. Watts Analyzed 385,903 YouTube videos using LLM classifications and unsupervised clustering, showing higher user engagement with non-traditional content creators than traditional media outlets for news.	2025

Political Slant and Causal Overreach in the Behavioral Sciences (under review) Calvin Isch, Upasana Dutta, Cory Clark, Phil Tetlock, Eric Uhlmann Analyzed 194,606 behavioral science articles with LLM classifications, finding that politically slanted articles roughly doubled from 1970 to 2020 and make more overreaching causal claims than apolitical ones.	2025
Scaling laws in empirical networks [Preprint] (under review) Upasana Dutta, Alexander Ray, Aaron Clauset Analyzed the scaling trends of 254 real-world networks and assessed the extent to which null models, such as the configuration model and the stochastic block model, explain their scaling behavior.	2024
Sampling random graphs with specified degree sequences [PDF] [Code] [Python Package] Upasana Dutta, Bailey K. Fosdick, Aaron Clauset Published in the Journal of Computational and Graphical Statistics 2025 Developed a method for sampling networks from the degree-preserving configuration model by detecting convergence in a double-edge swap Markov chain sampler.	2023
Analyzing Twitter Users' Behaviour Before and After Contact by Russia's Internet Research Agency Upasana Dutta, Rhett Hanscom, Jason Zhang, Richard Han, Tamara Lehman, Qin Lv, Shivakant Mishra Published in the Proceedings of the ACM on Human-Computer Interaction, CSCW 2021 [PDF] [Code] Employed statistical methods to analyze changes in user behavior on Twitter after they engaged with Russian bots backed by the Internet Research Agency (IRA) during the US 2016 presidential election.	2021

Awards

GAPSA Research Grant , University of Pennsylvania	2024
Abel Lukens Stout Fellowship , Department of Computer and Information Science, UPenn	2022
Bell Foundation Outstanding Research Award , Department of Computer Science, CU Boulder	2022
CS Annual Research Expo Award , Department of Computer Science, CU Boulder [Poster]	2022
NCWIT Collegiate Award 2022 Finalist [Video]	2021
CS Annual Research Expo Award , Department of Computer Science, CU Boulder [Poster]	2021
CS Publication Recognition Award , Department of Computer Science, CU Boulder [Publication]	2021

Talks

Penn Center on Media, Technology, and Democracy 2026 [Slides]	2026
PhD Symposium, The Web Conference 2026 [Slides]	2026
International Conference on Computational Social Science (IC2S2 2024) [Slides]	2024
ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2024) [Slides]	2024
Bernoulli-IMS 11th World Congress in Probability and Statistics 2024 [Slides]	2024
The Mitchell Centre for Social Network Analysis, University of Manchester [Slides] [Talk]	2021
A Joint Sunbelt and NetSci Conference (Networks 2021) [Abstract] [Slides] [Talk]	2021
International Conference on Complex Networks (CompleNet 2021) [Abstract] [Slides] [Talk]	2021

Professional Activities

Reviewer for ICWSM 2026 and 2027, and The Web Conference 2026	2026-2027
Reviewer for IC2S2 2024 and 2025, and NAACL Student Research Workshop 2025	2024-2025
Teaching Assistant for CIS 5300 Natural Language Processing at UPenn	2024
Gave a Networks Tutorial at the Summer Institute of Computational Social Science 2024 [Code]	2024
Google CS Research Mentorship Program 2022	2022
Complex Networks Winter Workshop 2021, University of Vermont [Slides]	2021

Skills

Data Science: Hypothesis testing, Statistical Inference, Anomaly Detection, Time-Series Analysis
Machine Learning: Regression Analysis, Supervised/Unsupervised Learning, Ensemble Methods, Active Learning, Deep Learning, Model Training and Fine-Tuning, Model Evaluation
NLP: Text Preprocessing, LLMs, Embeddings, Topic Modeling, Prompt Engineering
Comp. Social Science: Social Computing, Longitudinal Data Analysis, Observational Studies
Network Analysis: Centrality Measures, Null models, Community detection, Cascade models
Causal Inference: Experimental Designs, A/B testing, RCTs, Diff-in-Diff, Instrumental Variables