NEW YORK UNIVERSITY
PhD in Data Science
Aug 2023 | New York, NY
GPA: 3.8/4

MUMBAI UNIVERSITY
BEng in Computer Science
June 2018 | Mumbai, India
Cum. GPA: 9.23/10

COURSEWORK

DOCTORAL
Machine Learning
Causal Inference
Inference and Representation
Math Tools for Data Science
Unicode ML Summer School
(Instructor, funded by Google Research)

UNDERGRADUATE
Artificial Intelligence
Big Data Analysis
Operating Systems
Database Management Systems
Computer Networks
Unicode
(Co-founder, FOSS Training Program)

AWARDS

2023	Google exploreCSR Award
2023	WSDM Smart City Challenge
2022	AI2Amplify Fellowship
2022	MIDAS Future Leader Forum
2020	IRIS-HEP Fellowship
2018	Oxford MLHEP Challenge
2018	**India's Top 33 Engineers, Economic Times**

SKILLS

Python, Pytorch • MySQL • Bash • R
GCP • BigQuery • Unix/Linux • AWS
Docker • git • LAMP • SLURM
HTML/CSS • Javascript • C++ • Java

LINKS

Github:// **swapneelm**
Twitter:// **@swapneel_mehta**
YouTube:// **swapneel_mehta**

BOSTON UNIVERSITY AND MIT | Postdoctoral Associate
Oct 2023 - Present | Cambridge, MA

- **Investigating how to mitigate false ads in digital marketplaces**, building the Platform Governance Lab.
- Talks: **IC2S2, WISE, Stanford TSRC, Columbia MarkTech, Yale AI/ML Conf.**

SAKHI | Co-founder and Chief Scientist
July 2023 - Present | New York, NY

- MIT-incubated tech venture **improving women's health literacy** through AI-informed, personalized **behavioral interventions** on WhatsApp.
- Launched **IRB-backed maternal health pilot** with 140 families in Jalgaon, India, **menstrual health pilot** with 200+ families in Dhaka, Bangladesh.
- Won USD 62,500 in grants from AWS, MIT **IDEAS**, **Delta-V**, UNICEF, **UNDP**.
- Invited to present to **World Bank**, lead capstone teams at MIT **E-Lab** and **SEID**.

SIMPPL | Founder and President
Dec 2021 - Present | New York, NY

- Building research-backed tools to **improve trust on the social internet for 100M+ users in 6 countries** with the UN, New York Public Radio, Sunday Times, Deutsche Welle, Jagran News, 10+ other global partners.
- Leading team of 22 staff, won USD 200,000 in grants from AWS, **Google**, **Wikimedia Foundation**, **Mozilla**, **NYC Media Lab, Goethe Institut**.
- **12+ workshop and conference publications** at ICML, NeurIPS, AAAI, ICWSM, IC2S2, dg.o, Stanford TSRC, C+J, Truth and Trust Online.
- Selected as a **Google Research Innovator, Founder Fellow, Belfer Fellow** .
- Invited to lead capstone teams at MIT **A-Lab** by Sinan Aral, **NYU DS**.

CENTER FOR SOCIAL MEDIA AND POLITICS | Ph.D. Research
Feb 2021 – Present | New York, NY

- Thesis: **Towards Informed Interventions to Limit the Effects of Misleading Information on Social Networks**.
- Identify the causal effect of Reddit's outages on low quality news sharing.
- Identify the causal impact of Twitter's warning labels on Trump's tweets (**Stanford Trust & Safety Research Conference**).
- COVID infection modeling with probabilistic programs at **MIT PROBPROG**.

CERN, CMS EXPERIMENT | Technical Student (ML Research)
March 2018 - July 2019 | Geneva, Switzerland

- Developed graph neural networks for particle track reconstruction at the LHC
- Productionized the **DeepJet** ML framework; awarded by **Google, Oxford**.

INTERNSHIPS

SLACK, CENTRAL DS | Product Data Science Intern
May 2023 – Aug 2023 | New York, NY

- Identified 'social' learning drivers of feature adoption at Slack (NDA).

TWITTER, CIVIC INTEGRITY | Machine Learning Engineer
June 2022 – Aug 2022 | New York, NY

- Improved the precision of their civic and health misinfo. classifier (BigQuery ML) by **20%** scoring 1M+ tweets / week. NLP; recommender systems.

REVIEWER

TALKS

OXFORD, TORR VISION GROUP | Visiting Researcher
Aug 2021 – June 2022 | Oxford, UK

- Generative modeling of adversarial agent behaviors on Reddit recommendations presented in workshops at **ICML** and **DEFCON**, 2022.

ADOBE RESEARCH | Data Science Research Intern
May 2020 – Dec 2020 | New York, NY

- Trending Hashtag Recommendation for Videos; NLP, language models, data mining, weakly-supervised learning, multimodal graph neural networks.
- First-author paper accepted at IEEE ISM; U.S. Patent App P10634-US.

MICROSOFT RESEARCH (IIT-B) | Research Intern
May 2016 – May 2017 | New York, NY

- **Prof. S. Sahasrabudhe**, Automated evaluation of Blender 3D Modeling assignments (ICALT 2018); edX Prize Finalist, 2019.
- **Microsoft Research**: Google Analytics for **Lokacart** D2C farmer app.

SMOKESCREEN (ACQ. ZSCALER) | ML Intern
Nov 2016 – March 2017 | New York, NY

- Reducing the footprint of network scans and building Markov-chain based hostname generation for intelligent decoys.
- Automated deployment of NASSCOM-DSCI awarded software.

FELLOWSHIPS

GRANTS AND AWARDS

ORGANIZING

MEDIA

PUBLICATIONS

[1] A. Shah, P. Banerjee, J. Kuang, A. Nichols, N. Mazar, M. van Alstyne, S. Mehta. 'Market Design Interventions for Safer Agentic AI'. *Working Paper; Oral Talk, Yale AI, ML, and Business Analytics Conference, 2024.*

[2] H. Ranka, D. Shah, R. Sannikov, E. Brichetto, L. Ng, S. Mehta. 'Bridging Nodes and Narrative Flows: A Graph-Theoretic Analysis of Telegram's Disinformation Ecosystem'. *Working Paper; Poster at Stanford T&S Research Conference, 2024.*

[3] D. Shah, S. Shetty, S. Mehta. 'The Double-Edged Sword of User Agency: Empowerment and Risk in Decentralized Social Media Platforms'. *Working Paper; Poster at Stanford T&S Research Conference, 2024.*

[4] S. Mehta, M. van Alstyne. 'Improving Economic Welfare in Reputation-based Marketplaces with Truth Warrants'. *Working Paper. Oral Talk, Workshop on Information Systems and Economics, 2024.*

[5] S. Mehta, Z. Yang, C. Fan, D. Liu, N. Mazar, M. van Alstyne. 'Reducing Misleading Claims in Digital Marketplaces using Truth Warrants'. *Working Paper (under review), 2024.*

[6] R. Konka, P. Kurani, S. Mehta., L. Ng. 'Multimodal Analysis of Instagram Marketing Strategies Across Developed and Developing Countries'. *Working Paper, 2024.*

[7] G. Malpani, S. Mehta. 'The Effect of Influencers on the Independent re-sharing of News in an Alt-Right Ecosystem'. *Working Paper, 2024.*

[8] V. Dalal, A.V. Singh, C. Schroeder de Witt, S. Mehta. 'Using Large Language Model Debates to Predict X's Community Notes Reviews'. *Working Paper, 2024.*

[9] R. Jain, S. Mehta, F. Barez, P. Torr. 'Detecting Emergent Social Media Misinformation in the Era of LLMs'. *Working Paper, 2024.*

[10] S. Mehta, J. Bisbee, Z. Sanderson, R. Bonneau, J. Tucker, and J. Nagler. 'Identifying the Causal Effects of Twitter's Interventions on Trump's Tweets'. *Working Paper; Oral Talk, Stanford Trust and Safety Conference; Invited Talk, Twitter 2022-24.*

[11] S. Mehta, K. Aslett, Z. Sanderson, R. Bonneau, J. Tucker, and J. Nagler. 'Platform Outages Increase the Sharing of High-quality News on Reddit'. *Working Paper, 2023.*

[12] S. Mehta, S. Sarkhel, X. Chen, S. Mitra, V. Swaminathan, R. Rossi, A. Aminian, H. Guo, and K. Garg. 'Open-Domain Trending Hashtag Recommendation for Videos'. *In IEEE International Symposium on Multimedia (ISM), 2021.*

[13] S. Mehta, and N. Kasmanoff. 'Covid-19 Modeling and Control via Policy Interventions'. *In International Conference on Probabilistic Programming (PROBPROG), 2021.*

[14] S. Mehta, C. Raman, N. Ayer, and S. Sahasrabudhe. 'Auto-Grading for 3D Modeling Assignments in MOOCs'. *Oral Talk, International Conference on Advanced Learning Technologies (ICALT), 51–53. IEEE, 2018.*

[15] S. Mehta, P. Kothuri, and D.L. Garcia. 'A Big Data Architecture for Log Data Storage and Analysis'. Best Paper, ML Track. *Integrated Intelligent Computing, Communication and Security, 201–9. Springer, Singapore, 2018.*

[16] S. Mehta, P. Kothuri, and D.L. Garcia. 'Anomaly Detection for Network Connection Logs'. arXiv Preprint arXiv:1812. 01941, 2018.

WORKSHOP PAPERS

[17] S. Mehta, S. Wiedswang, Z. Yang, N. Mazar, M. van Alstyne. 'Improving the Governance of Digital Platforms with Interactive Marketplace Experiments '. *Oral Talk, Intl. Conference on Computational Social Science, 2024.*

[18] H. Ranka, M. Surana, N. Kothari, V. Pariawala, P. Banerjee, A. Surve, S.R. Sankepally, R. Jain, J. Lalwani, S. Mehta. 'Examining the Implications of Deepfakes for Election Integrity'. *Oral Talk, AAAI Workshop on AI for Credible Elections, 2024.*

[19] S. Mehta, A.G. Baydin, R. Bonneau, J. Nagler, and P. Torr. 'Estimating the Impact of Coordinated Inauthentic Behavior on Content Recommendations in Social Networks'. *Oral Talk, Workshop AI for Agent-Based Modelling, ICML 2022.*

[20] A. Surve*, A. Shamraj*, and S. Mehta. 'How Decentralization Affects User Agency on Social Platforms'. *Oral Talk, Data Challenge, AAAI Intl. Conference on Web and Social Media (ICWSM), 2024.*

[21] C. Shah, R. Konka, G. Malpani, S. Mehta, and L. Ng. 'Can Social Media Platforms Transcend Political Labels? An Analysis of Neutral Conservations on Truth Social'. *Oral Talk, DARE Workshop, AAAI Intl. Conference on Web and Social Media (ICWSM), 2024.*

[22] A. Surve, A. Rathod, M. Surana, G. Malpani, A. Shamraj, S.R. Sankepally, R. Jain, and S. Mehta. 'Multiagent Simulators for Social Networks'. *(Oral) In Multi-Agent Security Workshop, NeurIPS 2023, 2023.*

[23] D. Gandhi, R. Jain, J. Gala, J. Lalwani, and S. Mehta. 'Expanding Access to ML Research through Student-Led Collaboratives'. *In Workshop on Broadening Research Collaborations, NeurIPS 2022.*

[24] A. Ganeriwalla, and S. Mehta. 'Application of OLS Regression and VAR Models to Analyse the Economies of Varying Political Regimes'. Journal of Student Research 10, no. 4 (2021).

TECH POLICY RESEARCH AND REPORTS

[25] C. Vergara, J. Lalwani, H. Ranka, N. Kothari, and S. Mehta. 'A Principled Framework to Organize and Develop Social Media Interventions'. *Working Paper, 2024.*

[26] C. Vergara, R. Jain, and S. Mehta. 'A History of Transparency Reform to Shape Platform Regulation'. *Intl. Conference on Digital Governance (dg.O), 2024.*

[27] A. Sharma, A. Shetty, M. Surana, and S. Mehta. 'Harassment in the Digital Arena: Analyzing Online Abuse Against U.S. Politicians'. *Report, ADL Center for Tech and Society, 2024.*

[28] S.W. Baksh, D. Mungra, N. Ahasan, V. Pariawala, R. Jain, A. Surve, P. Banerjee, S.M. Eron, S. Mehta, S.R. Diya. 'How Facebook Has Become A Political Battleground In Bangladesh '. *Report (Tech Global Institute), 2024.*

[29] S.W. Baksh, D. Mungra, V. Pariawala, R. Jain, A. Surve, A. Das, M. Biswas, F. Afroz, S.M. Eron, S. Mahmood, S. Mehta, S.R. Diya. 'From homophobia to assault: The gendered landscape of Bangladesh's political disinformation'. *Report (Tech Global Institute), 2024; Meta issued takedown.*

[30] A. Sharma, A. Shetty, M. Surana, S. Mehta. 'Harassment in the Digital Arena: Analyzing Online Abuse Against U.S. Politicians'. *Report (ADL Center for Tech and Society), 2024.*

[31] H. Ross, B. Silverman, J. Allen, S. Mehta, et al. 'Shining a Light on Platform Transparency Best Practices'. *Report (Integrity Institute), 2023.*

[32] G. Ellingson, K. Harbath, J. Allen, Eric Davis, A. Leavitt, S. Mehta, et al. 'Defining and Achieving Success in Elections Integrity'. *Report (Integrity Institute), 2023.*