Swapneel Mehta

mehtaver.se | swapneel@mit.edu | LinkedIn | U.S. Permanent Resident

---

NEW YORK UNIVERSITY
Ph.D. in Data Science
2019 - 2023
GPA: 3.8/4

MUMBAI UNIVERSITY
BEng in Computer Science
2014 - 2018
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)

TRAVEL AWARDS

2025	TrustCon, ICML (multiple)
2023	WSDM Smart City Challenge
2022	AI2Amplify AI4Good
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**, launching the Platform Governance Lab which I helped grow from 9 to 28 members.
- Developed an e-commerce market like Amazon, and ran interactive behavioral experiments with over 4000 human participants and agentic AI sellers.
- Talks: **IC2S2, WISE, Stanford TSRC, Columbia MarkTech, Yale AI/ML Conf., Google, AMA Marketing and Public Policy, Harvard**

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

- MIT and UNICEF-incubated tech venture **improving women's health literacy** through AI-informed, personalized **behavioral interventions** on WhatsApp.
- Launched **pilots** with 500 families
- Won USD 62,500 in grants from AWS, MIT **IDEAS**, **Delta-V**, UNICEF, **UNDP**.
- Partnering with **Indian Council of Medical Research** for fully funded pilot, presented to **World Bank**.

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, USD 350,000 in grants and revenue from AWS, **Google, Wikimedia Foundation**, **Mozilla**, **NYC Media Lab, Goethe Institut**.
- **16+ 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** .

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

2025	Nature Patterns Journal
2025	IC2S2
2024	NSF
2024	IEEE DSAA MLJ
2024	Springer AI and Ethics
2024	ACM CHI
2024	EPJ Data Science
2024	Springer Nature ML
2023	NeurIPS Workshops
2023	AABI
2023	ICLR, AI4ABM Workshop
2023	AI+Local News Challenge
2022	ICML, AI4ABM Workshop
2022	NeurIPS, WBRC
2019	Elsevier JPDC

TALKS

2025	Harvard (BKC Spring Synth.)
2025	Harvard (BKC Spring Synth.)
2025	IC2S2 (AI and Economics)
2025	TrustCon (Global Majority)
2025	Microsoft NERD (LLMs)
2025	MIT Decentralized AI
2025	Prince Mahidol Awards Conf.
2025	Google (Monetization Policy)
2024	WISE Conf.
2024	Yale AI/ML Conf.
2024	MarkTech Columbia
2024	MIT Decentralized AI Summit
2024	Stanford (TSRC)
2024	TrustCon (AI + Elections)
2024	IC2S2 (Oral)
2024	ATIH, Embassy of Finland
2024	UNESCO, DW Akademie
2024	DW (Online Trust)
2023	DW Innovation Days
2023	IIT Madras Data Science
2023	SimPPL Fellowships
2023	Mandiant mWise (Google)
2023	McKinsey Guest Talk
2023	TrustCon (Community Building)
2023	MIDAS Future Leaders Summit (Responsible AI)
2023	CCNY SASE "TED Talks"
2023	MozFest by Mozilla
2023	NYU Data Science Seminar
2023	Deutsche Welle (multiple)
2023	NYC Media Lab
2022	Graphika
2022	Truth and Trust Online Conf.
2022	Twitter, Health ML
2022	MIT Media Lab
2022	Stanford, T&S Research Conf.
2022	NYU Graduate Seminar
2022	Misinfo. Village, DEFCON

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

2025	ITS Rio Fellow (Brazil)
2025	ISPI Future Leader (Italy)
2024	Community Advisory Board, Integrity Institute
2024	Responsible AI Affiliate, All Tech is Human
2023	Future Founders Fellowship (declined)
2023	Google Research Innovator
2023	Future Founders Startup Bootcamp
2023	CTS Belfer Fellow

AWARDS AND GRANTS

2024	Google (exploreCSR), with Prof. Pranit Bari	USD 75,000
2024	Finalist, Prototypes for Humanity (0.04%)	USD 20,000
2024	Finalist, Responsible Tech Youth Power Fund (1%)	USD 75,000
2024	MIT Delta V Accelerator	USD 20,000
2024	MIT PKG IDEAS Challenge	USD 13,500
2024	FTI: People to Watch in AI and Local News	
2024	Google PaliGemma Award	USD 5,000
2023	Google (exploreCSR), with Prof. Pranit Bari	USD 32,000
2023	Mozilla (Responsible Computing Challenge)	USD 25,000
2023	Algovera AI	USD 3,000
2023	NYUTech Venture Workshop	USD 1,000
2023	Belfer Fellowship	USD 40,000
2023	Goethe Institut	EUR 14,000
2023	Wikimedia Foundation 'WikiCred' Grant	USD 10,000
2023	Deutsche Welle	EUR 4,950
2022	Google Cloud Research Credit Award	USD 9,000
2022	Amazon (AWS) Research Credit Award	USD 5,000
2022	JournalismAI, The Sunday Times	USD 4,000
2021	UK Research & Innovation Grant (w/ Oxford TVG)	GBP 23,000
2021	NYC Media Lab AI + Local News Challenge	USD 7,500
2021	Google Research India	INR 85,000

TALKS

ORGANIZING

MEDIA

PUBLICATIONS

WORKING PAPERS

[1] S. Mehta, A. Nichols, N. Mazar, M. van Alstyne. 'Market Design Interventions for Safer Agentic AI'. *Working Paper; IC2S2 2025; Oral Talk, Yale AI, ML, and Business Analytics Conference, 2024.*

[2] H. Ranka, D. Shah, L. Ng, S. Mehta. 'Identifying Intervention Targets for Disinformation on Telegram'. *Working Paper; Poster at Stanford T&S Research Conference, 2024.*

[3] S. Mehta, Z. Yang, C. Fan, D. Liu, N. Mazar, M. van Alstyne. 'Truth Warrants Increase Welfare and Accelerate Sales in Digital Markets'. *Working Paper, Poster at AMA Marketing and Public Policy, 2025; Workshop on Information Systems and Economics, 2024.*

[4] R. Konka, P. Kurani, S. Mehta., L. Ng. 'Multimodal Analysis of Instagram Marketing Strategies Across Developed and Developing Countries'. *Working Paper, 2025.*

[5] G. Malpani, S. Mehta. 'The Effect of Influencers on the Independent re-sharing of News in an Alt-Right Ecosystem'. *Working Paper, 2025.*

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

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

[8] 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 2024.*

[9] 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, 2024.*

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

CONFERENCES

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

[12] C. Vergara, J. Lalwani, H. Ranka, N. Kothari, and S. Mehta. 'A Principled Framework to Organize and Develop Social Media Interventions'. *Research Conference on Communications, Information, and Internet Policy, 2025..*

[13] 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.*

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

[15] 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.*

[16] 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.*

WORKSHOP PAPERS

[17] D. Shah, S. Shetty, S. Mehta. 'How User Controls Affect User Activity on Decentralized Platforms'. *Working Paper; Poster at Stanford T&S Research Conference, 2024.*

[18] 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 (IC2S2), 2024.*

[19] 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.*

[20] 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.*

[21] 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.*

[22] 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.*

[23] 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.*

[24] 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.*

[25] 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 REPORTS

[25] 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.*

[26] 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.*

[27] 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.*

[28] 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.*

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

[30] 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.*