

Swapneel Mehta

swapneel.mehta@nyu.edu | [Linkedin](#) | [Github](#) | [Google Scholar](#)

EDUCATION

Ph. D. in Data Science, Center for Data Science, New York University, GPA 3.76 **2019 - 2023 (exp.)**

- Probabilistic Programming and Inference, Causality, and Social Network Analysis

Bachelor in Computer Engineering, University of Mumbai, GPA 9.23 **2014 - 2018**

- Applied Math, Graph Theory, Machine Learning, Artificial Intelligence, Algorithms

EXPERIENCE

Research Intern, Google Research (*upcoming, 2023*)

Founder, SimpPPL, Audit the Spread of Online Disinformation **Jan '22 - Present**

Developing [parrot.report](#), part of [simppl.org](#) for policymakers and journalists to audit online disinformation. Working with The Sunday Times (UK), Deutsche Welle, Yale Daily News, and One Fact Fdn. on disinformation. Awarded grants from AWS, Google Cloud, the Wikimedia Foundation, the [NYC Media Lab](#), and [JournalismAI](#).

Ph.D. Machine Learning Engineer Intern, Twitter Inc. **June - Aug'22**

Worked with Civic Integrity and sister teams at Twitter Health, improving the precision of their civic and health misinformation classifier (BigQuery ML) by 30% to be deployed into production to score millions of tweets a week.

Visiting Researcher, Oxford University and the Integrity Institute **Aug '21 - June '22**

Simulating the impact of influence operations on ranking and recommender systems on social networks. First-author workshop paper presented at the [AI4ABM](#) at ICML 2022; Misinfocon at [DEFCON](#) 2022.

Research Asst., Center for Social Media and Politics (CSMAP), New York **Feb '21 - Present**

Estimating the causal impact of Twitter's interventions on Donald Trump's tweets presented at [Stanford T&S Conf](#). Estimating the causal effects of Reddit's platform downtime on the spread of low quality news online (ongoing). Estimating the effects of coordinated behavior on recommender systems on social networks using simulations. Modeling policy interventions for epidemiology models accepted at [PROBPROG 2021, MIT](#).

Data Science Research Intern (Summer + Fall), Adobe Inc., San Jose **May - Dec '20**

Trending Hashtag Recommendation for Videos; web-scraping, weakly-supervised learning, graph neural networks. First-author [paper](#) accepted IEEE Intl. Symposium on Multimedia 2021; U.S. Patent App: P10634-US.

Technical Student, European Org. for Nuclear Research (CERN), Geneva **Mar '18 - Jun '19**

Built graph neural networks for particle track reconstruction using collision data from the Large Hadron Collider. Built, maintained, and deployed the '[DeepJet](#)' deep learning package into production across 42 global sites. Received a Google Cloud Award and the Oxford MLHEP Award for our work ([preprint](#)).

Openlab Intern, European Org. for Nuclear Research (CERN), Geneva **June '17 - Aug '17**

Developed an ensemble of unsupervised learning models for anomaly detection in database connections. Published a first-author paper, receiving a special mention in the ML track at Springer ICIICS, 2018.

Remote Research Intern, Dr. S. Sahasrabudhe, IIT-B, Mumbai **Aug '16 - May '17**

Automated the evaluation for 3D Modeling assignments in Blender; Course Instructor, edX Prize Finalist, 2019. Published a first-author short paper at IEEE International Conference on Advanced Learning Technologies, 2018.

Swapneel Mehta

SWE Intern, Smokescreen Tech. (acq. by Zscaler Inc.), Mumbai

Nov '16 - Mar '17

Combined low-latency network scans, hostname clustering, intelligent decoys, and automated deployment using custom-built modules integrated into the NASSCOM-DSCI award-winning IllusionBlack software, 3rd Gen., 2017.

SWE Intern, Falconry Software Solutions, Mumbai

July '16 - Sept '16

Developed and deployed real-time data streaming and subscription-based alerts in the three client connectors released as open-source codebases in [Java](#), [Python](#) and [Javascript](#).

Summer Intern, Microsoft Research, IIT Bombay, Mumbai

May '16 - July '16

Created a Google Analytics Dashboard to track and monitor user interactions on the “[Lokacart](#)” app. Presented the project to Microsoft Research and a committee of IIT-B faculty ([video](#)).

SELECTED PUBLICATIONS

- **Oral Presentation.** Estimating the Impact of Coordinated Inauthentic Behavior on Content Recommendation in Social Networks, S. Mehta et al., AI for Agent-based Modeling (AI4ABM) Workshop, ICML 2022 ([slides](#)) and Misinfocon, DEFCON 2022 ([blog](#)).
- **Oral Presentation.** Estimating the Causal Effect of Twitter's Interventions on Trump's Tweets, S. Mehta, J. Bisbee, J. Nagler, J. Tucker, Stanford Trust and Safety Research Conference, 2022 ([slides](#)).
- **Open-domain Trending Hashtag Recommendation for Videos**, S. Mehta, et al., IEEE International Symposium on Multimedia (ISM), 2021.
- **Compartmental Models for COVID-19 and Control via Policy Interventions**, S. Mehta and N. Kasmanoff, International Conference on Probabilistic Programming (PROBPROG), MIT, 2021.
- **Special Mention.** **A Big Data Architecture for Log Data Storage and Analysis**, S. Mehta, P. Kothuri, D. Garcia. Integrated Intelligent Computing, Communication and Security, 2019.
- **Anomaly Detection for Network Connection Logs**, S. Mehta, P. Kothuri, D. Garcia, 2018.
- **DeepJet: A Machine Learning Environment for High-energy Physics**, S. Mehta et al., 2018.

CONTRIBUTED TALKS AND EVENTS

- Co-organizer, [Workshop on AI for Agent-based Models](#), ICLR 2023 (upcoming)
- Co-organizer, [Workshop on Broadening Research Collaborations in Machine Learning](#), NeurIPS 2022
- Lead Organizer, [Workshop on AI for Everyone](#), Columbia Computation + Journalism Conference, 2022
- Board of Studies Member, Data Science and AI, D.J. Sanghvi College of Engg., 2021-Present
- **Invited Talk**, Detecting Coordinated Inauthentic Behavior Promoting Misleading Narratives Online, 2022
 - Graphika, ([slides](#)); Truth and Trust Online ([slides](#))
- **Invited Talk**, SimPPL: Using simulations to Combat Online Disinformation, 2022
 - MIT Media Lab ([video](#)); Misinfocon at DEFCON ([slides](#)); Oxford University
- **Invited Talk**, Interventions to Limit Misinformation in Social Networks, 2021
 - [Twitter Cortex](#); [Facebook AI Research](#)
- Tutorial on Probabilistic Programming, Computational Biology Group, Flatiron Institute, 2021.
- DeepJet: A Machine Learning Environment for High-energy Physics
 - Invited Poster Presentation at Nvidia GPU Tech Conference 2019, USA;
 - [Google Cloud Poster Award](#) at the Deep Learning Indaba 2018, South Africa;
 - Invited Tutorial at the CMS Machine Learning Workshop, July 2018.
 - Oral Presentation at the 2nd CERN IML Machine Learning Workshop, 2018;
 - Invited Talk, CVIT Lab, International Institute of Information Technology (IIIT), Hyderabad, India.

Swapneel Mehta

AWARDS

- WikiCred Grants, Wikimedia Research Foundation (USD 10,000)
- News Product Alliance Mentorship Program 2022 - Mentor: Alyssa Zeisler, VP Product, Dow Jones
- Digital Forensics Lab - Digital Sherlocks Program 2022
- Algovera AI Research, Community Grant Rounds #5, #7, #8 (USD 3000)
- Amazon Web Services Research Grant, Cloud Credits (USD 5,000)
- Google Cloud Research Award, Credits (USD 1,000)
- NYU Tech Ventures Workshop, 2022 (USD 1,000)
- Finalist, [AI + Local News Challenge](#), NYC Media Lab, 2022 (USD 7,500)
- UKRI Research Grant Funding, Oxford University (at the [Torr Vision Group](#)), 2021-22 (GBP 23,200)
- Google AI Research - Funding for [Unicode's ML Summer Course](#), 2021 (INR 85,000)
- IRIS-HEP Fellowship, 2020-21
- Center for Data Science Fellowship, 2019-24
- 2nd Prize, CodaLab Contest, Machine Learning in High-energy Physics Summer School, Oxford, 2018
- Economic Times Campus Stars Class of 2018 - Selected as one of **India's Top 33 Engineers**
- Finalist, Stanford Graduate School of Business, Reliance Scholarship 2017 (declined; USD 150,000)
- Indian Academies of Science Fellowship, 2017 (declined)

COMMUNITIES

- **Founder, Unicode Research, 2019 - Present:** Mentoring 25+ students to conduct independent, collaborative research on causal inference, social science. Created and taught a **Google Research India**-backed 10-week [machine learning summer course](#) for 100+ underserved Indian undergraduate students ([team](#)).
- **Lead Organizer, NYU AI School, 2020-23:** Teaching AI/ML to 400+ participants primarily from underserved communities and non-CS majors, across 25 countries. Fundraising over \$25,000 ([website](#), [Twitter](#), [LinkedIn](#)).
- **Panelist, NYU Center for Data Science, 2020 - Present**
Panelist, 'So you want to go to Grad School'; Student Panel: Ph.D. Orientation; Student Careers Panel
Podcast, Graduate School, 2020; [CERN](#), [Data Science](#), and [Ph.D.](#), 2021
- **Co-founder, DJ Unicode, 2017 - Present:** SWE training to 210+ undergraduates who are winning international hackathons, working at Google Summer of Code, CERN, MAANG, at Ivy League schools ([website](#), [Github](#)).
- **Research Mentor, Lumiere Education (Harvard/Oxford), 2020 - present:** Mentoring high-school students pursuing research on news sentiment analysis, conversational AI, politics, and causal inference ([webpage](#)).
- **Mentor, The Grand Challenge, CERN/RCA, 2018:** Guided 3 teams of students from the Royal College of Art in London on leveraging AI/ML for designing sustainable solutions in science for [The Grand Challenge](#).
- **Contributing Author**, Open Source For You Magazine, 2017 - 20: Reaching 200,000+ readers ([website](#)).