

KUNAL RELIA

Ph.D. Candidate - New York University

[updated: July 15, 2021]
NYU Tandon - 370 Jay St., 1122 F
Brooklyn, NY 11201, USA
☎ +1 (347) 433 0825
✉ krelia@nyu.edu
🌐 www.kunalrelia.com

EDUCATION

- Sep 2017 - Present **Ph.D., Computer Science**,
NYU Tandon School of Engineering, Brooklyn - USA.
Advisor: Julia Stoyanovich, Ph.D.
- Jan 2015 - Dec 2016 **M.S., Computer Science**,
NYU Tandon School of Engineering, Brooklyn - USA.
- Jul 2010 - Jun 2014 **B.E., Computer Engineering**,
Gujarat Technological University, Rajkot - India.

RESEARCH INTERESTS

Computational Social Choice, Algorithmic Fairness, Responsible Data Science,
Computational Social Sciences, Spatiotemporal Analysis.

PROFESSIONAL EXPERIENCE

Research Experience

- Jun 2016 - Present **Research Assistant**, NYU Tandon School of Engineering, Brooklyn - USA.
Research Projects: See page 4

Non-research Experience

- Sep 2015 - May 2016 **Team Lead/Analyst**, NYU Office of Graduate Admissions, Brooklyn - USA.
Led a team of 21 graduate assistants; cut application processing time by 15%, and augmented a transcript evaluation algorithm
- Sep 2015 - Dec 2015 **Teaching Assistant**, NYU Tandon School of Engineering, Brooklyn - USA.
Assisted Prof. Viji Srinivasan (Computer Architecture 1)
- Jun 2010 - Jan 2015 **Co-founder**, ReLife Pharma, Rajkot - India.
Co-founded a health care start-up manufacturing and marketing medicines in India
- Targeted the untapped customer base of doctors in Tier 3 towns of Gujarat state
 - Manufactured medicines at World Health Organization - Good Manufacturing Practices (WHO-GMP) certified facilities to ensure best quality products; only 1 in 10 facility in India is awarded this certificate
 - Turned profitable from the first year itself managing sales in excess of \$150,000 (within top 7 percentile of Indian income)

- Jul 2013 - Jun 2014 **Software Developer Intern, Johnson & Johnson Ltd., Rajkot - India.**
Conceptualized and developed a claim management software capable of managing annual claims worth \$4.5M of company's 1000+ dealers
- Automated the claim generation and management process reducing the required human hours by 80%

AWARDS

- Award **School of Engineering Fellowship - NYU Tandon, Sep 2017.**
- Award **Graduate Scholarship - NYU Tandon, Jan 2015.**
- Award **Academic Excellence Award - GTU, Jun 2014.**

PUBLICATION

[C] : conference; [J] : journal; [S] : in-submission; [P] : in-preparation.

* : authors listed alphabetically.

- [P1] **Kunal Relia.** "*DiRe Committee: Diversity and Representation Constraints in Multiwinner Elections*".
Presented at the 8th International Workshop on Computational Social Choice (COMSOC) Poster Session, 2021.
- [J1*] Vishal Chakraborty, Theo Delemazure, Benny Kimelfeld, Phokion Kolaitis, **Kunal Relia**, and Julia Stoyanovich. "*Algorithmic Techniques for Necessary and Possible Winners*." In ACM/IMS Transactions on Data Science (2021).
- [C5] **Kunal Relia**, Zhengyi Li, Stephanie H Cook, and Rumi Chunara. "*Race, Ethnicity and National Origin-based Discrimination in Social Media and Hate Crimes Across 100 U.S. Cities*." Proceedings of the Thirteenth AAAI International Conference on Web and Social Media (ICWSM-2019).
Media: [Market Insider](#), [Business Insider](#), [The Register \(UK\)](#), [VICE](#), [The Philadelphia Inquirer](#)
- [C4] Nabeel Abdur Rehman, **Kunal Relia**, and Rumi Chunara. "*Creating Full Individual-level Location Timelines from Sparse Social Media Data*." ACM SIGSPATIAL GIS'18.
- [C3] **Kunal Relia**, Mohammad Akbari, Dustin Duncan, and Rumi Chunara. "*Socio-spatial Self-organizing Maps: Using Social Media to Assess Relevant Geographies for Exposure to Social Processes*." Proceedings of the ACM on human-computer interaction 2, no. CSCW (2018).
- [C2] Mohammad Akbari, **Kunal Relia**, Anas Elghafari, and Rumi Chunara. "*From the User to the Medium: Neural Profiling Across Web Communities*." Proceedings of the Twelfth AAAI International Conference on Web and Social Media (ICWSM-2018). (*poster paper*)

- [C1] Tom Huang, Anas Elghafari, **Kunal Relia**, and Rumi Chunara. "*High-resolution temporal representations of alcohol and tobacco behaviors from social media data.*" Proceedings of the ACM on human-computer interaction 1, no. CSCW (2017).

MENTORING

- 2021 **Eunice Son and Afifa Tanisa**, *High-school students - NYU ARISE.*
2020 **Raisa Bhuiyan and Rachel Rose**, *High-school students - NYU ARISE.*
2019 **Kennan Gumbs and Heeyun Kim**, *High-school students - NYU ARISE.*
Kennan is now at MIT.
2018 **Zhengyi Li**, *Undergraduate student.*
Now Team-lead at Opine.

TALKS

- 2021 **COMSOC Seminar Series**, *Online International Seminar.*
Rump Session talk introducing the idea of the "Five Ws" to the COMSOC community
2017 **Natural Language Processing (NLP) Community Reception**, *NYU Center for Data Science, New York - USA.*
Lightening talk discussing the implementation of NLP techniques in social science applications that use social media data (Socio-spatial Self-organizing Map project)

COMMUNITY SERVICE

ACADEMIC.

- Program Committee AAAI ICWSM (2020, 2021).
Reviewer AAAI (2021), AAAI ICWSM (2020, 2021, 2022), ACM CHI (2020, 2021), ACM CSCW (2019, 2020, 2021), ACM WebSci (2020).
Selection Committee NYU ARISE (2020, 2021).

NON-ACADEMIC.

- Volunteer Part of UNICEF's "Leadership Circle" and "Guardian Circle".

RESEARCH PROJECTS

Sep 2019 - Present

The “Five Ws” of Fairness in Multiwinner Elections.

Fairness has recently received particular attention from the computer science research community. However, context-specific uses of a complex term like “fairness” narrate an incomplete story, possibly doing more harm than good. For example, an unintended consequence of fair elections with respect to candidates can be unfairness to voters, and vice versa. Hence, we propose to move towards completing the story of fairness by answering the *Five Ws*, namely, the *Who*, *What*, *When*, *Where*, *Why*, and *How*, which have been used in fields like scientific research, politics, and journalism to ensure thoroughness of the work done. We use the context of multiwinner elections, also known as committee selection, and advance to formally answer each question by designing algorithmic frameworks and giving axiomatic guarantees wherever possible. Our initial results show that finding a committee under our framework is generally computationally hard, and, for such cases, we study the inapproximability and parameterized complexity of the problems. As part of our work, we show that efficient algorithms exist under some real-world assumptions, such as when preferences follow a particular kind of structure. In addition to theoretical results, we conduct extensive empirical analysis. Overall, using multiwinner elections as an example, we show that any “fair” system with an incomplete story can have adverse outcomes, and with an advent of automated decision systems in all fields, even a context-specific “fair” system can unknowingly impact millions of people negatively. Immediate action is required to ensure that algorithm designers understand the responsibility associated with the claim of “fairness”. The goal of this work is to move towards a more complete story in this important domain.

Apr 2020 - Present

Demographic, Social, and Physiological Vulnerabilities related to Development of COVID-19 Symptoms: A Prospective Digital Health Study.

On July 17, 2020, the CDC updated their list of illnesses that are risk factors for COVID-19. Factors with “strongest and most consistent evidence” include hypertension, heart disease, cancer, and obesity, and demographic vulnerabilities are age, sex, minority race, and low SES. Social and psychological factors like stress, discrimination, and depression are less examined but may be critical factors in the disease contagion processes. In this project, we examine (i) psychophysiological vulnerability factors related to developing COVID-19 symptoms, (ii) mitigating factors associated with reduced symptomatic features, and (iii) the effect of randomly assigning people to stress reduction techniques aimed at reducing blood pressure or a control group.

Jan 2019 - Present

Algorithmic Techniques for Necessary and Possible Winners.

We investigate the practical aspects of computing the necessary and possible winners in elections over incomplete voter preferences. In the case of the necessary winners, we show how to implement and accelerate the polynomial-time algorithm of Xia and Conitzer. In the case of the possible winners, where the problem is NP-hard, we give a natural reduction to Integer Linear Programming (ILP) for all positional scoring rules and implement it in a leading commercial optimization solver. Further, we devise optimization techniques to minimize the number of ILP executions and, oftentimes, avoid them altogether. We conduct a thorough experimental study that includes the construction of a rich benchmark of election data based on real and synthetic data. Our findings suggest that, the worst-case intractability of the possible winners notwithstanding, the algorithmic techniques presented here scale well and can be used to compute the possible winners in realistic scenarios.

May 2018 - Jan 2019 **Discrimination in Social Media and Hate Crimes Across 100 U.S. Cities.**

We study malicious online content via a specific type of hate speech: race, ethnicity and national-origin based discrimination in social media, alongside hate crimes motivated by those characteristics, in 100 cities across the United States. We develop a spatially-diverse training dataset and classification pipeline to delineate targeted and self-narration of discrimination on social media, accounting for language across geographies. Controlling for census parameters, we find that the proportion of discrimination that is targeted is associated with the number of hate crimes. Finally, we explore the linguistic features of discrimination Tweets in relation to hate crimes by city, features used by users who Tweet different amounts of discrimination, and features of discrimination compared to non-discrimination Tweets. Findings from this spatial study can inform future studies of how discrimination in physical and virtual worlds vary by place, or how physical and virtual world discrimination may synergize.

Sep 2016 - Nov 2018 **Socio-Spatial Self-Organizing Map.**

Historically, neighborhoods have been defined using administrative boundaries like Zip codes. But, it is a growing understanding in social sciences that what we experience around us is different from what is captured by such administrative boundaries. Hence, we develop a novel pipeline, Socio-Spatial Self-Organizing Map (SS-SOM), that uses freely-available, sparse, geo-tagged social media data to assess relevant geographies to measure exposure to social processes. The pipeline uses (i) shallow neural network to classify Tweets, (ii) followed by dividing the city into grid-cells, and (iii) then using an augmented version of the Self-organizing Maps to create contiguous, non-overlapping, homogeneous clusters. We use prevalence of racism and homophobia in New York City as example social processes to measure the change in exposure to these social processes between Zip codes and SS-SOM clusters.

Sep 2016 - Oct 2018 **Filling User Timeline Using Sparse Social Media Data.**

A stochastic framework for predicting individual level mobility timelines using sparse location data from social media. The framework utilizes individual and community mobility patterns and prioritizes the effect of location data closer in time, to make predictions.

Jun 2016 - Dec 2016 **Predicting Age and Gender of Social Media Users.**

Demographic attributes like age and gender are considered important covariates in public health study. Moreover, while data from social media platforms like Twitter is increasingly used in public health research, companies' policy of not sharing users' demographic attributes act as a bottleneck in such analysis. Hence, we use the text of the Tweets made by users to infer their age using modern NLP techniques, and use user-names to infer the gender of the users.