

I am interested in researching ways to improve social processes and democratic governance with design and policy implications. An area I hope to explore is **computational social choice**, where I hope to identify structures of social networks to design robust mechanisms in order to support collective decision-making. I want to work with faculty who work in **algorithms and theory** and **social computing**. At NYU Tandon, I find the work of Professors Rumi Chunara, Damon McCoy, and Julia Stoyanovich compelling and relevant to my interests in fairness in sociotechnical systems and crowd dynamics. My past research was in **Human-Computer Interaction (HCI)**—online movements and mobile computing. I am working on Bayesian inference for my senior thesis.

My sophomore year, I was awarded Clark's LEEP Fellowship to investigate the #TeamTrees campaign on YouTube which fundraised for 20 millions trees with Dr. Shuo Niu. I built a data sanitization pipeline and performed Chi-squared tests to establish correlations between video popularity, recruitment methods, among other factors. Our work quantifies the impact of celebrity organizers in mobilizing communities that are usually uninvolved with activism campaigns, such as gaming or DIY communities. As our annotation scheme was not a textbook case, I proposed a pipeline using modified Kappa to measure coder's agreement which enabled multivariate analysis. Our paper was accepted at CSCW '21 [1].

Next, I leveraged the project's pipelines to analyze videos created during COVID-19. I released a YouTube crawler package that scraped video metadata and parsed JSON format, which was used across multiple projects in Dr. Niu's lab. Previously, my classmate and I used Naive Bayes classification on 5000 YouTube videos created during COVID-19 to see if they engaged directly with the pandemic. As this fit Dr. Niu's interests, we deepened the correlation analysis and applied Weiss' loneliness theory, which classified and determined the efficacy of different methods of online social integration and video-sharing as a pathway to social support. This work resulted in a paper at CHI '21 [2]. We expanded our scope on loneliness support by analyzing content surrounding drug addiction discourse on YouTube [3].

In a different flavor of HCI, in summer 2021 I worked with Dr. Jeff Huang at Brown HCI to use back-of-device pressure readings and 6DOF motion replay to predict users' emotion and attention levels. I performed cross-validation studies, trained CNN and Random Forest models, and analyzed performance increase with respect to pressure readings and phone motion. I co-authored and significantly rewrote a previous paper, reformulated the main research questions and incorporated our new findings. The work quantified how humans used non-contextual data to predict affect and the significance of using back-of-device pressure to infer immediate affect during phone use. We submitted the work to IMWUT in which I was the corresponding author

and presented the research at Clark's 2021 symposium.

While my research on social media campaigns has offered many insights about campaign propagation and structure, I realized sole reliance on retrospective analysis can sometimes lack predictive power and clear design implications, aspects that I have found myself more drawn towards. I became interested in exploring questions such as, How to combine insights or assumptions from smaller populations to build models to apply to much larger-scale data? How to understand actual demographics of online engagement when most of the data is collected from active participants? How do we design the optimal algorithm when social processes are non-static and have many confounding factors? I saw this as an opportunity to expand on my theoretical CS and Math foundations to investigate paradigms for analyzing crowd dynamics that excites me. I found this in computational social choice, with an eye towards applications in social networks analysis and civic participation.

For my senior thesis, I wanted to develop robust mathematical frameworks and deepen my understanding of statistical methods. I am working with Prof. Michael Satz on Bayesian inference through Markov Chain Monte Carlo (MCMC), examining different MCMC sampling algorithms, e.g. Metropolis-Hasting, to assess their convergence to target distributions. To maximize computational efficiency, I investigate reparameterization and different jumping rules. For complex (e.g. high-dimensional) target distributions, we find that traditional MCMCs remain inefficient because of their random walk behavior. Instead, we focus on Hamiltonian Monte Carlo, where jumping rules embed the gradient of the log-posterior density at every step. This combines conventional MCMC apparatus with deterministic calculations, resulting in faster convergence in complex models. In class, I had learned and proved MCMC's properties, but to be able to formulate a problem myself for MCMC to work has been incredibly fascinating.

My senior thesis has confirmed that I enjoy working in algorithms and my HCI experience has cultivated my inclination towards social computing. I'm fascinated by Dr. Chunara's works in equity in public health and investigations of telemedicine's impacts in healthcare systems. Dr. McCoy investigates online systems for privacy and security, and my works align with his recent works in online harassment and misinformation. I am intrigued by the security approach to measuring the web which I haven't sampled before. Dr. Stoyanovich also works in promoting privacy in data science research and conducted awesome works at the intersection of database and social choice. Therefore, I want to pursue my PhD at NYU Tandon because of many faculty working on information economics and sociotechnical systems and the collaborative environment with other engineering disciplines that is vital for interdisciplinary research.

## References

- [1] S Niu, **C Mai**, KG McKim, and S McCrickard. ”#TeamTrees: Investigating How YouTubers Participate in a Social Media Campaign.” In Proceedings of the ACM on Human-Computer Interaction 5, no. CSCW2 (2021) [\[pdf\]](#)
- [2] S Niu, A Bartolome\*, **C Mai\*** and NB Ha\*. #StayHome #WithMe: How Do YouTubers Help with COVID-19 Loneliness?. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. [\[pdf\]](#)
- [3] KG McKim, **C Mai**, D Hess, and S Niu. ”Investigating Drug Addiction Discourse on YouTube.” In Companion Publication of the 2021 ACM Conference on Computer Supported Cooperative Work and Social Computing. [\[pdf\]](#)