Research Statement

Kiran Garimella

MIT

My research interest is in the area of Computational Social science, which contributes to the understanding of large-scale human behavior on social networks based on solid theoretical foundations from computer science and use this understanding to solve critical societal problems of our time. I pursue this goal by developing a unique end-to-end research methodology, which can be summarized under two main categories: (i) conducting inter-disciplinary research that is based on solid formalisms & analysis, developing ethically responsible algorithmic techniques to collect and analyze societal problems such as political polarization, misinformation and human migration at scale; (ii) building tools to showcase the applicability of the developed techniques on real data and enable researchers from other disciplines such as political science, demography and journalism to easily access the data and methods. I will showcase three such examples in this statement, deliberately focusing on problems rather than on the methods used to solve these problems. This is not exhaustive of my work, but is designed to give a view of the focus and range of my work.

Polarization on Social Media: My PhD research focused on one of the key societal problems we face, political [or social] polarization. Polarization refers to the segregation within a society that may emerge from differences in opinion on various topics. Increased polarization in the society leads to balkanization, where groups of people enclose themselves in their own bubbles, leading to less deliberation and discussion, which might be dangerous for a democracy. My PhD thesis contributes to designing a comprehensive algorithmic framework for **identifying, tracking the evolution of and reducing online polarization**. This work [5, 6] got **two best student paper awards** at top conference venues in computer science and have been cited over 500 times in the last 3 years.

Misinformation: One of my primary projects at MIT has involved studying chat based messaging platforms such as WhatsApp and Telegram. These platforms are hard to access and monitor because of their private nature, lack of a public API and due to end-to-end encryption. However, platforms like WhatsApp are important components of social media especially in the global south, used by over 2 billion users every month. There have been high profile incidents of rumors killing dozens of people in India, or wide spread election manipulation through WhatsApp in Brazil. The COVID-19 pandemic has exacerbated such concerns with rampant health misinformation. Due to their encrypted and private nature, there is currently no way of knowing how bad the problem is, until we see real world consequences such as people getting killed or election manipulation. My research fills in the critical gaps in understanding these closed messaging applications by making three contributions.

- First, I developed a range of tools and techniques that enable the possibility of obtaining WhatsApp data at scale. This involved a wide range of innovations from solving technical challenges to automate working with Android phones, to conducting large-scale field surveys to obtain data from political workers in India to setting up tiplines encouraging users to submit messages for fact-checking. To enable real world impact of our research, I built WhatsApp Monitor [8], which provides dashboards for journalists to track and fact-check content flowing through WhatsApp. Our tool was **used by over 100 journalists and fact-checking organizations** during the Brazilian and Indian elections and was the source of many fact-checking stories.

- Second, I developed methods to study the prevalence and spread of *image* misinformation at scale on

WhatsApp. I specifically focused on images because they are the modal form of communication on WhatsApp making up roughly 40% of the content shared.

- Finally, I developed **solutions to identify and mitigate misinformation** [10], along with auditing whether the steps taken by platforms to curb misinformation work [1]. Mitigating misinformation on encrypted chat applications is non-trivial and can not be done using existing content moderation tools. We proposed an on-device fact-checking architecture that could help identify and flag such misinformation in real time before the content is encrypted [10].

Given the emergence of artificial intelligence assisted and nation state driven disinformation campaigns, our research is timely and provides the necessary insights for researchers, policy makers and journalists to act. My research fills the void in existing literature, looking at closed platforms, from a non-US context, studying non-textual content.

Human migration: The idea of using computational techniques for solving societal challenges also applies very well to demography, where traditionally, lack of access to data is a big issue. Survey data is of high quality, but is expensive to obtain and is only available once in a few years. Moreover, some countries, especially in the global south do not have demographic statistics due to lack of resources. To mitigate these problems, I have been working on **enabling the use of digital data sources to study demographic issues** such as migration through compiling diverse data sources as well as several problems that rise in this domain. Recent advances in Computational Social science have shown that digital trace data, such as tweets, email logs, call records, and advertising data can be used to estimate stocks of migrants and their assimilation. Such convenience samples of data can complement existing data sources in demography and also be incorporated in increasingly detailed models of migration flows that make it possible to better understand the dynamics and complexity of migration decisions. Working with an interdisciplinary team of demographers and social scientists, I built tools [3] and techniques [11, 9] to measure human migration statistics using digital data sources. A few demographic challenges I tackled in the past include, studying refugee sentiment by combining Twitter and call record data [7], measuring refugee assimilation [2], gender gaps in technology access [3], and public health [4]. Such models are extremely impactful in creating statistics, and hence develop policies for countries in the global south, which are likely to be impacted in the near future due to climate change and massive urbanization.

Since March 2020, I applied some of the skills I acquired over the years working with demographers to tracking and understand the concerns of people on the COVID-19. I have led two projects in this space, which had a broad impact.

- First, I worked with the Indonesia COVID-19 Impact Observatory at the World Bank, I led the development of an online monitoring system for WhatsApp, Twitter, Instagram and Google (trends) to get the pulse of the conversation about COVID-19. Using results obtained from the social media analysis, we conducted high frequency surveys through Facebook and Google ads. This work makes use of large scale passive monitoring from social media to identify key issues being discussed in real time, with active monitoring from surveys where you can dig in and ask more detailed questions. This allows us to scale up surveys, get targeted insights and that can quickly transfer into policy and helped shape World Bank COVID-19 response policy in Indonesia.

- Second, I am part of a team conducting a global survey on attitudes, beliefs and behaviors towards COVID-19 in collaboration with Facebook, Johns Hopkins university and the World Health Organization. The survey is ongoing and covers 67 countries and consists of over 750,000 respondents to date. The survey fields questions related to re-opening the economy and have been used extensively in informing WHO and policy makers response on COVID-19 in at least 22 countries. To make the estimates representative, we developed novel, privacy-preserving ways to make use of metadata that Facebook has about its users to generate weights for our responses. All the collected data has been released publicly for research purposes.[1]

---

[1]https://covidsurvey.mit.edu

To summarize, the three projects above showcase my ability to collaborate across disciplines, where I bring my expertise from data mining and network science to design algorithms that handle large-scale datasets and collaborate with experts from other disciplines: journalists, policy makers, political & social scientists, and demographers.

Future work

I have established lines of inter-disciplinary research involving multiple stakeholders in areas studying political polarization, misinformation and in digital demography, where I built tools, and systems harnessing the power of large-scale observational data of various types including text, images, mobility and networks. There are immediate next steps which can be done as extensions to some of my current work.

- On the misinformation project, I plan to develop **solutions for content moderation on end to end encrypted platforms**. Because of encryption, the platforms cannot see the content being shared. Hence, such moderation needs to have a different paradigm than the existing one, based on hiring thousands of content moderators. To address this issue, I propose a multi-disciplinary, multi stakeholder approach involving the creation of trusted authorities that are embedded within the communities sharing the misinformation, and, hence, can quickly identify misinformation and act as broadcasters of fact-checking information. Initial trials of such an approach have proven successful. My research is one of the handful with access to data and know-how on the platform and I have built up connections with a broad range of stakeholders in this process, including journalists/fact-checkers, NGOs/community organizations, platforms, researchers and the end users making me uniquely placed to tackle this problem.

- On the computational demography project, I am constantly working on exploring the use of new digital data and tools for enabling the study of migration. For instance, I am currently working on a project using satellite data to quantify the movement of migrants in war zones, making use of the presence of vehicles and their direction of travel during mass displacement events.

- On the methodological side, one area which is missing in my current work is **experimental studies**, which can lead to causal results. I am actively pursuing a few such studies and exploring ways to connect them to my expertise in big data analysis. For instance, we are running a study of around 100k Firefox users (in collaboration with Mozilla) to study whether the YouTube algorithm leads users down polarization rabbit holes.

- Finally, there are also even more interesting challenges of applying some of my research in a different context: the **global south**. The availability of cheap smartphones and access to 3G/4G internet has spawned a massive surge in new users using the web for the first time. It is estimated that around a billion users will enter the world wide web ecosystem in the next 5 years. These are people who were never online, and hence outside the realm of Computational Social science. We as researchers can be at the forefront to lead how this revolution happens and how we can use data for creating social good. In this context, I will continue to work on **creating a safe and sound information ecosystem** by reducing polarization, misinformation, hate, etc. and helping institutions like journalists/fact-checking organizations, and NGOs to fact-check effectively or help develop digital literacy to have community driven fact-checking. These problems have real world consequences in the global south: rumors and misinformation have killed people in India or the case of genocide against the Rohingyas in Myanmar.

This is not going to be a trivial extension of the work we've done in the western world, because we do not have access to certain types of data, knowledge of baselines do not exist. So we will have to build these ground up. There are also fundamental differences in the population itself like low digital literacy due to which people consume content in multimedia such as video/images and audio, and mostly using mobile phones, which create a few interesting technical, social and operational challenges.

In the next five years, I will continue my research in Computational Social science, with a focus on **Technological change, Population Dynamics and Development**. Access to the internet will definitely provide economic opportunities and can improve their livelihood. But it will also bring a lot more if not done properly. For instance, it could lead to an increase in fractures along various sociological lines, increase in inequalities and reinforcing existing systemic biases again certain communities like women or minorities etc. There is ample opportunity to shape the use of technology in a positive way that affects the lives of millions of people. The new ecosystem of apps and services that empowers these users, could lead to the creation and emergence of new institutions such as markets (micro work, virtual trading) digital commerce (AliPay, PayTM, mPesa), and social organization. We already find the transformative role the new ecosystem of online social platforms (e.g. WhatsApp, WeChat) play in such places by helping shape the economic, social and wellbeing of the people who contribute to them. My work will tap into this data revolution and investigate the role of technology in the changing social dynamics at population scale, and make sure such technology will be used for development and improving lives of users.

References

[1] P. de Freitas Melo, C. C. Vieira, **K. Garimella**, P. O. V. de Melo, and F. Benevenuto. Can whatsapp counter misinformation by limiting message forwarding? In *International Conference on Complex Networks and Their Applications*, pages 372–384. Springer, 2019.

[2] A. Dubois, E. Zagheni, **K. Garimella**, and I. Weber. Studying migrant assimilation through facebook interests. In *International Conference on Social Informatics*. Springer, 2018.

[3] K. Haranko, E. Zagheni, **K. Garimella**, and I. Weber. Professional gender gaps across us cities. In *Proceedings of the 11th International Conference on Web and Social Media*. 2018.

[4] **K. Garimella**, A. Alfayad, and I. Weber. Social media image analysis for public health. In *Proceedings of the Computer Human Interaction Conference*, 2016.

[5] **K. Garimella**, G. De Francisci Morales, A. Gionis, and M. Mathioudakis. The effect of collective attention on controversial debates on social media. In *Proceedings of the Web Science Conference*, 2017.

[6] **K. Garimella**, G. De Francisci Morales, A. Gionis, and M. Mathioudakis. Reducing controversy by connecting opposing views. In *Proceedings of the 9th International Conference on Web Search and Social Media*, 2017.

[7] N. Marquez, **K. Garimella**, O. Toomet, I. G. Weber, and E. Zagheni. Segregation and sentiment: estimating refugee segregation and its effects using digital trace data. In *Guide to Mobile Data Analytics in Refugee Scenarios*. 2019.

[8] P. Melo, J. Messias, G. Resende, **K. Garimella**, J. Almeida, and F. Benevenuto. Whatsapp monitor: A fact-checking system for whatsapp. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2019.

[9] H. Nguyen and **K. Garimella**. Understanding international migration using tensor factorization. In *Proceedings of the 26th International Conference on World Wide Web*. 2017.

[10] J. C. Reis, P. Melo, **K. Garimella**, and F. Benevenuto. Can whatsapp benefit from debunked fact-checked stories to reduce misinformation? *Harvard Kennedy School Misinformation Review*, 2020.

[11] E. Zagheni, **K. Garimella**, I. Weber, and B. State. Inferring international and internal migration patterns from twitter data. In *Proceedings of the International Conference on World Wide Web Companion*, 2014.