

Get the hammer out!

Breaking computational tools for feminist, intersectional “small data” research

Radhika Gajjala¹, Ololade Margaret Faniyi², Debipreeta Rahut¹, Emily Edwards³ and Sarah Ellen Ford⁴

¹ Bowling Green State University, USA

² Emory University, USA

³ St. Francis College, USA

⁴ Binghamton University, USA

✉ radhik@bgsu.edu

Abstract

This paper focuses on revealing how the interplay between algorithmic interactions and the intuitive ways humans navigate digital environments can be researched through a multi-method approach to collecting and critically examining data from online platforms. We use a case study that looks at the role that social media engagement by transnational activists, local activists and celebrities played in amplifying an offline protest by group of women in India. Grounded in a critical feminist perspective, this paper uses multiple methods to demonstrate how the amplification of local protesters work through an interplay of human action and platform algorithmics. We conduct an algorithmic ethnography involving the examination of computational systems shaping online interactions. We examine the digital emergence and recognition of the women of Shaheen Bagh as subaltern political agents/subjects. Understanding of the interplay between online and offline visibility and strategic planning is highlighted. We conduct close readings of small data clusters that emerge within big data networks. We challenge the overreliance on big data methodologies and the fetishization of in-person ethnography (Bishop 2018) over digital ethnography.

Keywords: multi-methods; feminist methods; algorithmic ethnography; political agents; critical analysis

1. Introduction

In this paper, we present a case study that explores how transnational and local activists and celebrities leveraged social media to amplify a group of women protestors in India. We used the globally accessible online data content that was focused on the group of offline protestors to conduct an algorithmic ethnography. The group of women and their allies were protesting onsite in Shaheen Bagh, Delhi, India by initiating a sit-in movement in response to the Citizenship Amendment Act and National Registry of

Citizens (CAA-NRC)¹ in India. Their presence was made visible to global audiences largely through social media activity. Our research focus overall is on looking at how the amplification occurred through the intertwining of human-machine actions and logics rather than on the actual offline events. In line with the theme of this special issue, we discuss our research methods, which involved the examination of both quantitative and qualitative data through a critical feminist perspective. The amplification of the local protesters, we note, occurred in an algorithmically shaped online space that lends itself to being examined through what Angèle Christin (2020) refers to as “algorithmic ethnography” which, she defines as “the ethnographic study of the computational systems enabling and shaping online interactions” (p. 109). Conducting such an ethnography required that we use select network visualization tools. These tools, while not strictly data analytics tools, have features that draw from the logics of data science and computational tools for analysis. While data science is more concerned with developing predictive models and solving complex problems, often involving unstructured or big data, data analytics is primarily focused on examining historical data to provide insights for informed decision-making. Both fields play important roles in leveraging data for different purposes, and their methodologies and skill sets do overlap.

Our research team in this sense was neither strictly doing data science nor data analytics but the tools we used – Gephi² and Netlytic³ – had some features that overlap with data analytics. Certainly, they are embedded in algorithmic logics that engage with both data science and data analytics. Given the relative largeness of even qualitatively readable data available on Twitter, we felt it necessary to use these tools to parse out this data. The tools also allowed us to extract Twitter data using Twitter API affordances that were available free of cost to academic researchers before the platform changed its policies and was renamed as “X”. Using this data and network visualizations as a starting point, our team conducted a multi-method investigation into specific scenarios that became visible to us through this algorithmic exploration. In this article, we describe specific themes made visible through a close look at the data through such an approach. As we proceed, we look at how the human-algorithm interaction amplifies particular fragments of the protest movement. The themes we focus on in this paper are centered on:

- a) an example of how we located select transnational social media activists who contributed to the global digital visibility,
- b) the impact of celebrity tagging, and,
- c) uncertain material/virtual archives.

Overall, our epistemological and theoretical approach is shaped by feminist understandings of knowledge as situated and shaped contextually but also by the location of the researchers.

This article is being written by five researchers. While we started our work from a Midwestern United States University in 2019 – we are all dispersed in various parts of the U.S now as this article is being written. Our collaborators from India who had connected with us via various digital tools have contributed significantly to our thinking and we are grateful for their contributions. We also engaged with various transnational activists to help make sense of the data. The co-authors of this article comprise the lead author, a professor of Indian descent and a U.S. citizen, with a long history of immersion online while

¹ The Citizenship Amendment Act (CAA) was passed by the Government of India in December 2019 and was to be implemented along with a proposed National Register of Citizens (NRC) which was to be an official record of legal citizens in India. This register was to include undocumented migrants of Hindu, Sikh, Jain, Parsi, Buddhist, and Christian religious backgrounds from Pakistan, Afghanistan, and Bangladesh who had entered India before 2014. Muslim undocumented migrants were excluded. Protests against the CAA/NRC happened in various parts of India. Not all the protests shared a common point of contention – but the protests at Shaheen Bagh, Delhi were from a low-income Muslim majority community.

² Gephi is an open-source network analysis and visualization software package.

³ “Netlytic is a cloud-based text and social networks analyzer that can automatically summarize textual data and discover communication networks from publicly accessible social media posts.” (<https://netlytic.org/>)

doing feminist research in both offline techno-mediated environments across online digital platforms; a doctoral student currently studying in the U.S., with a background as an activist in Nigeria; two white faculty members in the U.S who teach and research in areas of digital humanities and one Indian doctoral candidate studying in the U.S. The third co-author, who previously worked as a news journalist, is the only member among the five co-authors who was physically in India during the protests in 2019-2020 and the Covid-19 shutdown in March-April 2020. During this time, she actively engaged in parsing the nuances of what was happening by talking with various actors involved. The diversity of our social locations thus allowed us to discuss and debate different angles and possibilities as we examined the data. It is also important to note that our larger group of collaborators includes activists, scholars and onlookers who were present on site and in India during the protests and our work with them shapes how we write here. The different positionalities that we occupy allowed us to engage in discussion and debate productively around the data we were seeing as we pushed each other on our assumptions. As evident, a majority of the researchers and the co-authors of this article are also transnational audiences of this event. Our experience as transnational participants and/or observers therefore shaped the way we conducted the research. Interpretation of online data, therefore, is not solely based on networked data but is also informed by offline interviews with activists and onlookers onsite.

Similar to scholars such as Yu-Shan Tseng (2022) we reiterate that algorithms impact content visibility, filtering of data that is collected. Therefore, we examined algorithmic rankings and layouts alongside different human practices and adopted a reflexive approach that highlights our awareness that back-end algorithms of the data processing and machine learning models also play a role in the telling of the story. Our intention is also to provide a clear understanding of how qualitative and quantitative approaches can be intentionally used to complement each other when analyzing social media text data. We thus employ techniques like social network analysis while also conducting qualitative interviews. The tenets of data feminism (D'Ignazio and Klein, 2020; Poirier, 2021) and situated data analytics (Rettberg, 2020) as well as several other feminist and postcolonial interventions in the use of computational tools for critical research (Risam 2018, Jackson et al, 2020) inform our approach. These tenets require us to center the experience of human in context through a commitment to action and intersectional feminist approach as we proceed with our algorithmic ethnographic investigation. The Twitter datasets were therefore both collected and curated by our research collective.

To describe the methodological process, we specifically draw on our examination of the dataset around the #womenofshaheenbagh, which captures select Twitter activity during the anti-CAA/NRC protests that happened in Shaheen Bagh, Delhi, India (Gajjala et. al, 2023a and Gajjala et. al, 2023b, Edwards et. al, 2024)⁴. The data cast light on the amplification of the protest movement via social media, while the interviews with activists informed our close reading of the digital data. We have solicited permission to reference our interviewees and their participation. However, the institutional review board confidentiality clause and guidelines, requires that we do not name them. Public figures and organizations such as Bollywood celebrities and globally known activist groups are however mentioned. Regarding such public Twitter user profiles, we follow research methods guidelines provided by Richard Rogers (2019), where he notes that “since the social media user agreed to the terms of service, the researcher can fall back on those terms and use them to cover” (p.33).

Following the introduction, literature review and a methods section, we discuss three main themes that are focused on the transnational amplification of the protests that happened in Shaheen Bagh, Delhi. Two sections of the literature review concentrate on outlining literature around digital ethnography, algorithmic ethnography and on feminist and race theorists' examinations of the hierarchies embedded in

⁴ We acknowledge the work of the BGSU “Tooling Around Research Lab” members past and continuing who contributed to our thinking through this project.

algorithmic infrastructures. A third literature review section refers to relevant work about social media use by transnational activists.

2. Literature review

The literature review for this article is structured into four brief sections:

- a) An overview of selected literature advocating for digital and algorithmic ethnographic approaches and emphasizing the need to incorporate both qualitative and quantitative tools for data collection and interpretation.
- b) A brief discussion of the importance of feminist interventions into such research methods.
- c) A summary of relevant literature on the use of social media by transnational activists.
- d) A brief overview of academic publications about the Shaheen Bagh protests.

2.1 From digital ethnography to algorithmic ethnography

Scholars such as Antoinette Rouvroy have called for a problematization of the contemporary “infatuation for Big Data and the presentation or the making of the real through its algorithmic modelling” (Rouvroy in Rouvroy & Stiegler, 2016, p. 7). Madeleine Elish and dana boyd (2018), in turn, have suggested that we “reimagine AI as a new form of qualitative inquiry” to understand the potential of machine learning and data science for big social data research. They claim that “such a reframing suggests new opportunities for methodological development” (2018).

Digital environments, thus, have continually compelled researchers to rearticulate methods with each interface and platform that surfaced. The struggles around defining what the “field” is, for field work, when researching online interaction dates back to the late 1980s and early 1990s. Mann & Stewart (2000) have critically explored some of these earlier debates on the suitability of the internet as a field, and the scholar consensus emphasizing the need for new methodologies to address social interaction and self-presentation in computer mediated communication. Qualitative and quantitative researchers alike have had to reframe what it means to do research when examining online social space and networks. In this reframing – where the Internet is simultaneously seen as social space, global/local cultures, extractable big data, and mappable/quantifiable networks - we see that some researchers have argued for methodological approaches such as digital ethnography (Pink et.al, 2015) and more recently algorithmic ethnography (Christin, 2020; Tseng, 2022).

Sarah Pink et al., (2015) note that experience is integral to ethnographic practice, echoing the logic of feminist and indigenous scholars who have argued that lived experience can be an important analytical lens to understand people and communities. In addition, digital media experiences – whether with the gadgets and appliances, digital platforms or while living immersively in digitally mediated environments as described in Secondlife ethnographies (see for instance, Boellstorff, 2015 and Gajjala, 2011) - are increasingly becoming a part of the research process. Hence, self-reflective awareness of the embodied experiences of living through senses - affective digital materialities (Pink et al, 2017) - are also integral to digital ethnography. The ethnographer’s immersion in the multiple fields – offline and online – shapes the understanding of the overall research site.

Algorithmic ethnography, in a sense is an extension and variation of digital ethnography – as it responds to the move to closed black boxed algorithm-driven gadgets and platforms. Christin (2020) notes that in doing “algorithmic ethnography,” we need to be cognizant of the concept of “platform” and what it entails in conjunction with other digital infrastructures. Ethnographers should examine the data closely while being aware that these digital platforms are never actually neutral. They target users through

profit-oriented business models that are built-in and designed into environments that direct the users' choices (Pedwell, 2022). Ethnographers should also examine metrics such as platform engagement through actual user practices around given platform features such as “stars, likes, hearts, smiling emojis” (Christin, 2020, p. 110). Context collapse and temporal collapses occur when users are affectively engaging each other across time and geography. When examining platforms, algorithmic interfaces as well as the intertwined human-machine logics and intuition through an ethnographic lens, it becomes crucial for both users and researchers to be attentive to and actively engage with “minor affective registers” (Seigworth and Pedwell, 2023, p. 26).

2.2 Data feminism, racial bias, and algorithmic infrastructures

Feminist researchers using data analytics, data science and machine learning tools have argued that data science needs feminism. Catherine D'Ignazio and Lauren Klein (2020), for instance, argue that in the world of data science, power is not evenly distributed. Lindsay Poirier (2021) also emphasizes that to effectively assess data bias; data analysts must possess the ability to scrutinize datasets as cultural products originating from inherently power-influenced semiotic systems.

Scholars have also written about the importance of paying attention to how representations of race, ethnicity and nation are also algorithmically shaped in online platforms. Safiya Noble (2018) and Ruha Benjamin (2019), for instance, effectively demonstrate how algorithmic interfaces themselves are implicated in reproducing and sustaining uneven social hierarchies based in historical oppressions and negative stereotypes of particular groups of people. Benjamin explores the automation of anti-blackness, arguing that racism is in-built in the technology industry and operates through systemized discrimination that renders Black and other people of color illegible within certain technologies including facial recognition software. Her work explored black representation within technoscientific developments. Noble, on the other hand, deconstructs the myth of neutrality and objectivity associated with algorithms. She notes how biases are normalized when they are embedded into algorithmic platforms by their creators.

2.3 Transnational activists use of social media

Regardless of implicit bias of the algorithms and power hierarchies that shape platform cultures, they have nonetheless facilitated new forms and practices of consciousness-raising. Even as private ownership as well as state control and surveillance continue to restrict, shadow-ban and reroute information – these platforms still have allowed the mobilizing of protests. While digital activism has been critiqued as “slacktivism” (Rotman et. al, 2011), it has been a significant driver of transnational digital protest since the Arab Spring (Howard and Hussain, 2014). There also exists a long tradition of transnational, progressive Indian digital protests (Dey, 2020). It is worth noting, however that this literature is based on examination of protests that predate the current moment of surveillance and banning of some types of activist content. Even though the celebration of the role of social media in earlier movements was critiqued – as in the work of Rebecca Mackinnon (2012) and Ramesh Srinivasan (2013) - the critiques were more about access and use than about the power and control of the platforms and private ownership of the platforms. These protests however predate the contemporary moment where platforms are actively working to shadow-ban content from protest movements.

Thus, while scholars have noted that in a post-2007 world, social media does play a key role in mobilization and visualization of local activism to global audiences, they have also expressed concerns around surveillance that have underpinned tactics for transnational digital activism even as far back as during the Arab Spring.

2.4 The Shaheen Bagh protests

While the overall activism around the pass of the CAA – whether against or for the ruling party was spread all over India – the Shaheen Bagh protest movement has engaged the interest of academics around the world. From 2020 on a variety of academic researchers have written on this topic. These include local researchers and activists (Salam, 2020; Mustafa, 2020; Sinha et. al, 2023) as well as internationally located researchers interested in global south activist movements, feminist researchers highlight women’s political activity (Kadiwal, 2021) as well as researchers intrigued by how digital platforms were used during these protests (such as our team).

There were several protests occurring worldwide around this time when the world was also faced with Covid-19 shutdowns and a massive move to online communication as a primary mode of connecting with local and transnational communities. In addition to the CAA/NRC protests in India a few of these included the Black Lives Matter movement (Jackson et al, 2020) mostly in the U.S. and U.K, the Endsars movement in Nigeria (Faniyi, 2023) and the Farmers’ protest in India (Mishra et al., 2021), the continuing protests around the Hijab in Iran (Kermani 2023) among others⁵. These studies shed light on how historically marginalized communities as well as global south activists use Twitter to reach global publics. However, with Twitter’s transformation to “X” the scraping of data is too expensive and inaccessible to academics for research.

3. Methods employed in our project

The tools we used in our visualization of data around the online amplification of the Shaheen Bagh Protests of 2020 were Gephi and Netlytic. Our overall multi-methods approach however was developed through a series of projects. In this paper, however, we primarily focus on datasets collected from our work between December 2019 and the summer of 2020. These datasets shed light on the community-based, women-led protest movement in Shaheen Bagh, Delhi, India, which had gained significant online visibility. We demonstrate our approach by providing illustrative examples in the next section of this paper.

In December 2019, we were exposed to the surge of social media activity originating from India in response to the passing of the CAA/NRC. Our primary challenge at that time was managing the collection of rapidly proliferating digital data. Our motivation for collecting this data initially was to preserve the ephemeral digital content on these social media platforms. We were concerned about the potential loss of all texts related to the Shaheen Bagh protest due to systematic state erasure and Internet blackouts, a trend we had observed. However, we found ourselves asking questions about how the algorithm was throwing up content selectively. We also observed several transnational actors grouping together via message sharing apps such as WhatsApp in attempts to try and amplify content from the protest site. This led us to enter conversation with such transnational actors and to observe their practices within the algorithmic interface.

Algorithmic conditions rely on an intuition that is a “mode of sensing, knowing, anticipating, and navigating the world that exceeds rational analysis” (Pedwell, 2022, p. 3). As Carolyn Pedwell argues, we are seeing a redistribution of cognition through human and technological engagement with machine learning technologies. This suggests that conducting an algorithmic ethnography requires the honing of intuitive reading through and with the platforms and not just what is seen as the symbolic evidence of human output. Further, as Colman et al., (2018), note, the use of algorithms in societies not only results in a change from practical knowledge generation and acquisition but also creates a symbolic logic that materializes into reality. The double effect here is that the algorithmic interface simultaneously and

⁵ While we recognize that our list here is not comprehensive, we have noted some that were globally visible through Twitter and that reflect research interests of our team members.

affectively invites intervention – the user of the tool must intervene by doing different things with the data – while also privileging the symbolic – the text and the statistical results. We therefore develop an approach to algorithmic ethnography that focuses on the interplay between algorithmic interactions and the intuitive ways humans navigate digital environments. We emphasize the tangible impacts of context, affect, and intuition, as well as the underlying logics that either contain or amplify certain bodies online and center the importance of telling this as part of our algorithmic ethnographic narrative.

Our research journey, in essence, formed a cyclical path, commencing with our initial curiosity about Shaheen Bagh. This path then led us to track individual profiles and hashtags, delve deeper into data analysis and visualization, collect relevant contextual information, and ultimately reach a point where we conducted ethnographic interviews. Our subsequent analysis required multiple forms of data collection and a multi methods approach (Murthy, 2017). This approach involved utilizing tools that enabled us to closely examine the networks of users and coalitions and use the information we had gathered through interviews and examinations of news reports. Further, it is an accepted fact that methods for studying digital environments “are often experimental and situational, because they develop in tandem with the medium conditions, and occasionally are built on top of other devices” (Rogers, 2019, p. 49).

We, however, understood that we could not use data analytics tools alone to come to conclusions about how the movement was made visible globally through the use of social media. Our engagement with Twitter involved analyzing both the text and metadata of speech acts that are generated from constantly evolving offline contexts and the users' daily practices, as well as the underlying algorithmic logics that facilitate the textual exchanges on the platform. Thus, we found that taking an either/or approach that dichotomizes research into purely “quantitative” or “qualitative” methods was counterproductive. Rather than isolate knowledge production from the everyday use and development of knowledge by binarizing these two as the only ways to discuss methods, we focused on an iterative process of information gathering to help us with contextual readings. Thus, epistemological concerns (and skepticism toward the flatness of data) guided our approach to the analysis as we gathered evidence through multiple methods. The techniques for data collection followed by methods for organizing the data are therefore linked through an investigation of context and historical understanding. The situated contextual understandings that each of our co-authors and the larger group of collaborators brings to the project are fundamental. Sarah Florini (2019) has emphasized the importance of knowing cultural contexts to understand the specificity of users' practices or how technology shapes users' behaviour.

Donna Haraway in 1988 noted how knowledge production is always already situated. Taking this further into the algorithmic realm, we see that the situatedness is embedded, captured, into the infrastructure and the very ways in which the platforms function. Even if we do not consider the personal socio-economic and cultural location of the researchers –the data itself is very much situated and “biased” even as we scrape it. When using our own devices for data collection, we have already trained the platforms to our own first-person orientation (Elmer et al., 2015), and if we use the platforms on public computers and refuse to log in to any of our accounts while gathering the data, they too are oriented to users somewhere. The gathering of data is always already shaped by the situated knowledge orientations of the tools as much as of the researcher.

For us, as feminist researchers who typically use qualitative methods, the “largeness” of the data was a relative question. However, for this study, “large” meant thousands of Tweets, a division that would eventually show us the arbitrary nature of the dialectic between “small” and “big” data (Strom, 2012; Shahin, 2016). In addition to compiling large datasets, we also employed a feminist methodological principle of close reading of a select number of texts (Lukić and Espinosa, 2011), referring to tweets, Instagram posts, and ethnographic interviews conducted with activists and protesters at Shaheen Bagh in India and in London. Our aim was not to supersede the descriptions provided by the data, but to “thicken” them, allowing for critical interpretations to emerge (Latzko-Toth et al., 2017). For instance, data collection is based on criteria inputted by the researcher, such as the number of retweets or location, even if the said researcher is inattentive to the algorithmic underbelly of digital platforms. Thus, as Rogers

notes, even the collecting of data is very clearly a “search as research” (Rogers, 2019, p. 43). We maintain that this highlights the contextual gaps in pure quantitative big data archiving/analysis, which creates contexts that prioritize data over users, researcher location and often non-static, shifting narratives.

A meta layer of reflexivity on the part of the researcher involves the awareness of how our own “algorithmic thought” shapes our intuition in the process of examining the data even as we also pay attention to the Twitter users and their specific interactive practices that produce the data that researchers scrape (Pedwell, 2019, p. 3). As Pedwell (2019) notes, “we become increasingly algorithmically mediated by digital capital at the micro-level of affect, gesture and habit” (p. 3).

Yet, for our team the problem with Twitter research is not just centered on access to the data itself. It is also an issue of approach and methodology. We have noticed that the data we work with often appears to be “low hanging fruit” (see Burgess & Bruns, 2015) due to the way it is approached. There is a common assumption that it is easy to analyze data using computational software to run statistics or that we can simply produce topic models and sentiment analysis because we can code and run esoteric software program features. Thinking on this further – even the hands-on steps that are set standardized for us – what to “clean” and how to clean the data, for instance, are standards that we have adapted to, not originally from our creations (we are not data scientists, after all), but specific rules that are already configured into the software we engage.

For instance, when looking at the intersection of #womenofshaheenbagh and #sheinspiresus (which was a counterpart hashtag created to undermine and sugar coat the affective circulation of dadis (grandmothers) raging against the Indian state during CAA-NRC protests), we would typically run the “connected components” feature in statistics feature on Gephi to show how strongly or weakly people and other metadata in the dataset are connected via a path. We would then apply the strongest connected partition to the appearance of our visualized graph and filter each group as needed. While our method of filtering and selecting layout algorithms humanizes our methods, we still act as assistants to this software that inducts us into layered algorithmic automation and conditioning of contexts. As noted by Colman et al., (2018), the algorithmic condition drives us towards privileging the symbolic – whether numbers, image, or text. Therefore, while relying solely on traditional textual or quantitative analysis might give us a sense of discovery and accomplishment, it nonetheless keeps us confined to our symbolic environment.

A feminist algorithmic ethnographic approach to using these tools, however, puts the symbolic environments in perspective. We understand that social network data comes to us in personalized modes through algorithms that have come to “know” us, and through interlocutors whom we summon into interviews, even as we try to mitigate the “exploitative power dynamics” (Sobande, et al., 2020) by centering the significance of speaking with people based on who we were able to access for interviews. Thus, not only do we need to be self-reflexive about how we gather and examine our multi-pronged data, but we also need to become aware of how the machine (and even our interlocutors) likely already present data to us in particular and situated ways.

4. Themes visualized

The following key themes show how the different methods we used contributed to our understanding how the online and global visibility of women protesters from Shaheen Bagh was happening through small networks or variously situated and often scattered actors. We note that platform-specific practices and intentional political/activist strategies offline mediate and/or shape the emergence of specific women as globally and publicly visible political agents (see Edwards et al 2024, Gajjala et al 2023a; Gajjala et al 2023b). Examples here are not intended to overshadow nor claim absolute knowledge of the strategy and labor of the offline community members at Shaheen Bagh. Rather, they serve as illustrations of transnational amplification of the movement. The examples here are taken from Twitter data mining tools

provided by platforms such as Netlytic and Gephi which are designed to mine, organize, and visualize large amounts of data.

The themes also reveal productive tensions and communicative gaps as western-located transnational/diasporic activists connect with geographically distant movements and the struggles of the on-site local activists as they seek to reach out to global/digital audiences. The interaction between individual and text takes place through mobile devices, while each of the bodies in a cluster/network itself may occupy physical space elsewhere. Within the social media archive a tweet initially conveys a sense of urgency and prompts immediate responses. It initially enters the public sphere as an individual textual/visual statement. It then remains there, awaiting revival by another user either as retweets or replies to it. The Twitter archive is both a live representation of the present and serves as a record of a spoken statement, essentially offering a "receipt" of what was said. This unique feature of social media, combining immediacy and archiving, has given rise to the practice of "showing receipts" (Brock, 2020, p. 19). These practices underscore the fluid boundary between public and private, as well as the intertwined but sometimes concealed existence of personal and public aspects in social media space. These practices significantly contribute to the strategies employed by protest movements in using social media, particularly in Twitter's public sphere. However, they also introduce an element of unpredictability over time, as different emotional responses to Twitter posts can influence the algorithmic structures in various ways.

4.1 Transnational coalitions amplifying digital visibility

The process of identifying actors, represented as nodes in the datasets, involved an understanding of the movement acquired through various sources. These included informal discussions with activists and researchers, formal interviews, and attention paid to the coverage of certain actors in mainstream news media. In this sense, our work was inductive as we got information and themes from interviews using a constructivist grounded theory approach (Charmaz, 2006) to contextualize what was visible in the Twitter data.

Once we recognized certain actors and patterns through such a process, we moved recursively back and forth between searching for reports and people to interview further (see figures 1 and 2 below). This is how we learned about transnational coalitions that were involved in trying to nudge the algorithm to make certain hashtags "trend." For example, an interviewee told us there had a concerted effort to hijack the #sheinspiresus (a hashtag initiated by pro-government actors) and to connect it visibly with #womenofshaheenbagh (a hashtag for amplifying the presence of the protesting women). They were attempting tweet storms so that the latter hashtag would connect and trend. Yet it is worth noting that numerically the number of tweets and the size of the networks around #sheinspiresus was significantly larger than the size of the networks around #womenofshaheenbagh. Yet the transnational visibility of the women of Shaheen Bagh led to them being profiled in international media outlets such as the Times Magazine. The interplay of traditional media and online media internationally therefore points to media amplification strategies that go beyond social media.

Still, as social media researchers we focused on the claims made in interviews regarding transnational activist efforts to make the hashtag trend. One way we did this was by looking through the networked data for what are known as "in-degree" and "out-degree" centralities (that is, the number of inbound and outbound connections directed to a user that establishes them as an authority in a dataset), and strongly connected clusters around the usernames mentioned in the interviews with members of transnational coalitions of activists.

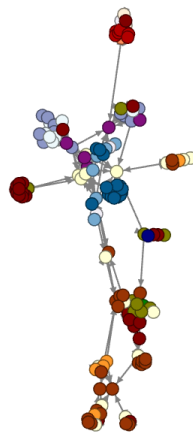


Figure 1. A cluster from the larger dataset after running the modularity algorithm and then the ForceAtlas2 layout. Running this layout helped us spread out the nodes and identify relationships between users and also to see what they tweeted.

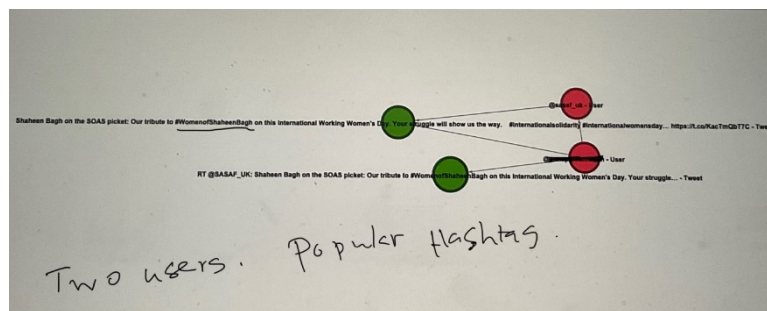


Figure 2. A sample of how we identified users to contact for interviews. We made notes on the figure so in order to map those we would try to interview.

To get at the strongly connected nodes and clusters, we used a range of features in Gephi. First, we ran some statistics through the statistics menu in Gephi – these identified actors with a high frequency of Twitter activity in that cluster and those nodes (not all nodes were Twitter users) connected strongly to each other and so on. Then we ran an algorithm layout called "ForceAtlas2"⁶ layout which spatialized the network to visually spread-out nodes and view them relationally. As a result of running this layout algorithm, the nodes were spread out but also clustered together into communities of connected nodes. Then we filtered out strongly connected node cluster to examine each one even more closely. Figures 1 and 2 show a small cluster that filtered from the larger dataset after we had run the modularity and connected components algorithms. The modularity algorithm is usually run to examine how a network is divided into groups/clusters. We then ran degree and eigenvector centrality algorithms to see which specific nodes seemed to connect to each other through the tagging and retweeting of posts. Tagging and retweeting of course did not necessarily signify consensus or disagreement.

⁶ "ForceAtlas2" is a force-directed layout close to other algorithms used for network spatialization. It is a default algorithm to use on Gephi. We found that running this algorithm and following it with the running of the "Label Adjust" algorithm to be most useful in helping us closely examine smaller clusters of strongly connected notes to observe how they connected with others – whether through being tagged or retweeted or because they were tagging several others and so on.

We were, however, less interested in the numbers here because we were not merely doing a description of the network that became visible from the timebound scrape we did. Generalizing from what we saw as activist networks was not as important to us as trying to see which actors were tweeting and retweeting certain content and identifying whether the tweeting and retweeting was being done in support of or in opposition to the messages coming from the community-based activists in Shaheen Bagh. Thus, what we were doing was detailed *small* data work from parts of the larger datasets. Worth noting here that the relationship of the smaller data clusters within the larger datasets was important, however. It would not have been enough to just copy paste select tweets and examine them as texts outside of the context of the larger datasets. Examining the smaller datasets as part of the larger scenario is what allowed us a contextual understanding of these fractured moments of activism.

The numbers alone do not tell the story of the impact or support for the activists. The success or failure of an offline social justice movement cannot be assessed based on just what is tweeted at any moment in time. In fact, any generalizations about the larger protest movements that we might have drawn just from that section of data would have been wrong. Twitter is a dynamic site and actors and content shift in relation to real life events offline. Within the eco-system of Twitter, the movement might encounter context collapse (Marwick & boyd, 2011) or even time collapse as people find old tweets and restart discussion and debate by bringing it into the current moment (Gajjala et al, 2023). This is also why our discussion of the data is centred on questions of intervention in this discursive public space. Were the tweets targeting a global, international audience or a local Indian audience – or both simultaneously? How were they working with and against the platform affordances? While interviewing a range of local and international people involved in creating social media presence for the movement, we tried to get a sense of what these actors were engaged in both online and offline and how they viewed their contributions to the protest movement.

Figure 2 above gives us a closer look at the smaller section after we ran a “Label Adjust” layout so that the labels were spread out. The nodes and connections were more visible to the naked eye. The visualization in Figure 1 was a screenshot taken to share on a WhatsApp group where we were discussing what our next steps should be. One of our collaborators at that time was from U.K. (we refer to her as “Y” in this article). She recognized nodes belonging to members of the activist group that she herself was a part of. She connected us to members of the activist group, and we together interviewed a few of them. Figure 3 below shows a screenshot of us asking Y to connect us with other members of the group once we realized who the nodes represented.

What is significant here in terms of our methods is that it was not until we visualized the data and observed the clusters of nodes that we fully recognized the reality and actuality of the efforts being made by Y and other transnational actors to amplify #womenofshaheenbagh. Thus, the visualization of the nodes representing these actors served to verify the qualitative statements made during the interviews and conversations. Seeing this visual nodal evidence in the data encouraged us to continue interviews with members of the transnational activist group. We were able to understand and historicize the role of transnational/diasporic progressive activists by doing more in-depth interviews, revealing that these transnational/diasporic activists were not just clicktivists who got involved at a surface level, as is sometimes assumed. Rather, they had a long-standing history of providing support to those fighting against oppression in the South Asian subcontinent.

It must be noted, however, that even though in this paper we only write of progressive anti-fascist groups, the role of diasporas in Indian politics is not limited to these types of groups. The conservative right-wing government has a larger diasporic support structure with significantly more financial backing than those who oppose the government. While we discuss strategies deployed by diasporic feminist activists working to amplify the Shaheen Bagh protests on Twitter from afar, to make the movement visible to a sympathetic global audience, these tactics are not exclusive to progressive collectives. Rather, the social media platforms provide these affordances to all users.

These same principles we have documented are today being utilized by far-right Indian Hindu nationalists coalitions located in the US (Edwards, 2023) There is a vibrant far-right Indian Twittersphere (Bhatia. 2022), and far-right Indian Hindu nationalist users in India and in the diaspora deploy similar strategies to foster transnational connections and visibility (Gajjala et. al in-progress).

Our process demonstrates the utility in deploying a feminist multi methods approach to track and illuminate political movements within and against the algorithmic conditions Twitter presents. Making visible historically marginalized women’s political agency is fraught with obstacles – offline and online, in personal space and public space. So, the why question must be raised – why did transnational and other activists work to create visibility for this group of subaltern women from a low-income community in Delhi? Part of the answer lies in the offline movement’s strategy. The highlighting of a certain image of an older Indian nationalist peace loving, community-oriented woman was strategic⁷ (Gajjala et al., 2023). Thus, it is the offline community-based strategy and actors that ultimately shaped how the movement got amplified. As one of our interviewees noted, this was another strategy to invoke the non-violence resistance of Gandhi to the present times when this protest was termed “anti-national” by certain segments of the present government. Groups like United Against Hate came forth to recognize the “compassionate” and “peaceful” nature of the protest. One such older woman, Bilkis Bano – an elderly woman, became the face of this. Bilkis was included in Time Magazine's list of 100 most influential people in the world in 2020, where she was described as the "voice of the marginalized" and a symbol of resistance.

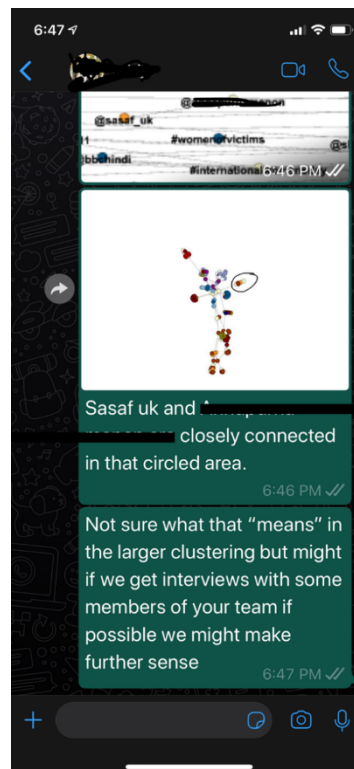


Figure 3. Sample message sent by team leader on WhatsApp to an interviewee and co-researcher discussing the cluster from Figure 1.

⁷ Once again it must be emphasized that “strategy” does not imply insincerity or something that is not real or intentional and spontaneous.

4.2 The internet celebrity

In the previous theme we noted transnational activist interventions. Here we discuss how we were able to visualize the specifics of how the presence of a perceived local celebrity contributed to the amplification. As other researchers have noted when celebrities either participate in Twitter storms or they are tagged by others there is significant amplification for the cause (see for instance Mishra et al., 2021).

Yet the fractured nature of how the protests happened offline was reflected through the data gathered as well. With various spontaneous actors joining the movement either performatively or by coordinating their efforts with the grassroots activists the protest movement was by no means strategically controlled by one set of activists. One of our interviewees (who also agreed to be a co-author on two of our academic publications) told us that groups of activists on the physical site of protest were also fragmented. She noted that individuals participating in the protest became more noticeable not only because of their physical presence at the site but also because of how social media users on both sides of the political spectrum highlighted their offline statements about the protests. The reasons for why certain offline actors were amplified in social media vary but certainly when celebrities appeared at the protest site there was much social media activity about their presence. When celebrities visited the protest site to show their support crowds quickly gathered around them onsite and several people shared this information while tagging the social media handle of the celebrity. In such cases, if those sharing and tagging also used the hashtag connected with the movement, the amplification was significant. Figure 4 is a visualization from our search for the participation of known celebrities – particularly Bollywood actors. Here the one node that tags several Bollywood actors is a Twitter post node that comes from a user who is clearly pro-CAA/NRC. The text actually calls out these Bollywood actresses for supporting the protests. In this instance doing a mere textual/qualitative reading would not have given us an understanding of the larger impact while doing just a quantitative description would have given a wrong impression. The culture of calling out on Twitter leads to this practice – however, since the algorithmic logics of the platform amplify content based on the largesse of follower count among other metrics – the hashtag from the activists still got amplified. The linguistic meaning of the actual tweet – “all these women are part of the ill-famed #womenofshaheenbagh... Losers indeed!” – becomes secondary or irrelevant or differently relevant.



Figure 4. Celebrity users tagged to amplify content. Tweet by a pro-CAA/NRC Twitter user that tags the Bollywood actresses who supposedly support the women of Shaheen Bagh. The text of the Twitter post states: “all these women are part of the ill-famed #womenofshaheenbagh... Losers indeed!”

Thus, we see that even at the local site, the movement happened through a combination of spontaneous and planned actions of various smaller groups and individual actors, including celebrities and students who came to the site. The local community and their ethic of care – as they maintained an inviting and peaceful protest site, handing out tea and snacks to visiting protestors – was particularly noted (Gajjala et al., 2023). The pro-CAA groups started to belittle the activists who came to support the community by accusing them of coming just to have “biryani” - a delectable dish, made of rice and meat, famous in South Asia. One of our interviewees noted this fragmented nature of the overall movement on-site. She said that the protest sites were fractured because of what became visible to outsiders.

It can be argued that the activities and relationships facilitated by social media, both on and off-site, “generate both individual and collective action and conceive of the individual and collective as the result of a socio-technical genesis” (Renzi & Langlois, 2015, p. 205). Thus, the social media praxis cannot be considered extraneous and outside of the local community activism. It is integral to the internal offline community-based activism and the strategic and careful staging of citizenship, community care and loyalty. This is a form of *transindividuation*, which Alessandra Renzi and Ganaele Langlois (2015) describe as the emergence of individual supporters and large crowds, which can be attributed to various factors, including the “data analysis and complex algorithms that extract surplus value from social justice causes (while supporting them), viral information circulation, and the emotions circulating on social media networks converge to establish a new relationship between the one and the many” (p. 206).

Even efforts to erase and silence the events took place both offline and online. Eventually, the Covid-19 shutdown presented authorities with a chance to dismantle and clear the protest site. Ironically, though, while the movement's termination occurred through multiple means offline, the digital images gathered by different activists, whether disseminated on social media or not, became some of the concrete evidence that the movement had taken place.

4.3 Uncertain material and virtual archives

In March 2020, we observed some key actors on Twitter tagging the Delhi police Twitter handle regarding the violence that had occurred as soon as the Covid-19 shelter-at-home was announced in India. The authorities used this as an opportunity to shut down all protests with force. The protest graffiti on the walls of Shaheen Bagh was being erased as there were orders to whitewash the walls. A few of the protestors and onlookers continued to Tweet out these conditions to the world – enabling a virtual witnessing of what was occurring. Some of the Tweets were threaded together with the hashtag “InquilabJaariHai” (the revolution lives). Even though this was a fairly small data set we came to know of this hashtag because one of the transnational activists we were working with called our attention to it. When our attention was drawn towards this hashtag, of course we proceeded to collect data using the key word “InquilabJaariHai.” Some of the Twitter handles that had the highest visibility in the dataset scraped, have since then either deleted their accounts or had their accounts suspended. Figure 5 shows how we began to identify the most active accounts. The Twitter handle labels is hidden under the black ink since we do not wish to reveal the activist names whose accounts were subsequently deleted. Figure 6 shows how we narrowed down the search using features available on Netlytic platform, and finally in Figure 7 we see how central the Delhi Police's Twitter handle is to this dataset. However, what is key here is that this moment of violence against the protestors under the guise of Covid-19 lockdown would not have been recorded or known beyond the local community were it not for the Tweet storms that these users organized. The series of images from Figure 5 to Figure 7 thus are meant to illustrate how we combined knowledge gained from interviews to locate data points and get a further understanding of the importance of what might be seen as uncertain archives (Thylstrup et al, 2021).

As Nanna Bonde Thylstrup and co-authors point out, big data archives are uncertain because of the ways in which the private owners of the platforms work with various state actors to suppress and disappear evidence of protest movements. When we examined this cluster again two years after we collected the

data, we searched for the prominent actors and the hashtag. Some of the content we had in our offline data collection was missing in online publics. Yet in the case of Shaheen Bagh, we saw that material evidence (archives in the form of graffiti and protest art) was also erased. Our team and our collaborators therefore accidentally became archivists holding this evidence of the protest movement in offline digital repositories.



Figure 5. Dataset around InquilabJaariHai. Locating the most visible Twitter handles in the data set " InquilabJaariHai"



Figure 6. Sticky notes used to annotate Netlytic data. Sticky note feature on Netlytic used as place markers to take notes while examining the connections between the most visible nodes.

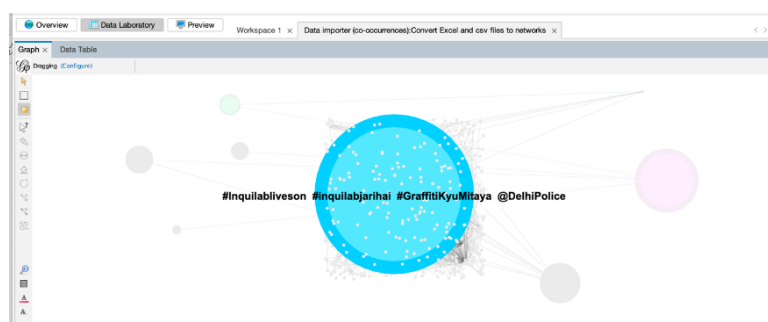


Figure 7. Dataset from Netlytic uploaded to Gephi for closer examination. The dataset was downloaded from Netlytic as Graphml file and uploaded to Gephi to highlight the central nodes. Different ways of looking at the data provided insight and help us locate significance as we did follow up interviews.

5. Conclusion

Our experience as feminist researchers operating in digital spaces has highlighted the importance of adopting a multi-methods approach that encompasses multiple sites. Specifically, in our case, we collected data from both social media platforms and the offline site, thus gathering insights from a diverse range of subjects. These included social media users with varying political affiliations and the protestors on the ground. We were engaged in an algorithmic ethnography that distinguishes itself from normative approaches to tracking and analyzing digital social movements. This nuanced engagement with data distinguishes our research from flattened big data approaches and inspires us to name our approach as a feminist intersectional small data analysis.

By taking this holistic approach, we gained a comprehensive understanding of the sudden emergence and recognition of the women of Shaheen Bagh as subaltern political agents/subjects. Drawing from Dhiraj Murthy's observation that Twitter can be thought of as a "field" in the Bourdieusian sense, we focused on immersing ourselves in the environment and "gaining experiential/cultural knowledge about [the] tweet corpus" (Murthy, 2017, p. 1). Only by entering this digital field were we able to fully consider the practices and social and material capital (access and literacy) needed for engaging with Twitter discussion spaces. Twitter engagement is not accidental, but intentional and an intervention into world views. While Twitter has both naïve and spontaneous everyday users, it also has strategic and planned actors such as influencers and activists who have developed strategies to create algorithmic visibility, and this shifting between on and offline visibility and strategic planning only becomes legible when researchers and collectors look within and beyond data corpuses to actors. Through our feminist intersectional small data analysis, we pushed back against both the reliance on big data methods and what Sophie Bishop refers to as the "fetishization" of "hard" in-person ethnography in comparison to digital ethnographies (Bishop 2018, 143). We do this by placing significant emphasis on our algorithmic intervention which involves a circular process of scraping, immersing, collecting, and then diving back into the data, allowing us to trace these connections outward to the people and collectives engaged in activist work and preservation.

Further, much of public Twitter engagement is planned through WhatsApp groups and other offline modes of group communication. During the anti-CAA/NRC protests, we witnessed some of this unfolding as several groups of supporters gathered in groups on various messaging platforms to discuss and plan responses to the situation and to offer help. Some were collecting visual evidence of police brutality on-site, and others were storing away archives of protest art and signage, fearing that they would be destroyed, confiscated, or whitewashed at any moment. Our findings highlight the significance of digital activist archiving. As Hui (2015) argues, the will to archive transforms both digital and on-site platforms into sites of power. The iterative process of response and collection carried out by supporters becomes identifiable and meaningful when we focus on relational, networked connections as manifested in the data.

By engaging in radical performances of care-oriented context creation and the reconstruction of memories, both online users and on-site actors further underscores the importance of a holistic approach to data collection that captures even the smallest cultural nuances and specificities, such as the hashtag *InquilabJaariHai*, that constitute the dense mass of big social data. We argue that to access multiple field sites and refrain from privileging or isolating one context requires addressing the imbrication of the online and offline in contexts of digital social movements. It is, therefore, imperative for feminist scholars to embrace an approach to algorithmic ethnography, that narrates how we break through the algorithmic conditions of digital platforms and network analysis tools and situated biases to unveil the nuances of digitally mediated social movements.

References

- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. John Wiley & Sons.
- Bhatia, K. V. (2022). Hindu nationalism online: Twitter as discourse and interface. *Religions*, 13(8), 739.
- Bishop, S. (2018). Fetishisation of the “offline” in feminist media research. *Feminist Media Studies*, 18(1), 143–147. <https://doi.org/10.1080/14680777.2018.1407120>
- Boellstorff, T. (2015). *Coming of age in Second Life: An anthropologist explores the virtually human*. United States: Princeton University Press.
- Brock, A., Jr. (2020). *Distributed Blackness*. NYU Press.
- Burgess, J., & Bruns, A. (2015). Easy Data, Hard Data: The Politics and Pragmatics of Twitter Research After the Computational Turn. In G. Langlois, J. Redden, & G. Elmer (Eds.), *Compromised data: From social media to big data* (pp. 93–111). Bloomsbury Publishing USA.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.
- Christin, A. (2020). Algorithmic ethnography, during and after COVID-19. *Communication and the Public*, 5(3-4), 108-111.
- Colman, F., Bühlmann, V., O’Donnell, A., & van der Tuin, I. (2018). Ethics of coding: A report on the algorithmic condition [EOC]. European Commission. <https://dspace.library.uu.nl/handle/1874/396432>
- D’Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press.
- Dey, A., (2020). Sites of exception: Gender violence, digital activism and Nirbhaya’s zone of anomie in India. *Violence Against Women* 26(11), 1423–44
- Edwards, E. L. (2023). *Digital Islamophobia: Tracking a Far-Right Crisis*. Walter de Gruyter GmbH & Co KG.
- Edwards, E., Ford, S., Gajjala, R., & Bhatia, K. V. (2024). Shaheen Bagh: Making sense of (re)emerging “Subaltern” feminist political subjectivities in hashtag publics through critical, feminist interventions. *New Media & Society*, 26(1), 473-494. <https://doi-org.ezproxy.bgsu.edu/10.1177/14614448211059121>
- Elish, M. C., & boyd, danah. (2018). Situating methods in the magic of Big Data and AI. *Communication Monographs*, 85(1), 57–80. <https://doi.org/10.1080/03637751.2017.1375130>
- Elmer, G., Langlois, G., & Redden, J. (2015). *Compromised data: From social media to big data*. Bloomsbury Publishing USA.
- Faniyi, O. (2023). *A herstory of #EndSars: Nuances of intersectionality in Nigeria’s movement against police brutality* [Master’s thesis, Bowling Green State University]. Ohio Link Electronic Theses and Dissertations Center.
- Florini, S. (2019). *Beyond hashtags: Racial politics and Black digital networks*. NYU Press.
- Gajjala R (2011) Snapshots from sari trails: Cyborgs old and new, *Social Identities*, 17(3), 393-408, DOI: 10.1080/13504630.2011.570977
- Gajjala R, Edwards EL, Rahut D, et al. Transnationalising Dadis as Feminist Political/Activist Subjects. *Feminist Encounters: A Journal of Critical Studies in Culture and Politics*. 2023;7(1), 08. <https://doi.org/10.20897/femenc/12886>,
- Gajjala, R., DeGalan, A., Rahut, D., Akbar, S. Z., & Jain, J. (2023b). Women activists imaged through social media publics: The “Fiesty Dadis of Shaheen Bagh” as political subjects. In M. Lunenburg & B. Rottger-Rossler (Eds.), *Affective formation of Publics: Places, networks, and media* (pp. 105–128). Routledge.
- Haraway, D. (1988). Situated knowledges: The science question in Feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575–599. <https://doi.org/10.2307/3178066>
- Howard, P.N., & Hussain, M.M. (2013). *Democracy’s fourth wave?: Digital media and the Arab spring*. Oxford University Press.
- Hui, Y. (2015). A contribution to the political economy of personal archives. In G. Langlois, J. Redden, & G. Elmer (Eds.), *Compromised data: From social media to big data* (pp. 226–246). Bloomsbury Publishing USA.
- Jackson, S.J., Bailey, M. & Welles, B. F. (2020). *#HashtagActivism: Networks of race and gender justice*. MIT Press.
- Kadiwal, L. (2021). Feminists against Fascism: The Indian female Muslim protest in India. *Education Science*, 11(12), Article 12. <https://doi.org/10.3390/educsci11120793>
- Kermani, H. (2023). #MahsaAmini: Iranian twitter activism in times of computational propaganda. *Social Movement Studies*, 0(0), 1-11. <https://doi.org/10.1080/14742837.2023.2180354>
- Latzko-Toth, G., Bonneau, C., & Millette, M. (2017). Small data, thick data: Thickening strategies for trace-based social media research. In L. Sloan & A. Quan-Haase (Eds.), *The SAGE handbook of social media research methods* (pp. 199–214). SAGE.
- Lukić, J., & Espinosa, A. S. (2011). Feminist perspectives on close reading. In R. Buikema, G. Griffin, & N. Lykke (Eds.), *Theories and methodologies in postgraduate feminist research* (pp. 105-118). Routledge.
- MacKinnon, R. (2012). *Consent of the networked: The worldwide struggle for internet freedom*. Basic Books.
- Mann, C., & Stewart, F. (2000). *Internet communication and qualitative research: A handbook for researching online*. SAGE.
- Marwick, A. E., & boyd, d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society*, 13(1), 114-133.
- Mishra, D., Akbar, S. Z., Arya, A., Dash, S., Grover, R., & Pal, J. (2021). Rihanna versus Bollywood: Twitter influencers and the Indian farmers’ protest (arXiv:2102.04031). arXiv. <https://doi.org/10.48550/arXiv.2102.04031>

- Murthy, D. (2017). The ontology of Tweets: Mixed-method approaches to the study of Twitter. In L. Sloan & A. Quan-Haase (Eds.), *The SAGE handbook of social media research methods* (pp. 559-572). SAGE.
- Mustafa, S. (ed.) (2020). *Shaheen Bagh and the idea of India: Writings on a movement for justice, liberty and equality*. Speaking Tiger Books.
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Pedwell, C. (2019). Digital tendencies: Intuition, algorithmic thought and new social movements. *Culture, Theory and Critique*, 60(2), 123–138. <https://doi.org/10.1080/14735784.2019.1579658>
- Pedwell, C. (2022). Speculative machines and us: More-than-human intuition and the algorithmic condition. *Cultural Studies*, 0(0), 1-31. <https://doi.org/10.1080/09502386.2022.2142805>
- Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2015). *Digital ethnography: Principles and practice*. SAGE.
- Pink, S., Sumartojo, S., Lupton, D., & Heyes LaBond, C. (2017). Empathetic technologies: digital materiality and video ethnography. *Visual Studies*, 32(4), 371-381. <https://doi.org/10.1080/1472586X.2017.1396192>
- Poirier, L. (2021). Reading datasets: Strategies for interpreting the politics of data signification. *Big Data & Society*, 8(2). <https://doi.org/10.1177/20539517211029322><https://doi.org/10.1177/20539517211029322>
- Renzi, A., & Langlois, G. (2015). Data activism. In G. Langlois, J. Redden, & G. Elmer (Eds.), *Compromised data: From social media to big data* (pp. 202–225). Bloomsbury Publishing USA.
- Rettberg, J. W. (2020). Situated data analysis: A new method for analysing encoded power relationships in social media platforms and apps. *Humanities and Social Sciences Communications*, 7(1), 1–13.
- Risam, R. (2018). *New digital worlds: Postcolonial digital humanities in theory, praxis, and pedagogy*. Northwestern University Press.
- Rogers, R. (2019). *Doing digital methods*. SAGE.
- Rotman, D., Vieweg, S., Yardi, S., Chi, E., Preece, J., Shneiderman, B., Pirolli, P., & Glaisyer, T. (2011). From slacktivism to activism: participatory culture in the age of social media. In CHI '11 extended abstracts on Human Factors in Computing Systems (pp. 819-822). <https://doi.org/10.1145/1979742.1979543>
- Rouvroy, A., & Stiegler, B. (2016). The digital regime of truth: From the algorithmic governmentality to a new rule of law. *La Deleuziana*, (3), 6–29.
- Salam, Z.U. (2020). *Shaheen Bagh: From a protest to a movement*. Bloomsbury Publishing.
- Seigworth, G. J., & Pedwell, C. (Eds.). (2023). *The affect theory reader 2: Worldings, tensions, futures*. Duke University Press.
- Shahin, S. (2016). A critical axiology for big data studies. *Palabra Clave*, 19(4), 972–996. <https://doi.org/10.5294/pacla.2016.19.4.2>
- Sinha, K., Jhalani, P., Khan, A. & Mukherjee, P.C. (2023). Influencers as institutions: Impact of digital politics in the Global South. *Global Policy*, 14(5), 912-924. <https://doi.org/10.1111/1758-5899.13188>
- Sobande, F., Fearfull, A., & Brownlie, D. (2020). Resisting media marginalisation: Black women’s digital content and collectivity. *Consumption Markets & Culture*, 23(5), 413-428. <https://doi.org/10.1080/10253866.2019.1571491>
- Srinivasan, R., (2013). Bridges between cultural and digital worlds in revolutionary Egypt. *The Information Society*, 29(1), 49-60. <https://doi.org/10.1080/01972243.2012.739594>
- Strom, D. (2012, August 3). Big data makes things better. *Dice Insights*. <https://www.dice.com/career-advice/big-data-makes-things-better>
- Tseng, Y. S. (2022). Algorithmic empowerment: A comparative ethnography of two open-source algorithmic platforms—Decide Madrid and vTaiwan. *Big Data & Society*, 9(2), 20539517221123505. <https://doi.org/10.1177/20539517221123505>
- Thylstrup, N. B., Agostinho, D., Ring, A., D’Ignazio, C., & Veel, K. (2021). *Uncertain archives: Critical keywords for big data*. MIT Press.