Journal Homepage: https://discoursereview.com

JOURNAL OF DISCOURSE REVIEW

Open Access Humanitarian Scholarship

July 2025
Volume-1, Issue-2

Research Article

Surveilled Selves: Gendered Datafication and the Discourse of Control in the Digital Public Sphere

Mahera Imam[1]* , N. Manimekalai[2] & S. Suba[3]

[1] *Department of Women's Studies, Khajamalai campus, Bharathidasan University, Tiruchirapalli, India*
[2] *Centre for Women's Development Studies, New Delhi,110001*
[3] *Department of Women's Studies, Khajamalai Campus, Bharathidasan University, Tiruchirappalli, India*

Abstract

In the era of pervasive digital technologies, surveillance has evolved into a distributed, algorithmically driven force. This paper argues that datafication, the conversion of human lives into quantifiable digital data, manifests in deeply gendered forms of surveillance and control within the digital public sphere. Drawing on Zuboff's (2019) "surveillance capitalism" and Benjamin's (2019) "New Jim Code," this study critically examines how algorithmic systems not only perpetuate but intensify existing patriarchal and intersectional oppressions. Through a feminist lens, it investigates how women's digital bodies are surveilled, commodified, and policed, restricting autonomy and reinforcing normative control, particularly amidst caste, religious, and class inequalities in India. Empirical evidence like UN Women (2020) and Freedom House (2024) reports highlight these vulnerabilities. Utilizing discourse analysis, the paper interrogates the power embedded in digital architectures, contending that the digital public sphere, far from democratizing, is stratified and exclusionary, echoing Fraser's (1990) critique. This study contributes to critical feminist digital studies, advocating for digital justice rooted in care and consent, and calls for gender-responsive digital governance.

Article info

Received 15 June 2025
Accepted 23 June 2025
Published 22 July 2025

Keywords: *Gendered Surveillance, Datafication, Digital Public Sphere, Algorithmic Bias, Feminist Digital Justice, Technology-Facilitated Violence.*

1. Introduction

The digital turn a term used to describe the sweeping transformation of society through the integration of digital technologies has dramatically altered how individuals interact, communicate, work, and are governed. At the centre of this shift lies the process of datafication, which Mayer-Schönberger and Cukier (2013) define as the conversion of qualitative aspects of life into quantitative, machine-readable data. This includes everything from GPS locations and social media activity to emotional expressions and biometric information. Datafication enables the collection, analysis, and prediction of human behaviour on an unprecedented scale, promising efficiency and optimization but also raising critical concerns around privacy, control, and power. Closely tied to datafication is the rise of digital surveillance, which refers to the continuous monitoring, tracking, and profiling of individuals through digital systems. As David Lyon (2018) argues, modern surveillance is not limited to state apparatuses but is embedded in everyday interactions mediated by corporations, algorithms, and digital platforms. This surveillance capitalism, a term popularized by Shoshana Zuboff (2019), operates by extracting behavioural surplus

**Corresponding author*, ✉ maheraimam8@gmail.com https://orcid.org/0009-0005-2191-4631

Cite this Article: Imam, M., Manimekalai, N., & Suba, S. (2025). Surveilled Selves: Gendered Datafication and the Discourse of Control in the Digital Public Sphere. *Journal of Discourse Review, 1*(2), 169-178.

data from users to generate predictive insights, monetize attention, and shape future actions. Importantly, this form of surveillance is often opaque, ambient, and normalized under the guise of convenience and personalization.

While the digital turn is frequently celebrated for its democratizing potential enabling new forms of participation, communication, and social mobilization it also produces novel hierarchies and exclusions. Scholars like Ruha Benjamin (2019) and Safiya Umoja Noble (2018) caution against techno-optimism, arguing that digital systems often reproduce and amplify existing inequalities rather than dismantle them. This is particularly true for marginalized groups, including women, Dalits, Adivasis, religious minorities, and LGBTQ+ individuals in the Global South, who experience datafication and surveillance in deeply gendered and intersectional ways. Surveillance is not a neutral act; it is socially patterned and politically consequential. As Simone Browne (2015) observes in her work on "racializing surveillance," the history of monitoring is entangled with systems of domination. Applying her insight to the Indian context reveals how caste, religion, and patriarchy intersect to structure who is visible, who is invisible, and who is rendered hyper-visible in digital spaces. For example, the cases of *Bulli Bai* and *Sulli Deals* where Muslim women's photos were circulated via online "auctions" demonstrate how digital infrastructures are weaponized to target marginalized identities. These incidents, documented by the *Internet Freedom Foundation* and *Human Rights Watch*, explore the limitations of legal frameworks and platform governance in addressing technology-facilitated gender-based violence. In this context, the present paper investigates how datafication and digital surveillance shape gendered experiences in the digital public sphere. It explores how technologies, far from being neutral, are embedded with patriarchal logics and discursive practices that regulate, commodify, and control women's bodies and behaviours. By focusing on India a nation marked by rapid digitization, stark socio-economic divides, and vibrant feminist movements this study seeks to unpack the paradox of digital empowerment and digital control. The aim is to understand how surveillance functions not merely as a technical process but as a social and political tool of discipline, often disguised in the language of safety, security, and efficiency. The following sections outline the theoretical framework (Section 2), the methodology (Section 3), and empirical insights into gendered datafication (Section 4). Subsequent sections explore the discursive structures of digital violence (Section 5), feminist resistance and counterpublics (Section 6), and conclude with reflections on digital justice and policy directions (Section 7).

2. Conceptual and Theoretical Framework

This paper draws on an interdisciplinary conceptual framework that integrates theories from feminist studies, surveillance studies, and critical digital media scholarship. Each theoretical lens offers tools to analyse how surveillance and datafication are not merely technological phenomena, but also deeply social and political processes with distinct gendered and intersectional implications.

2.1. Surveillance Capitalism – Shoshana Zuboff:

Zuboff's (2019) theory of surveillance capitalism articulates how digital platforms extract, commodify, and monetize behavioural data to predict and manipulate human behaviour. This process, driven by the economic imperative of profit, transforms users into raw material. While Zuboff's work centres primarily on economic exploitation, this paper extends her framework to understand how surveillance capitalism intersects with patriarchal control, especially in the context of gendered monitoring, digital profiling, and the erosion of bodily autonomy online. In the Indian context, this manifests in the commodification of women's digital presence and the monetization of visibility, often without consent or recourse.

2.2. Intersectionality – Kimberlé Crenshaw:

Crenshaw's (1991) concept of intersectionality enables a layered understanding of how digital surveillance disproportionately targets individuals based on the intersecting axes of gender, caste, religion, class, and sexuality. In digital environments, these identities are not erased but encoded resulting in intensified forms of exclusion, harassment, and control. For instance, Muslim women in India face targeted online attacks that combine gendered abuse with communal hate speech, as seen in the Sulli Deals and Bulli Bai cases. Intersectionality thus provides a vital framework for identifying and analysing the compounded vulnerabilities produced through algorithmic systems.

2.3. Feminist Surveillance Studies – Rachel Dubrofsky, Simone Browne:

Feminist scholars like Dubrofsky (2011) and Browne (2015) challenge the assumption that surveillance is neutral or objective. Browne's notion of "racializing surveillance" is particularly relevant for understanding how digital technologies extend colonial and patriarchal systems of control into virtual spaces. In India, where caste-based surveillance persists in offline structures, these dynamics are replicated in online settings through content moderation, facial recognition biases, and the marginalization of dissenting voices. Dubrofsky emphasizes how surveillance is often framed as care or protection, especially in relation to women, masking its disciplinary intent.

2.4. Algorithmic Oppression – Safiya Umoja Noble:

Noble (2018) introduces the concept of "algorithmic oppression" to reveal how supposedly neutral search engines and recommendation systems reproduce societal biases. Her work is instrumental in examining how digital infrastructures perpetuate misogyny, racism, and heteronormativity under the guise of objectivity. In India, casteist and sexist content often trends due to algorithmic amplification, and platform designs rarely reflect the needs of marginalized users. This framework explores the role of design choices, data curation, and moderation policies in shaping digital realities that are neither neutral nor equitable.

2.5. Public/Counterpublic Sphere – Nancy Fraser:

Fraser's (1990) critique of Habermas's idealized public sphere introduces the concept of "subaltern counter publics" alternative discursive spaces where marginalized groups can articulate their concerns and contest dominant narratives. In the digital realm, feminist activists, journalists, and collectives like Khabar Lahariya or Point of View have created such counter publics to challenge gendered oppression and reclaim voice. However, these spaces are not immune to surveillance and backlash, revealing the tension between visibility and safety. Fraser's theory helps contextualize digital participation as a terrain of both empowerment and contestation.

Collectively, these theories guide this paper's analysis of how digital surveillance functions through gendered, racialized, and class-based logics. They offer critical insight into the structural forces behind online violence, algorithmic exclusion, and the political economy of data enabling a detailed critique of the digital public sphere and a call for feminist digital justice.

3. Methodology

This study employs a qualitative and critical methodological approach, firmly rooted in feminist digital epistemologies that prioritize lived experiences and critically interrogate the structural inequalities embedded within technological systems. Rather than conceptualizing digital infrastructures as neutral or value-free, this approach understands them as complex socio-technical systems inherently shaped by intersecting power relations across gender, caste, class, and religion. This perspective acknowledges that technology is not merely a tool but a site where social inequalities are reproduced, amplified, and contested.

Critical Discourse Analysis (CDA) serves as the primary analytical framework for this investigation. Drawing on the foundational work of scholars such as Norman Fairclough (1992) and Teun A. van Dijk (2001), CDA is utilized to systematically examine how language, images, and narratives circulating on digital platforms construct, legitimize, and circulate power. This methodology enables a detailed unpacking of how specific discourses contribute to the naturalization of inequality, shape public perception, and influence the treatment of various social groups online. By analyzing the discursive strategies employed by both dominant and marginalized actors, CDA helps reveal the ideological underpinnings of digital interactions and their socio-political implications.

The empirical data sources for this study are diverse and multifaceted, encompassing a range of digital artifacts and public records. These include:

1. **Social Media Content:** A purposive sample of posts, comments, and interactions from platforms like Twitter, Facebook, and Instagram, focusing on discussions related to gender, surveillance, and online harm in India.
2. **Media Reports:** News articles and investigative journalism pieces from reputable national and international media outlets covering incidents of digital violence and surveillance.
3. **Legal Documents:** Key policy frameworks, notably the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, are analyzed for their discursive construction of harm, responsibility, and rights.
4. **Platform Interfaces:** Examination of the design features, moderation policies, and user experience flows of selected digital platforms to understand their inherent biases and power dynamics.

High-profile cases, such as the egregious instances of the *Sulli Deals* and *Bulli Bai* apps (which targeted Muslim women with online 'auctions'), and the increasing proliferation of deepfakes, are subjected to detailed critical discourse analysis. These cases are particularly salient as they vividly demonstrate the complex intersections of gendered surveillance, algorithmic bias, and communal hate in the Indian context. This multi-layered methodological approach allows the study to explore not only the explicit acts of digital violence but also their symbolic dimensions, ultimately demonstrating how a robust feminist discourse analysis can effectively expose, deconstruct, and resist the logics of exclusion and control that fundamentally structure the contemporary digital public sphere.

4. Gendered Datafication: Mechanisms and Manifestations

The pervasive process of datafication, frequently presented as a neutral and inevitable technological advancement, is, in reality, deeply imbued with gendered and intersectional logics. This section argues that datafication operates as a powerful mechanism that selectively renders certain bodies hyper-visible for intensive data extraction and control, while simultaneously rendering others invisible, excluded, or subject to systematic marginalization. Women's digital presence, in particular, is increasingly subjected to a multi-layered regime of surveillance, commodification, and manipulation through sophisticated technical mechanisms. These processes reduce complex human identities to discrete data points, thereby stripping away crucial context and individual agency. As Haggerty and Ericson (2000) powerfully articulate through their concept of the "surveillant assemblage," individuals are disaggregated and reconfigured into "data doubles" – fragmented digital identities that circulate autonomously, often beyond the control or even awareness of the original subject. These data doubles become the primary targets for algorithmic analysis, categorization, and predictive modeling, fundamentally altering how individuals are perceived and acted upon in the digital realm.

Within digital spaces, profiling and predictive algorithms play an increasingly critical role in shaping how users are represented, categorized, and ultimately treated. These algorithmic systems, far from being objective, are almost invariably trained on vast datasets that reflect and often amplify existing societal biases. Consequently, they reinforce dominant social norms by associating specific gendered behaviors, expressions, or even appearances with predetermined categories such as "risk," "deviance," or "market value." For instance, women who actively engage in political discourse, advocate for feminist ideals, or challenge patriarchal norms online are disproportionately flagged by automated systems or reported by malicious actors as potential "violators of community standards." Conversely, male aggression, misogynistic content, or even explicit threats often receive less scrutiny or are moderated inconsistently, creating a glaring double standard that disproportionately penalizes women's speech and activism. This algorithmic bias is not merely theoretical; a 2021 report by the Algorithmic Justice League, for example, provided compelling evidence that facial recognition systems consistently misclassified women and darker-skinned individuals, unequivocally demonstrating a structural bias embedded within the core technology itself. Beyond facial recognition, similar biases manifest in credit scoring, hiring algorithms, and even targeted advertising, illustrating how datafication can automate and scale discrimination. Furthermore, major platforms like Facebook and Twitter have faced widespread and sustained criticism for their opaque and inconsistent content moderation practices. These practices frequently result in the silencing of victims of online abuse, particularly women and marginalized groups, while simultaneously allowing perpetrators to continue operating with relative impunity, undermining trust and safety.

The datafication of women's online behavior extends beyond mere content to encompass their emotional labor, aesthetic expressions, and intricate social networks. All of these facets of digital life are meticulously captured, exhaustively analyzed, and ultimately monetized by platform operators. Social media platforms, for example, extract immense value from women's expressions of care, their creative endeavors, and their efforts in community building, transforming these intangible forms of engagement into quantifiable metrics that are highly profitable for advertisers and platform owners. As Sarah Banet-Weiser (2018) astutely observes, even feminist expression online, intended to challenge power structures, can be subtly co-opted and commodified. This process reinforces neoliberal ideals of individual empowerment – where self-branding and personal visibility become paramount – while simultaneously diverting attention from, and failing to address, systemic inequalities. The very act of "performing" feminism online can become a source of data extraction, where engagement metrics overshadow genuine political impact.

The surveillance of marginalized identities is even more acute, particularly at the critical intersections of caste, religion, and sexuality. In India, Dalit activists, queer individuals, and Muslim women face heightened scrutiny, not only from state apparatuses seeking to monitor dissent but also from hostile digital publics engaging in coordinated harassment. According to a grim report by Amnesty International India (2020), Muslim women activists, in particular, receive disproportionate levels of online abuse, a toxic cocktail combining misogyny with virulent Islamophobia. This intersectional targeting illustrates how existing societal prejudices are amplified and weaponized through digital platforms. Queer individuals, meanwhile, frequently experience algorithmic erasure through heteronormative platform design, which fails to recognize or support diverse gender identities and sexual orientations, making it difficult for them to find community or express themselves authentically. Similarly, Dalit voices, when they speak out against casteist discrimination or assert their rights, are often deplatformed under the spurious pretext of "hate speech violations," even when their content is a legitimate response to systemic casteist abuse. This selective application of content policies serves to silence the oppressed while protecting the perpetrators of hate.

Violations of consent and privacy are rampant and endemic within the digital landscape, disproportionately impacting women. The non-consensual circulation of intimate images, large-scale data leaks exposing personal information, and doxxing (the public release of private or identifying information about an individual) are pervasive forms of technological violence. These acts not only cause immense psychological distress but also pose significant physical safety risks. The Cyber Peace Foundation (2022) reports a significant rise in India of cases involving morphed images, deepfakes, and unauthorized data sharing, often with minimal legal recourse available to victims. These violations are not isolated incidents attributable to rogue actors; they are systemic, facilitated by a confluence of factors: weak privacy protections embedded in legal frameworks, inadequate enforcement mechanisms by both state and platform actors, and a profound lack of gender sensitivity in the very design and regulation of technology. The default settings of many platforms, for instance, often prioritize sharing over privacy, making it easier for personal data to be exposed without explicit, informed consent.

5. Digital Violence and Discursive Control

Technology-facilitated gender-based violence (TFGBV) is not an unintended side-effect of the digital age but a deliberate and systemic form of social control. It operates by policing boundaries of acceptable behaviour, punishing dissent, and reinforcing patriarchal norms. Digital violence ranging from trolling and doxxing to cyberstalking, deepfake creation, and online shaming is embedded in broader cultural scripts that devalue women’s autonomy, particularly in public or political roles. This section draws on illustrative case studies to examine the discursive functions of digital violence. For example, journalist Rana Ayyub, a vocal critic of communalism and patriarchy, has been subjected to relentless abuse, including the circulation of morphed pornographic content designed to humiliate and discredit her. Similarly, women targeted in the *Sulli Deals* and *Bulli Bai* cases were "auctioned" on online platforms an act that symbolically and digitally marked them as property. These attacks, widely covered by media and rights groups like *Human Rights Watch* and the *Internet Freedom Foundation*, are examples of how gendered and communal forms of hatred converge online to silence vulnerable voices. The discourse surrounding such attacks often blames the victim, subtly or overtly. When women report harassment, questions around their online behaviour, dress, or presence are raised shifting the burden from the perpetrator to the survivor. This victim-blaming discourse serves to normalize violence, delegitimize resistance, and justify surveillance. Language plays a crucial role in this process. The euphemization of abuse as "trolling," the infantilization of perpetrators as "boys will be boys," and the framing of dissent as "anti-national" are discursive strategies that reinforce gender hierarchies and protect perpetrators.

Further, digital platforms themselves reproduce these silences through inconsistent enforcement of community guidelines. Survivors often report posts only to be told that the content “does not violate” platform policies. This institutional neglect creates what Sara Ahmed (2017) terms a “non-performative” response where mechanisms exist in name but not in action. The resulting silence discourages women from speaking out, further entrenching their marginalization. Drawing from qualitative narratives collected by NGOs and research organizations, it becomes evident that TFGBV is experienced not merely as individual trauma but as collective disciplining. A Delhi-based activist, interviewed by the Internet Democracy Project (2022), described withdrawing from political commentary after repeated threats of rape and acid attack. "It’s not just me it’s a message to every girl who thinks she can speak her mind online," she said. Through these examples, this section demonstrates how digital violence is a form of discourse, one that enforces silence, shapes norms, and sustains structures of domination. Far from being isolated incidents, TFGBV represents a patterned, normalized, and tolerated phenomenon that is deeply woven into the fabric of the digital public sphere. Recognizing this violence as systemic and discursive is essential to devising interventions that go beyond legal remedies and address the broader cultural and technological frameworks that sustain it.

6. The Politics of Visibility and Invisibility

In the contemporary digital landscape, the concept of visibility transcends mere presence; it is inherently political, serving as a critical determinant of who gains recognition, legitimacy, and safety online. Far from being a neutral or universally accessible space, digital platforms are profoundly shaped by intricate social structures that dictate the conditions under which individuals and groups can be seen, heard, or acknowledged. This section argues that despite the prevalent discourse of digital democratization and participation, online environments are, in fact, hierarchically structured through complex mechanisms of algorithmic amplification, platform design, and entrenched cultural norms. These mechanisms actively elevate certain voices and narratives while systematically marginalizing others, thereby reproducing and often intensifying existing power imbalances.

Drawing critically from Nancy Fraser's (1990) foundational work on the public sphere, particularly her concept of "subaltern counterpublics," this analysis extends her critique into the algorithmic age. Fraser posited that in stratified societies, marginalized groups often form parallel discursive arenas where they can formulate oppositional interpretations of their identities, interests, and needs. These counterpublics serve as vital spaces for articulating alternative forms of political participation and challenging dominant ideologies. In the digital realm, this theoretical framework finds powerful resonance. Marginalized communities, including women, Dalits, Muslims, queer individuals, and Adivasis in India, have adeptly leveraged digital tools to construct their own alternative discursive spaces. These digital counterpublics, exemplified by movements such as #MeTooIndia, the impactful work of Dalit Camera, and the community-driven journalism of Khabar Lahariya, function as crucial platforms for collective storytelling, organized resistance, and fostering solidarity. They operate in a dynamic tension, existing both within and against the hegemonic mainstream digital public sphere, strategically seeking to reshape public discourse from the margins by introducing counter-narratives and challenging established power structures.

However, the acquisition of visibility in digital spaces is a profoundly double-edged sword. While it undeniably offers avenues for empowerment through amplified voices, community-building, and the mobilization of collective action, it simultaneously exposes individuals to significantly heightened risks. The very act of becoming visible can render users vulnerable to intensified surveillance, targeted harassment, and various forms of digital violence, including doxxing and deplatforming. As sociologist Sarah Banet-Weiser (2018) incisively warns, digital visibility for feminist and dissenting voices often comes at a steep cost. She argues that the increased exposure, while potentially empowering, paradoxically makes marginalized users more susceptible to being attacked, their personal information exposed (doxxed), or their platforms removed entirely (deplatformed). This precarious balance between empowerment and vulnerability is particularly evident in the Indian context. Numerous public figures and collectives have faced severe online backlash simply for challenging deeply entrenched patriarchal and communal norms. Prominent activists like Trupti Desai, who has consistently campaigned for women's entry into religious sites, and journalists like Arfa Khanum Sherwani, known for her critical commentary on socio-political issues, have been subjected to relentless online vitriol, including death threats, character assassination, and organized trolling campaigns. Similarly, collectives such as Pinjra Tod, advocating for women's hostel freedoms, have endured coordinated online attacks aimed at silencing their dissent.

Crucially, this uneven distribution of visibility is not an accidental byproduct of digital interaction; it is deeply systemic, embedded within the very architecture of digital platforms. Algorithms, designed to maximize engagement and profit, frequently privilege dominant identities and suppress dissenting voices. This manifests in various forms, such as the frequent "shadow-banning" of feminist hashtags, which reduces their reach and discoverability, while hate speech directed against marginalized groups often remains inadequately moderated, or is even amplified due to its high engagement metrics. The underlying logic of platform design, driven by engagement metrics, often values outrage, controversy, and spectacle over detailed discussion, equitable representation, and the promotion of diverse perspectives. This creates a feedback loop where content that generates strong, often negative, reactions is prioritized, further marginalizing voices that advocate for social justice and equity. This algorithmic bias perpetuates a form of "algorithmic injustice," where the digital infrastructure itself contributes to the systemic neglect and discursive marginalization of certain groups.

Moreover, even within the ostensibly liberatory spaces of feminist movements and digital counterpublics, internal hierarchies and power dynamics can persist. Critics have rightly pointed out that early iterations of #MeTooIndia, while undeniably powerful in bringing attention to sexual harassment, often centered the narratives and experiences of urban, upper-caste women. This inadvertently left the crucial experiences of Dalit, queer, and working-class women underrepresented, highlighting the persistent challenge of intersectionality within social movements. As scholars like Nishant Shah and Shweta Radhakrishnan argue, for a truly inclusive and transformative digital counterpublic to emerge, there must be an active recognition of the plurality of feminist struggles. This necessitates a conscious effort to dismantle internal hierarchies and ensure that diverse voices, particularly those from the most marginalized intersections of identity, are not only heard but genuinely prioritized and integrated into the broader discourse.

In this complex terrain, invisibility takes on dual meanings. On one hand, it can signify erasure – a direct consequence of systemic neglect, biased algorithmic design, and deliberate discursive marginalization. When voices are systematically silenced or rendered unfindable, it amounts to a form of digital disappearance, denying individuals and communities the right to participate in public discourse and shape their own narratives. On the other hand, strategic invisibility is sometimes purposefully employed by marginalized users as a crucial tactic of resistance and self-protection against surveillance and harm. This might involve using pseudonyms, engaging primarily in private or encrypted online groups, or carefully curating their digital footprint to limit exposure to hostile actors. This tension between forced erasure and deliberate, strategic withdrawal explores the complex and often perilous landscape of digital participation, where the pursuit of empowerment and the experience of vulnerability are inextricably intertwined.

Therefore, the politics of visibility in the digital sphere extends far beyond the simplistic notion of merely "being seen." It is fundamentally about the power to define the terms of one's own visibility – to be seen with dignity, in spaces that respect diversity, and within environments that prioritize safety. Reimagining digital publics through the transformative lens of feminist justice necessitates not only increasing access to technology but, more profoundly, transforming the underlying structural biases that dictate who is rendered visible, how they are portrayed, and to what ultimate end. This requires a critical examination of platform governance, algorithmic accountability, and the cultivation of digital spaces that genuinely foster equitable participation and protect the most vulnerable.

7. Digital Justice and Feminist Alternatives

The efficacy of current policy frameworks addressing online harm and digital governance in India is demonstrably limited, fragmented, and predominantly reactive, failing to adequately contend with the complex challenges of the contemporary digital landscape. The Information Technology Act, 2000 (IT Act), which serves as the foundational legislation governing digital space in India, is increasingly anachronistic. Enacted at a nascent stage of internet evolution, its scope is inherently constrained and ill-equipped to address the sophisticated and pervasive issues of algorithmic surveillance, systemic platform bias, and multifaceted forms of digital violence that characterize the current digital age. Its provisions, largely focused on cybercrimes and electronic transactions, offer insufficient legal and regulatory tools to grapple with the detailed implications of datafication, the opaque nature of algorithmic decision-making, or the rapid proliferation of technology-facilitated gender-based violence. The more recent IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, while ostensibly designed to regulate digital platforms and address online harms, have paradoxically exacerbated concerns rather than alleviating them. These rules have been widely criticized for their potential for governmental overreach, a pervasive lack of transparency in their implementation, and the inclusion of vague definitions of "harm" and "objectionable content." Such ambiguity creates a chilling effect on legitimate online expression, often leading to the silencing of dissent and critical voices rather than genuinely protecting users from harm. Critics argue that the rules grant excessive power to the state and platforms to censor content, without robust independent oversight or clear due process. This regulatory approach, rather than fostering a safe and equitable digital environment, risks transforming platforms into instruments of state control and corporate censorship, further marginalizing vulnerable communities whose voices are often the first to be suppressed.

Compounding these legislative shortcomings is a profound lack of gender sensitivity within existing policies, coupled with inadequate mechanisms for redress, particularly for marginalized communities. Despite a growing body of evidence detailing the prevalence and devastating impact of technology-facilitated gender-based violence (TFGBV) – ranging from online harassment and cyberstalking to doxxing and non-consensual image sharing – Indian digital policies offer minimal specific provisions or tailored support for victims. This oversight is particularly egregious given the intersectional vulnerabilities faced by women from marginalized caste, religious, and class backgrounds, who often experience compounded forms of online abuse and have even fewer avenues for seeking justice or protection. The current frameworks largely fail to acknowledge the gendered nature of online harm, treating it as a generic cybercrime rather than a manifestation of systemic gender inequality amplified by digital technologies.

In direct response to these critical limitations, scholars, activists, and civil society organizations have articulated a compelling vision for "feminist digital justice." This emergent framework fundamentally re-centers the design, governance, and regulation of digital technologies around core principles of care, consent, and contextuality. It moves beyond a narrow focus on individual harm or technical fixes, advocating for a systemic transformation that prioritizes human dignity, equity, and the empowerment of marginalized communities.

This vision is deeply informed by critical theoretical perspectives. Helen Nissenbaum's (2010) seminal theory of "contextual integrity" offers a powerful critique of traditional, binary notions of privacy. Nissenbaum argues that privacy is not simply about control over information or its absence, but rather about the appropriate flow of information within specific social and relational contexts. She contends that data collection and dissemination should adhere to the norms governing those contexts, respecting the expectations of individuals regarding how their information is shared and used. From a feminist digital justice perspective, current policy failures often violate contextual integrity by allowing for the indiscriminate collection and use of personal data, particularly from women and marginalized groups, without their informed consent or in ways that breach their contextual expectations. This leads to surveillance, exploitation, and the weaponization of personal information, undermining autonomy and safety.

Virginia Eubanks's (2018) groundbreaking work on "automated inequality" further illuminates how data-driven systems, far from being neutral, often entrench and exacerbate existing social inequalities. Through meticulous empirical research, Eubanks demonstrates how these systems disproportionately punish poverty and marginalization, automating discrimination and reinforcing cycles of disadvantage. Her work makes a compelling case for moving beyond techno-solutionism – the belief that technology alone can solve social problems – towards a demand for structural accountability. In the Indian context, where caste, class, and religious disparities are deeply entrenched, Eubanks's insights are particularly pertinent. Algorithmic systems, if not designed with a critical awareness of these power dynamics, can easily replicate and intensify existing biases, leading to automated forms of discrimination in areas ranging from access to public services to online policing.

Building on these critiques, Ruha Benjamin (2019) passionately advocates for a radical shift from merely "inclusion" into harmful systems to a fundamental "reimagining" of those systems through what she terms "abolitionist tools for the new Jim Code." Benjamin's "New Jim Code" refers to the ways in which seemingly neutral technological systems can encode and perpetuate racial and social hierarchies, mirroring the discriminatory practices of the Jim Crow era. Her call for "abolitionist tools" implies a transformative approach that seeks to dismantle oppressive technological structures rather than simply integrating marginalized groups into them. For feminist digital justice, this translates into actively challenging the extractive logic of surveillance capitalism, resisting the commodification of personal data, and reimagining digital participation as a site for collective liberation and dignity, rather than just a space for consumption or surveillance.

These theoretical principles find concrete expression in various global and regional initiatives. The "Feminist Internet Principles," for instance, articulate a comprehensive vision for a feminist internet that emphasizes universal access, agency over one's digital life, freedom of expression, and the centrality of consent in all digital interactions. Crucially, they also highlight the collective responsibility for challenging discrimination and promoting equity online, moving beyond individualistic notions of digital rights.

In India, a vibrant ecosystem of organizations has been at the forefront of translating these principles into practice. Organizations such as Point of View, Digital Empowerment Foundation, Internet Democracy Project, and IT for Change have consistently advocated for gender-sensitive digital policy reforms, promoted critical digital literacy among diverse communities, and spearheaded community-led technological interventions. Their work effectively bridges the gap between abstract theory and practical implementation, offering tangible models for rights-based, feminist technology use and governance. These organizations engage in policy advocacy, legal aid, research, and grassroots capacity building, empowering individuals to navigate the digital world more safely and assert their digital rights.

Furthermore, initiatives like DigitalHimalaya, the aforementioned Khabar Lahariya, and Feminism In India demonstrate the immense potential of alternative digital platforms to center marginalized voices and reclaim digital narratives. DigitalHimalaya, for example, focuses on digital inclusion in remote regions, while Khabar Lahariya provides a platform for rural women journalists to report on local issues, challenging mainstream media narratives. Feminism In India serves as a crucial online space for feminist discourse and activism, amplifying diverse perspectives. These grassroots initiatives not only provide much-needed representation for communities historically excluded from mainstream media and digital spaces but also offer compelling models of ethical platform design and moderation. These platforms are often rooted in principles of care, collective agency, and participatory governance, demonstrating how digital spaces can be intentionally built to foster safety, respect, and genuine empowerment, rather than merely maximizing engagement or profit.

A feminist approach to digital justice is not merely about providing protection from online harm; it is fundamentally about transforming the very structures that produce vulnerability and inequality in the digital realm. It demands a radical shift towards designing digital technologies with inherent care, governing digital spaces with robust accountability mechanisms, and building digital communities in ways that are deeply collaborative and inclusive. This vision calls for a concerted effort to challenge the pervasive logic of surveillance capitalism, resist the extractive practices of data commodification, and ultimately reimagine digital participation as a site of profound empowerment, collective dignity, and social transformation.

8. Conclusion

This paper has examined how digital surveillance and datafication are not neutral processes but are shaped by deeply embedded gendered, casteist, and communal structures. Through the lenses of surveillance capitalism, intersectionality, algorithmic oppression, and subaltern counter publics, it has explored how digital technologies not only reflect but actively reproduce inequalities often under the guise of empowerment or security. Technology-facilitated gender-based violence, unequal visibility, biased algorithmic infrastructures, and weak policy frameworks converge to make the digital public sphere an exclusionary and precarious space for many. The normalization of digital violence, the commodification of women's presence, and the strategic invisibility of dissenting voices highlight the urgent need to rethink the foundations of digital governance. As this paper argues, surveillance must be recognized as a gendered and political phenomenon. It is not enough to improve platform moderation or update legal frameworks; we must fundamentally reimagine the values and structures that shape our digital futures. A feminist approach to digital justice emphasizes care over control, consent over coercion, and community over commodification. To move toward truly inclusive digital participation, governments, platforms, and civil society must collaborate to ensure accountability, centre marginalized voices, and design technologies that affirm dignity and safety. Only through such transformative governance can the promise of the digital age be realized not just for a few, but for all.

Acknowledgement

This publication is an outcome of ICSSR-sponsored doctoral research. We are sincerely grateful to the Indian Council of Social Science Research (ICSSR), Ministry of Education, Government of India, New Delhi, for their invaluable financial support, which made this work possible. However, we bear sole responsibility for the information presented, the views expressed, and the findings of this study.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Ahmed, S. (2017). *Living a feminist life*. Duke University Press.

Amnesty International. (2020). *Troll Patrol India: Exposing online abuse faced by women politicians in India.*

Banet-Weiser, S. (2018). *Empowered: Popular feminism and popular misogyny*. Duke University Press.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity Press.

Browne, K. (2015). Negotiations and fieldworkings: Friendship and feminist research. *ACME: An International Journal for Critical Geographies, 2*(2), 132–146. https://doi.org/10.14288/acme.v2i2.690

Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review, 43*(6), 1241–1299. https://doi.org/10.2307/1229039

Dubrofsky, R. E. (2011). *The surveillance of women on reality television: Watching The Bachelor and The Bachelorette*. Lexington Books.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. Longman.

Fraser, N. (1990). Rethinking the public sphere: A contribution to the critique of actually existing democracy. *Social Text, 25/26*, 56–80. https://doi.org/10.2307/466240

Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *The British Journal of Sociology, 51*(4), 605–622. https://doi.org/10.1080/00071310020015280

Internet Democracy Project. (2019). *Online gender-based violence: A submission to the UN Special Rapporteur*. https://internetdemocracy.in/reports/

Internet Freedom Foundation. (2022). *Gendered disinformation: Challenges and policy responses*. https://internetfreedom.in

Lyon, D. (2018). *The culture of surveillance: Watching as a way of life*. Polity Press.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Eamon Dolan/Houghton Mifflin Harcourt.

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

Point of View. (n.d.). *Feminist principles of the internet*. https://www.apc.org/en/pubs/feminist-principles-internet

Shah, N., & Radhakrishnan, S. (2021). Digital feminist activism in India: Navigating visibility and erasure. *Feminist Media Studies, 21*(3), 476–492. https://doi.org/10.1080/14680777.2021.1875012

van Dijk, T. A. (2001). Critical discourse analysis. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 352–371). Blackwell.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.