

Platform Governance and Political Closure: Shadow Banning, Algorithmic Suppression, and Free Speech Violations in India, UK, and USA

Shilpa Agarwal¹, Vikram Kaushik²

¹ Research Scholar, Department of Mass Communication, Guru Jambheshwar University of Science & Technology, Hisar, Haryana, India.

² Professor, Department of Mass Communication, Guru Jambheshwar University of Science & Technology, Hisar, Haryana, India.

Corresponding Author:

Vikram Kaushik^{2*}

Abstract: Social media platforms have changed how political discourse takes place in a civilized society. What started off as neutral and democratic spaces for political debates are now battlegrounds of information warfare. Rules of content moderation and algorithmic ranking now govern these platforms, determining which voices deserve amplification and which face silencing. The state, under the guise of regulating the platforms, decides what content gets removed and what accounts get suspended, resulting in censorship. This is how the right to free speech in a democracy takes a hit and technology becomes a vice. This paper examines 49 documented cases of content removal, algorithmic suppression, account suspensions, and internet blackouts in India, the UK, and the USA between 2014 and 2024. It places these incidents within three unique but converging regulatory environments across the three democracies. By utilizing comparative policy analysis and critical discourse analysis, the study demonstrates how in all three democracies, government practices target opposition actors, journalists, activists, and marginalized communities the hardest while simultaneously protecting the establishment and corporate speech. India, where internet shutdowns rest on statutory authority as per the IT Rules 2021, saw 84 internet disruptions and 134 free speech violations in just 2024 that were state-ordered. In the USA, while free speech is protected by the First Amendment and the press is free from state control, the state uses indirect tactics such as pressurization, platform self-regulation, account suspension, and shadow banning in collusion with the tech giants and media elites to silence dissent. The UK's Online Safety Act does more to provide Ofcom wide discretion over "harmful" and "illegal" content than it does for online safety or free speech itself. The paper argues that the platforms should be viewed as a part of a technological intelligentsia consisting of a network of state agencies, corporations, media houses, and technical experts that collectively decide which narratives become visible, credible, and acceptable depending on the political aspirations and outcomes that suit the acting powers. It also suggests that the current regulatory framework is inefficient and often extends platform authority and surveillance capacities, contributing further to democratic erosion...

Keywords: platform governance, free speech violations, algorithmic suppression, technological intelligentsia, internet shutdowns

INTRODUCTION

On 19 May 2024, the Indian government ordered internet shutdowns in sixteen states during farmer protests and episodes of communal tension (Access Now, 2024; Singh, 2024). In the months before, officials had already blocked 177 social media accounts, with no clear legal explanations (Newslandry, 2024). Around the same time the Election Commission of India instructed platforms to remove deepfakes and election-related "misinformation" (ECI, 2024). 2024 ended with India logging 84 internet shutdowns, the highest among democracies, and 134 recorded free speech violations (Free.

Speech Collective, 2024). The takedown of the “Bolta Hindustan” YouTube channel via a government order (Medianama, 2024) and blocking of Hindutva Watch and India Hate Lab, two independent initiatives documenting communal violence (Article 19, 2024), are prominent examples of state-ordered censorship. Several journalists left India after sustained pressure linked to critical reporting (The Guardian, 2024; Peoples Dispatch, 2024).

In the USA, platform governance coincided with electoral cycles and political crises. The 2020 and 2024 elections (Council on Foreign Relations, 2024; ACLED, 2024), the COVID-19 pandemic (Benkler et al., 2022), protests against police violence (Carnegie Endowment for International Peace, 2026), and recurring immigration debates (Migration Policy Institute, 2024; Scientific American, 2024) all became sites of digital conflict (Trithara, 2024). Platforms suspended accounts belonging to activists, environmental groups, and public critics of government policy (Human Rights Watch, 2021). Shadow banning content and accounts, that is, quietly reducing the reach without alerting the user, was reported on X, Facebook, and Instagram (Vox, 2024; Washington Post, 2024; Murphy, 2025; Time, 2024; The Atlantic, 2022). In 2023, a federal appeals court concluded that the White House had likely violated the First Amendment by urging platforms to take down content (Missouri v. Biden, 2023). Yet this ruling led to few concrete consequences for either the government or the companies.

The UK showed similar examples. The Online Safety Bill, brought in as a way to tackle “harmful content,” gave Ofcom expansive powers over online speech (Institute of Economic Affairs, 2022; Online Safety Act, 2023). This was exemplified by the removal of content around Brexit debates (Schnyder, 2026), climate protests (NBC News, 2019), and anti-government organizing on social media (Article 19, 2023). Many accounts belonging to Labour Party supporters were suspended, content related to climate activism was removed, and pages that were created to coordinate protests were shut down even before the new legislation was fully in place (BBC News, 2023; Reuters, 2025).

These incidents are neither rare nor isolated. They are proof of how states and social media platforms now have the power to shape, filter, and, if needed, block communication to their favor. This paper argues that platform governance through content moderation, account suspension, algorithmic suppression, and state-ordered shutdowns functions as a mechanism of “political closure” (Bakir, 2013). Political closure is the deliberate narrowing of acceptable political opinions, the sidelining of opposition voices, and the construction of an illusion of consensus around dominant state or majority narratives (Beck, 1992; Blumler, 1995). Social media platforms help build and sustain these closures under the pretext of “safety” or “community standards” but in practice tend to reinforce establishment narratives and punish dissent.

This paper aims to provide a comparative mapping of 49 platform-related free speech violations across India, the UK, and the USA. It also aims to show how, despite having different legal and institutional setups, the outcomes nonetheless converge on similar closures. And it also intends to develop the view of platforms as a part of the technological intelligentsia (Wang, 2007), a set of state, corporate, and technical actors who, together, craft and enforce narratives in ways that limit democratic freedom.

Review of Literature

2.1 Digital Platforms as Political Infrastructure

Content moderation is presented as a technical maintenance task but is in fact a political practice. This is how platforms pick and choose the narratives that will dominate any conversation, and everything else is moderated into the margins. Thus, drawing invisible yet effective lines between

acceptable and unacceptable speech (Gillespie, 2019; Roberts, 2019; Trithara, 2024). What qualifies as a “violation” is not a set of rules but a flexible “understanding,” which keeps evolving to cater to platform interests and broader political pressures. Similarly, “community standards” are defined to operate as regulatory regimes over democratic spaces of discourse (Gorwa et al., 2020; Klonick, 2018; Santos et al., 2025; Tech Policy Institute, 2025; Cato Institute, 2024). The platforms absorb domestic politics and profit from market temperaments while evading legal obligations (Cremer & Sheehan, 2022).

Political actors use whatever means at their discretion, formal and informal, to influence the platform moderation policies and decisions, including the removal of critical or oppositional content (Chaudhry, 2024; Gorwa et al., 2020). The impact of these dynamics is intensified due to the high concentration in ownership and control of the main channels through which information and political debate disseminate, making user surveillance and narrative control easier yet systematic at a larger scale (Helbing et al., 2019).

2.2 Internet Shutdowns and State Control

Internet shutdowns have become more common across the world. India, in recent years, has accounted for 41% of all recorded shutdowns, despite a constitutional commitment to free speech (Access Now, 2023). Officials often defend shutdowns as an emergency measure taken to prevent communal violence or to handle urgent security threats (Human Rights Watch, 2023). But the timing of these shutdowns almost always coincides with elections (Mishra, 2023; Internet Freedom Foundation, 2022), major protests, or other political crises. The shutdowns are also inconsistent in execution, impacting certain regions and communities more than others (Chaturvedi & Singh, 2024; SSRN, 2024).

Research has proven shutdowns to be a powerful instrument of political control (Ruijgrok, 2022) that disrupts the flow of information (Freedom House, 2025). As a result, the counter-narratives fail to gain traction. They cut off communication among protesters and weakened the capacity for coordination (Zegart, 2018; Dunn Cavelty & Wenger, 2020). Shutdowns work both as immediate tools of repression and as signals that dissent carries serious costs.

2.3 Algorithmic Governance and Visibility

Researchers have made significant progress in studying algorithms as mechanisms of governance. Bucher (2018) and Gillespie (2014) concluded that ranking, recommendation, and demotion systems are not neutral in their function. They act on the values of their designers and operators (Stark & Diakopoulos, 2020; Gritsenko & Wijermars, 2022; Hunt, 2019). External forces such as regulatory threats, advertiser preferences, and coercion from powerful political actors also impact them (Stanford HAI, 2026).

“Shadow banning,” a recent practice where content/user accounts are downranked, removed from search, or excluded from recommendation feeds by platforms to reduce visibility without notifying users (Leerssen, 2023; Savolainen, 2022). This practice is seen to be done more to accounts that belong to dissenting and marginalized groups (Helbing et al., 2019; King et al., 2013). Thus, algorithmic governance serves as a tool for controlling participation, granting full visibility to some voices while reducing the reach of others (Yeung, 2018; Leerssen, 2023).

2.4 Political Closure and Democratic Erosion

Several studies on echo chambers and polarization have concluded that digital media plays a prominent role in deepening ideological divides and limiting exposure to alternative views in online discourse (Hampton et al., 2014; Gentzkow & Shapiro, 2011). Such behavior leads to the creation of

“political closures,” actively shrinking the range of legitimate political perspectives (Bakir, 2013; Blumler & Gurevitch, 1995). Such behavior makes opposing ideas almost impossible to exist in the mainstream channels, as they are framed as irrational, dangerous, or entirely illegitimate (Klonick, 2023; Zuckerman, 2024). Undisclosed censorship is not the only way this is being done. It is very much a part of the platform infrastructure and practices (Sullivan, 2025). Thus, political closure is a direct outcome of the way governance rules, technical systems, and state power interact (Bucher, 2018; Bakir, 2013).

2.5 The Technological Intelligentsia and State–Platform Collusion

“Technological intelligentsia” is a group of powerful actors who come together to shape how information is produced, spread, and interpreted. The group includes political actors, governments, platform owners, data brokers, tech consultancies, media giants, etc. (Wang, 2007; Woolley & Howard, 2018; Zuboff, 2019). These power blocs work within shared data-driven ecosystems. They collectively determine which narratives expand, who they reach, and what are the boundaries of acceptable speech. (Yeung, 2018; Gorwa et al., 2020; De Gregorio & Streeck, 2024).

And this is how the boundaries between state surveillance, platform moderation, and commercial data-driven marketing blur, producing coordinated forms of influence (Zuboff, 2019). This is how the technological intelligentsia monitors, controls, and fine-tunes political communication to their favor (Loeb & Loeb LLP, 2022; Centre for Digital Democracy, 2022; Chaudhry, 2024). Hence, even in formally free societies, one can find constraints in democratic discourse.

2.6 Research Objectives

This paper seeks to:

Map patterns of platform-related free speech violations in India, the United Kingdom, and the United States (2014–2025), identifying their frequency, forms, targets, and stated justifications.

Examine disparities in platform enforcement, focusing on differences in how opposition groups, marginalized communities, protesters, journalists, and establishment actors are affected.

Analyze regulatory and public justifications for platform governance such as national security, misinformation, hate speech, and child safety and assess how their ambiguity may enable political closure.

Conceptualize platforms as part of a technological intelligentsia, highlighting the combined influence of platform governance, state pressure, and market incentives in shaping political closure across regulatory regimes.

2.7 Research Questions

RQ1: What patterns of platform-related free speech violations emerge across India, the United Kingdom, and the United States between 2014 and 2025?

RQ2: Which political actors and social groups are disproportionately targeted by platform governance, and how does this phenomenon contribute to political closure?

RQ3: How are platform governance actions justified by states and platforms, and how do these justifications obscure political motivations?

RQ4: How do different legal and regulatory frameworks shape platform governance while producing similar outcomes of political closure?

Research Methodology

3.1 Research Design

The study adopts a mixed-methods approach that combines comparative policy analysis with critical discourse analysis. While the primary focus is qualitative, basic quantitative mapping is used to illustrate the distribution of incidents across countries and categories. This integrated approach helps with the identification of broader patterns while also capturing the political significance of individual cases.

3.2 Data and Sampling

The dataset consists of 49 documented instances of free speech restrictions in India (16), the United States (21), and the United Kingdom (12) between 2014 and 2025. Each case involves a state-ordered or platform-led action that limits political speech or organizing. The actions include internet blackouts, account suspensions, content removals, and algorithmic suppression, as well as legal threats and intimidation directed at platforms.

Only incidents with clear documentation of actors, actions, and some form of stated justification were included so that they could be coded in a consistent way.

Secondary sources were used to contextualize and interpret the dataset. These include:

Legal and policy documents (IT Rules 2021, the Online Safety Bill/Act, First Amendment case law, congressional hearings, and testimonies).

Official statements by governments and platforms explaining or defending moderation decisions.

Reporting by major news organizations on speech-related incidents and controversies.

Reports by groups such as Freedom House, Amnesty International, and the Electronic Frontier Foundation, which track trends in censorship and digital repression.

3.3 Variables and Coding Scheme

Each incident was coded along several dimensions. The aim was to capture who acted, against whom, in what context, using which tools, and with what stated rationale.

3.3.1 Incident-level variables:

Country: India, USA, or UK.

Year: from 2014 to 2025.

Type of restriction:

Internet shutdown or blackout

Account suspension or permanent ban

Content removal or blocking

Algorithmic suppression / shadow banning / demonetisation

Legal threats or harassment of platforms

3.3.2 Actor(s) involved:

State (executive, courts, or regulators such as ministries or commissions)

Platform (e.g., Facebook/Meta, Twitter/X, YouTube, Instagram)

Hybrid (state pressure plus platform action)

3.3.3 Target group:

Journalists or media outlets

Opposition politicians or parties

Protest organisers or activists (e.g., farmers' movements, climate protests, labour struggles, BLM, anti-war campaigns)

Marginalised communities (Dalit, tribal, religious minorities, racialised groups)

Individual users or civil society actors

3.3.4 Issue context:

Electoral/campaigning

Protest/mobilization

National security/counter-terrorism

Pandemic/health

Other, specified where relevant

3.3.5 Justification (as stated by state or platform):

J1: National security / terrorism / foreign interference

J2: Misinformation / disinformation / false information

J3: Hate speech / communal violence / incitement

J4: Child safety / sexual exploitation

J5: Community standards / terms of service violations

J6: Electoral integrity / voter suppression prevention

J7: Other (specified in the case record)

3.3.6 Political closure indicators:

PC1: Visible disproportionate targeting (dissenters and opposition hit more than establishment)

PC2: Vague, discretionary justifications that can be weaponised in future

PC3: Precedent-setting actions that normalise restrictions

PC4: Suppression of counter-narratives (e.g., blocking fact-checking or alternative framings)

PC5: Intimidation or consequential impact on future dissent

These codes were assigned using incident descriptions, source documentation, and, where available, official or platform statements associated with each case.

3.4 Analysis Procedure

The analysis unfolded in four stages.

Stage I: Descriptive mapping

All 49 incidents were organized by country, year, restriction type, and target group to identify distributional patterns and address the first research question concerning where and when violations

occur.

Stage II: Comparative legal-policy analysis

The IT Rules 2021 in India, the Online Safety Bill/Act in the UK, and First Amendment rights and case law in the USA were examined, as each regime defines speech, regulates platforms, and enables discretionary enforcement. This stage connected individual incidents to broader legal structures.

Stage III: Discourse analysis of justifications

A subset of 15 incidents (five per country) was selected for close reading. Official statements, platform communications, judicial material, and detailed reporting were analyzed thematically to identify recurring justificatory frames, particularly misinformation, hate speech, harmful content, and national security. This was done to understand how technical and safety language framed political interventions.

Assessment of political closure

Each case was evaluated against closure indicators (PC1–PC5), including disproportionate targeting, vague justification, precedent-setting effects, suppression of counter-narratives, and intimidation. This stage linked empirical findings to the concept of political closure and addressed the remaining research questions.

Findings and Discussion

4.1 Cross-National Patterns in Platform-Related Free Speech Violations

India: Statutory and Visible Closure

India displayed clear and visible state control over digital spaces. The authorities justified these actions using reasons such as public order and national integrity under the **IT Rules (2021)**. Because many interventions occurred through formal orders or court directives, state control appeared legalistic and bureaucratically routine.

USA: Informal and Fragmented Closure

In the United States, strong First Amendment protections limit direct censorship, making closure more indirect. The state influence often occurs informally through communication and signaling, while platforms frame moderation as neutral policy enforcement. Practices such as algorithmic demotion remain difficult to measure but still affect visibility.

UK: Pre-emptive and Legalistic Closure

The United Kingdom reflects a preemptive model of regulation. Broad definitions of harm encourage precautionary intervention, illustrated by early monitoring linked to warnings about potential deepfake election risks.

Comparative Insight

Despite differing legal systems, all three contexts reveal similar outcomes: restrictions framed as public protection, reduced visibility for dissenting voices, and continued dominance of states, major media, and corporate actors within digital spaces.

4.2 Disparate Targeting: Who Is Silenced?

A key pattern in the dataset is unequal targeting. The incidents are not random and repeatedly affect specific groups:

Protest organizers and activists.

In India, many blocked accounts were directly linked to farmers' protests. In the United States, suspensions and visibility reductions affected Black Lives Matter organizers, environmental campaigns, and anti-war voices. In the United Kingdom, climate groups and protest coordinators experienced removals of content and organizing pages.

Opposition politicians and parties.

In India, restrictions on opposition-linked accounts were noted during election periods. In the United States, some campaign content faced reports of downranking during the 2020 cycle. In the United Kingdom, removals affecting labor supporters suggest uneven enforcement patterns.

Marginalized communities.

Indian cases included suppression of Dalit, tribal, and Muslim activist accounts. US incidents involved disproportionate moderation of posts from Black activists and immigrant communities. In the UK, minority and non-establishment voices appeared more vulnerable to removals or heightened scrutiny.

Journalists and fact-checkers.

In India, platform takedowns and legal pressures contributed to some journalists leaving the country. In the United States, critical media figures reported unexplained reach reductions. In the UK, reporting that scrutinized state actions or security policies occasionally faced platform friction.

By contrast, government and establishment actors were rarely suspended or meaningfully restricted, while corporate and institutional accounts aligned with dominant narratives maintained strong visibility. Overall, enforcement appears uneven, with stricter consequences for voices challenging power, illustrating the dynamics of political closure in practice.

4.3 Justifications and Their Political Function

The language used to justify interventions is central to understanding platform governance. Four recurring frames appear across the dataset:

1. The Misinformation Frame

"Misinformation" and "disinformation" function as broad labels applied to contested claims about elections, health, policy, or identity. In India, this included deepfakes and unverified election content; in the United States, critiques of COVID-19 policy or election integrity; and in the United Kingdom, debates around Brexit, immigration, and climate change. While addressing falsehoods is necessary, the label often extends to critical interpretations, allowing dissenting but factual perspectives to be treated as suspicious.

2. The Hate Speech and Safety Frame

Interventions are frequently justified through harm prevention and anti-hate measures. In India, communal speech was flagged as dangerous, contributing to actions against monitoring organizations. In the United States, discussions involving marginalized groups, such as Black Lives Matter, were sometimes framed as incitement. In the United Kingdom, the Online Safety framework enables action against content causing emotional or physical harm. Although framed as protection, this approach can blur the line between preventing violence and shielding power from scrutiny.

3. The Electoral Integrity Frame

Concerns about voter manipulation and foreign interference underpin interventions during

election periods. Takedowns and content labeling in India (2024), the United States (2020 and 2024), and the United Kingdom were presented as safeguards for democratic processes. However, this frame can also be applied to opposition, protest, or journalistic content, potentially constraining legitimate political debate.

4. The National Security Frame

National security is a dominant frame when it comes to India. Internet shutdowns are routinely defended from the lens of security threats, maintenance of public order, or protection of national integrity. These evidently extend to protests, regional unrest, or political disputes. Over time this normalizes exceptional control of dissenting voices being portrayed as threats.

Overall insight

Across contexts, these frames reveal a consistent pattern: political speech is often moderated not explicitly as dissent but through safety-oriented categories such as misinformation, harm, or security. This reframing casts restrictive actions as protective measures, making censorship appear neutral and more difficult to challenge.

4.4 Platforms as Technological Intelligentsia

The findings support the study's central claim that digital platforms are not neutral intermediaries but part of a broader technological intelligentsia shaping public narratives. This is reflected in several patterns:

State–platform coordination and collusion.

In India, the IT Rules (2021) formalized platform compliance with state directives. In the United States, court records and investigative reporting indicate ongoing communication between government officials and platforms regarding moderation priorities. In the United Kingdom, negotiations between Ofcom and major platforms during Online Safety Act implementation occurred largely outside public scrutiny.

Algorithmic alignment with power.

Platform algorithms prioritize engagement and revenue, often amplifying established sources while reducing visibility for contentious or oppositional voices. Practices such as shadow banning, recommendation filtering, and search ranking adjustments subtly shape the reach of dissent and can reflect regulatory or political pressures.

Control over key categories.

Platforms play a decisive role in defining and enforcing categories such as misinformation, hate speech, and harmful content. Although external expertise may inform policy, corporate actors retain ultimate authority over enforcement, significantly influencing whose speech gains visibility.

Overall insight

Taken together, these dynamics position platforms alongside states, intelligence networks, consultants, and data brokers as key intermediaries in shaping political closure. Rather than simply transmitting information, platforms interpret external pressures and embed them into technical design and organizational practice.

Conclusion

Between 2014 and 2025, platform governance shifted from a marginal concern to a defining

feature of democratic politics in India, the UK, and the US. In India, widespread interventions, including 84 internet shutdowns and 134 free speech violations in 2024, alongside the IT Rules (2021), highlight a strong link between state policy and platform control. In the US, constitutional protections limit direct censorship, but informal government pressure continues to shape moderation practices. In the UK, the Online Safety Act has introduced a pre-emptive, risk-based model granting regulators broad oversight of online content.

Despite different legal systems, all three contexts show similar outcomes: platform governance narrowing political debate and reducing the visibility of dissent. Frames such as misinformation, hate speech, electoral integrity, and national security often present political interventions as safety measures, making them less open to scrutiny.

Ironically, many reform efforts reinforce existing power structures, meaning attempts to regulate platforms can also expand the influence of those already in control.

Limitations

Several limitations need to be acknowledged.

Documentation bias: The dataset relies largely on English-language media and national or international databases, which may overrepresent highly visible incidents while overlooking local, small-scale, or non-English cases across all three countries.

Attribution problems: Determining whether moderation actions stem from state directives, platform policy, or their interaction is often difficult. The analysis therefore depends on publicly available justifications and evidence, which may not fully capture underlying motivations or negotiations.

Algorithmic opacity: Practices such as shadow banning and algorithmic suppression are inherently difficult to observe. Researchers typically infer them through reach patterns or leaked information, meaning the dataset likely reflects only a partial view of such interventions.

Temporal limits: The 2014–2025 period encompasses rapid legal, technological, and policy shifts. As a result, the study offers a focused snapshot of a transformative decade rather than a comprehensive historical account.

Recommendations

For democratic governance:

Regulation should decentralize narrative power and limit discretionary control by both states and platforms. Oversight mechanisms must be independent, transparent, and open to civil society involvement.

For platforms:

Moderation rules should be clear and specific rather than vague. Users need accessible explanations and appeals when content is restricted, and algorithms should be evaluated for their impact on viewpoint diversity, not just safety or engagement.

For civil society and journalism:

Greater support for independent journalism, fact-checking, and digital rights advocacy is essential. Civil society can help document incidents, assist affected users, and push for transparency.

For academic research:

Future research should build larger cross-national datasets that capture both visible moderation actions and less visible forms of algorithmic suppression

Reference

- Access Now. (2023). Shutdowntracker: A comprehensive database of internet shutdowns globally. <https://www.accessnow.org/>
- Access Now. (2024). Keep It On: Annual report on internet shutdowns. <https://www.accessnow.org/keepiton/2024-report/>
- ACLED. (2024, September 17). Will past US election turbulence strike again in 2024? <https://acleddata.com/report/will-past-us-election-turbulence-strike-again-2024/>
- Article 19. (2023, November 2). UK: The Online Safety Act is now law, Ofcom's powers as online safety regulator have officially commenced. <https://www.article19.org/resources/uk-online-safety-act/>
- Article 19. (2024, April 15). India: Government blocks websites monitoring hate speech. <https://www.article19.org/resources/india-government-blocks-hate-speech-monitors/>
- Bakir, V. (2013). News, media and terrorism: Public and political responses to 9/11. I.B. Tauris.
- BBC News. (2023, March 27). Jeremy Corbyn banned from standing as candidate for Labour. <https://www.bbc.com/news/uk-politics-65102128>
- Beck, U. (1992). Risk society: Towards a new modernity (M. Ritter, Trans.). Sage.
- Benkler, Y., Faris, R., & Roberts, H. (2018). Network propaganda: Manipulation, disinformation, and radicalization in American politics. Oxford University Press.
- Benkler, Y., Peters, B., Ravi, N., Boylston, J., & Chen, W. Y. (2022). The distribution of partisan news on social media during COVID-19. Misinformation Review. <https://doi.org/10.37016/mr-2020-89> (extends 2018 work)
- Blumler, J. G., & Gurevitch, M. (1995). The crisis of public communication. Routledge.
- Bucher, T. (2018). If... then: Algorithmic design as ethical practice. World Picture, 13, 1–20.
- Carnegie Endowment for International Peace. (n.d.). Global protest tracker. Retrieved January 22, 2026, from <https://carnegieendowment.org/features/global-protest-tracker>
- Cato Institute. (2024, May 20). A guide to content moderation for policymakers. <https://www.cato.org/policy-analysis/guide-content-moderation-policymakers>
- Centre for Digital Democracy. (n.d.). Surveillance marketing 2.0. <https://democraticmedia.org/publishings/surveillance-marketing-2-0>
- Chaturvedi, S., & Singh, A. (2024). A case of internet shutdowns in India. Media and Communication, 12(4), 1–12. <https://www.cogitatiopress.com/mediaandcommunication/article/download/8511/4096>
- Chaudhry, A. G. (2024, July 30). Regulation or repression? Government influence on political content moderation in India and Thailand. Carnegie Endowment for International Peace. <https://carnegieendowment.org/research/2024/07/india-thailand-social-media-moderation?lang=en>
- Council on Foreign Relations. (2024, April 10). Preventing U.S. election violence in 2024. <https://www.cfr.org/reports/preventing-us-election-violence-2024>
- Cremer, J., & Sheehan, D. (2022). Platform rules: The politics of content moderation. Institute of Economic Affairs. <https://iea.org.uk/wp-content/uploads/2022/06/AN-UNSAFE-BILL.pdf>
- De Gregorio, G., & Streeck, J. (2024). Algorithmic borders of visibility: Platform governance, legacy media, and the Israel-Palestine conflict. Communication, Culture & Critique. <https://doi.org/10.1093/cc/c/tzaf001>
- Dunn Cavelty, M., & Wenger, A. (2020). Cyber security politics: A brief history. Internet Policy Review, 9(3), 1–15.
- Election Commission of India. (2024, May 6). ECI/PN/72/2024: Press note on responsible and ethical use of social media platforms by political parties. <https://elections24.eci.gov.in/docs/5yIWJLjQBX.pdf>
- Free Speech Collective. (2024, May 1). India has seen at least 134 violations of free speech in 2024.

<https://freespeechcollective.in/india-has-seen-at-least-134-violations-of-free-speech-in-2024/>

Freedom House. (2024). India: Freedom on the Net 2024 country report. Freedom House. <https://freedomhouse.org/country/india/freedom-net/2024>

Gentzkow, M., & Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126(4), 1799–1839.

Gillespie, T. (2010). The politics of platforms. *New Media & Society*, 12(3), 347–364.

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, distribution, and difference* (pp. 167–194). MIT Press.

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1). <https://doi.org/10.1177/2053951719897945>

Gritsenko, D., & Wijermars, M. (2022). Algorithmic governance: A modes of governance approach. *Regulation & Governance*, 16(1), 169–183. <https://doi.org/10.1111/rego.12367>

Hampton, K. N., Rainie, L., Lu, W., Dwyer, M., Shin, I., & Purcell, K. (2014). Social media and the "spiral of silence." Pew Research Center.

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., & Zuboff, S. (2019). Will democracy survive big data and artificial intelligence? In D. Helbing (Ed.), *Towards digital enlightenment* (pp. 73–98). Springer.

Human Rights Watch. (2021, January 13). Social media companies must end unequal treatment of Black Lives Matter protesters. <https://www.hrw.org/news/2021/01/13/social-media-companies-must-end-unequal-treatment-black-lives-matter-protesters>

Human Rights Watch. (2023, June 14). "No internet means no work, no pay, no food": Internet shutdowns deny access to basic rights in "Digital India." <https://www.hrw.org/report/2023/06/14/no-internet-means-no-work-no-pay-no-food/internet-shutdowns-deny-access-basic>

Hunt, R. (2019). Algorithmic regulation in media and cultural policy. *International Journal of Communication*, 13, 307–325.

Institute of Economic Affairs. (2022, June 26). Online Safety Bill hands unprecedented censorship powers to Secretary of State and Ofcom, warns new research. <https://iea.org.uk/media/online-safety-bill-hands-unprecedented-censorship-powers-to-secretary-of-state-and-ofcom-warns-new-research/>

Internet Freedom Foundation. (2022). Internet blocking and shutdowns in India and international law. <https://ippr.in/index.php/ippr/article/download/282/118/425>

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *Proceedings of the National Academy of Sciences*, 110(37), 14755–14762.

Klonick, K. (2018). The new governors: The people, rules, and processes governing online speech. *Harvard Law Review*, 131(6), 1598–1670. <https://harvardlawreview.org/print/vol-131/the-new-governors-the-people-rules-and-processes-governing-online-speech/>

Leerssen, P. (2023). An end to shadow banning? Transparency rights in the European DSA. *Computer Law & Security Review*, 49, Article 105789. <https://doi.org/10.1016/j.clsr.2023.105789>

Loeb & Loeb LLP. (2022, January 7). Commercial surveillance: Technology, government, and civil rights implications. <https://www.loeb.com/en/insights/publications/2022/08/commercial-surveillance-technology-government-and-civil-rights-implication>

Medianama. (2024, April 5). 'Bolta Hindustan' YouTube channel taken down after MIB order under Section 69A. <https://www.medianama.com/2024/04/223-bolta-hindustan-youtube-channel-taken-down/>

Migration Policy Institute. (2024, December 16). Top 10 migration issues of 2024. <https://www.migrationpolicy.org/programs/migration-information-source/top-10-migration-issues-2024>

Mishra, K. (2023). Impact of 4G internet on political outcomes in India. Indian Statistical Institute. <http://www.isid.ac.in/~acegd/acegd2023/papers/KaranMishra.pdf>

Missouri v. Biden, 83 F.4th 350 (5th Cir. 2023). <https://www.ca5.uscourts.gov/opinions/pub/23/23-30445-CV0.pdf>

Platform Governance and Political Closure: Shadow Banning, Algorithmic Suppression, and Free Speech Violations in India, UK, and USA

- Murphy, M. (2025, May 29). Emerging from the shadows: Shadow banning and misinformation. <https://www.mattamurphy.com/academics/shadowbanning>
- NBC News. (n.d.). London police under fire after banning Extinction Rebellion climate protests. <https://www.nbcnews.com/news/world/london-police-under-fire-after-banning-extinction-rebellion-climate-protests-n1067251>
- Newslaundry. (2024, February 20). The IT ministry blocks 177 social media accounts and links related to farmer protests. <https://www.newslaundry.com/2024/02/20/it-ministry-blocks-177-social-media-accounts-links-related-to-farmer-protests-hindustan-t>
- Online Safety Act 2023, c. 50. (2023). Online Safety Act 2023. UK Public General Acts. <https://www.legislation.gov.uk/ukpga/2023/50>
- Peoples Dispatch. (2024, June 24). Forced departure of foreign correspondent from India sparks concerns over press freedom. <https://peoplesdispatch.org/2024/06/24/forced-departure-of-foreign-correspondent-from-india-sparks-concerns-over-press-freedom/>
- Reuters. (2025, January 22). Facebook removes, then reinstates, page of British anti-racism organisation. <https://www.reuters.com/technology/facebook-removes-then-reinstates-page-british-anti-racism-organisation-2025-01-22/>
- Roberts, S. T. (2019). Behind the screen: Content moderation in the shadows of social media. Yale University Press.
- Ruijgrok, K. (2022, December 19). The politics of internet shutdowns. *Verfassungsblog*. <https://verfassungsblog.de/the-politics-of-internet-shutdowns/>
- Santos, A., Cazzamatta, R., & Napolitano, C. J. (2025). Holding platforms accountable in the fight against misinformation: A cross-national analysis of state-established content moderation regulations. *Convergence*. Advance online publication. <https://doi.org/10.1177/17480485251348550>
- Savolainen, L. (2022). The shadow banning controversy: Perceived governance of social media speech. *Media, Culture & Society*, 44(5), 860–877. <https://doi.org/10.1177/01634437221077174>
- Schnyder, G. (n.d.). Brexit impact tracker. Retrieved January 22, 2026, from <https://www.gerhardschnyder.com/brexit-impact-tracker>
- Scientific American. (2024). 2024 election will define America's stance on immigration. <https://www.scientificamerican.com/article/2024-election-will-define-americas-stance-on-immigration-with-consequences/>
- Singh, P. (2024, May 22). Internet shutdowns across 16 Indian states amid farmer protests and communal violence. *The Wire*. <https://thewire.in/government/internet-shutdowns-16-states-farmers-protests>
- SSRN. (2024). Targeted disruptions: Internet shutdowns in India. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5129468
- Stanford HAI. (2026, January 18). Tuning our algorithmic amplifiers: Encoding societal values in social media AIs. <https://hai.stanford.edu/news/tuning-our-algorithmic-amplifiers-encoding-societal-values-social-media-ais>
- Stark, E., & Diakopoulos, N. (2020). Governing platforms: Extracting and communicating public meanings of algorithmic systems. *AlgorithmWatch Communications Study*. <https://algorithmwatch.org/en/wp-content/uploads/2020/05/Governing-Platforms-communications-study-Stark-May-2020-AlgorithmWatch.pdf>
- Sullivan, G. (2025). Algorithmic governance of 'terrorism' and 'violent extremism'. *London Review of International Law*, 13(1), 47–71.
- Tech Policy Institute. (2025, March 30). No perfect solution: Content moderation will always reflect politics. <https://techpolicyinstitute.org/publications/content-moderation/no-perfect-solution-content-moderation-will-always-reflect-polit>
- The Atlantic. (2022, April 22). Shadowbanning is Big Tech's big problem. <https://www.theatlantic.com/technology/archive/2022/04/social-media-shadowbans-tiktok-twitter/629702/>
- The Guardian. (2024, April 23). Reporting in India is "too difficult" under Modi, says departing Australian journalist. <https://www.theguardian.com/media/2024/apr/23/abc-south-asia-correspondent-avani-dias-narendra-modi-government>
- The News Minute. (2024, May 1). India records 134 counts of free speech violation in four months: Report. <https://www.thenewsminute.com/news/india-records-134-counts-of-free-speech-violation-in-four-months-report>

Time. (2024, March 25). Instagram's political content limit: Everything to know. <https://time.com/6960587/meta-instagram-political-content-limit-off-setting-default/>

Trithara, D. (2024). Agents of platform governance: Analyzing U.S. civil society organizations' role in content moderation. *Telecommunications Policy*, 48(5), Article 102761. <https://doi.org/10.1016/j.telpol.2024.102761>

Vox. (2024, November 2). Were Bad Bunny and Ricky Martin really shadow banned by Instagram? <https://www.vox.com/culture/381925/instagram-shadow-ban-political-election-posts-ricky-martin-bad-bunny-twitter-x-censorship>

Wang, Y. (2007). The technological intelligentsia and China's information society. *Journal of Contemporary China*, 16(52), 279–296. <https://doi.org/10.1080/10670560701251792>

Washington Post. (2024, October 16). Everything we know about 'shadowbans' on social media. <https://www.washingtonpost.com/technology/2024/10/16/shadowban-social-media-algorithms-twitter-tiktok/>

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505–523. <https://doi.org/10.1111/rego.12158>

Zegart, A. B. (2018). Spies, lies, and algorithms: Why the intelligence community is terrible at spotting security threats and what to do about it. *Foreign Affairs*, 97(3), 81–96.

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. *Public Affairs* (broader framework for data/power nexus)

Zuckerman, E. (2024). The ethics of social media: Why content moderation is a moral issue. *Journal of Practical Ethics*. <https://journals.publishing.umich.edu/jpe/article/id/6195/>