

De-platforming and Free Speech: Regulating Private Content Moderation Under Constitutional Law

Katelyn Elisabeth Waller

B.S. Criminal Justice, University of Central Florida, Orlando, FL, United States

Corresponding Author; Katelyn Waller; Email ID:
kewaller2004@gmail.com Corresponding Authors Contact
Number: +1 (618) 964-4285

Abstract

As digital communication platforms have supplanted traditional public forums, private technology corporations exercise unprecedented governance over global speech. This paper examines the point at which constitutional principles should logically constrain private platforms, asking how constitutional democracies can regulate private content moderation without undermining either free speech or private autonomy. Drawing on comparative regulatory analysis, constitutional doctrine, and institutional case studies, this paper synthesizes perspectives from the United States, the European Union, India, and Ghana to examine three interlocking dimensions of the problem. This includes the design of regulatory frameworks that balance state intervention against platform independence, the doctrinal question of whether constitutional free-speech norms can or should apply horizontally to private, non-state actors, and the practical mechanisms (notice, appeal, and reasoned decision-making) through which transparency and accountability can be operationalised inside private platform governance. This paper argues that no single jurisdiction has produced a regime that is both speech-protective and respectful of platform autonomy, and that meaningful accountability requires state intervention to constrain process rather than dictate outcome. This is supported by institutional separation between content-flagging bodies and the government whose conduct is in question.

Keywords: Content moderation, private platforms, free-speech, accountability, unprecedented governance

1. Introduction

In the digital age, social media platforms have become the primary arenas for political and social discourse worldwide. Platforms such as Facebook, YouTube, X, and TikTok are no longer mere websites; they function as the modern public squares where citizens exercise their fundamental right to expression. This shift, however, presents a profound constitutional problem. These platforms are private companies, yet they increasingly perform functions once associated exclusively with public forums. Their unilateral content moderation and de-platforming decisions directly affect core constitutional values such as freedom of expression. The United States Supreme Court's decision in *Moody v. NetChoice, LLC* underscored the deep tension between a private platform's First Amendment right to editorial discretion and state efforts to prevent alleged political censorship. Similarly, in India, while the landmark Supreme Court ruling in *Shreya Singhal v. Union of India* protected online free speech from government overreach by striking down Section 66A of the Information Technology Act, it left unresolved the immense power private intermediaries hold over users. In Ghana, where Article 21(1)(a) of the 1992 Constitution guarantees the right to free speech, the unchecked ability of tech platforms to govern democratic discourse presents an identical, growing challenge. This is a growing issue. Excessive moderation suppresses constitutional rights and silences dissent, while insufficient moderation accelerates online misinformation, hate speech, terrorism, child exploitation, and democratic instability. Establishing the accountability of powerful digital actors is therefore critical to the health of democratic self-governance. The primary purpose of this research is to determine the point at which constitutional principles should logically constrain private platforms. Ultimately, this study asks: how should constitutional democracies regulate private platforms without undermining either free speech or private autonomy? By answering this question, the study contributes a balanced, comparative framework that safeguards the integrity of modern democratic discourse across jurisdictions with markedly different constitutional traditions and enforcement capacities.

Although traditional constitutional law relies on a state-action dichotomy to restrain sovereign tyranny, unilateral private actions, such as shadow-banning, demotion, demonetization, and permanent de-platforming, now exert severe effects on expression

that rival administrative censorship. Recent jurisprudence underscores this global tension: *Moody v. NetChoice, LLC* in the United States reaffirmed platforms' First Amendment editorial rights against state carriage mandates; *Kaushal Kishor v. State of U.P.* in India extended fundamental rights under Articles 19 and 21 directly against private entities, complementing *Shreya Singhal v. Union of India*; and in Ghana, Article 21(1)(a) protects expression yet leaves discourse exposed due to the lack of a comprehensive digital framework. Ultimately, while private platforms should not be classified as traditional state actors nor every removal as a constitutional breach, state intervention must set structural guardrails, regulating the procedural process and transparency of moderation rather than dictating outcomes to protect free expression, preserve corporate autonomy, and prevent executive censorship.

2. Literature Review

2.1 Constitutional Foundations of Free Speech

Many constitutional guardrails for freedom of speech were set in place in which nation-states and their political subdivisions wielded much or absolute power and could, in effect, censor or regulate the speech of people living within their borders. Almost all of such constitutional guardrails are rooted in principles of free speech, where the government is mandated to refrain from inhibiting communication that poses no legitimate threat of harm. Among earlier principles of free speech is John Stuart Mill's utilitarian position on individual liberty (Balkin, 2021; De Gregorio, 2022). Mill's utilitarian viewpoint undergirds an absolute freedom of opinion on all matters, arguing that the only reason for legitimate exercise of dominion over any member of a civilized society against his will would be to prevent danger or harm to other members of society. Accordingly, his position for the protection of free speech and expression in a democratic society is grounded on the premise that liberty is a necessary tool for truth-seeking, moral development, and social progress. While Mill views free speech as an individual right essential for personal flourishing and autonomy (Lubin, 2023), other

philosophers like Meiklejohn present a different perspective, viewing freedom of expression through the lens of democratic self-governance where the idea of individual autonomy is forfeited in favor of a collective agenda. In essence, what is important in a system of freedom of expression is not individual opinions, but rather that what is considered worth saying is spoken (Balkin, 2021). A departure from the dominant libertarian view is the idea of the State as a facilitator of free speech. Philosophers like Owen Fiss argue that the First Amendment right with regard to free speech exists to facilitate collective self-determination. According to Fissian Equality, it behooves the State to ensure access to a diverse marketplace of ideas in order to facilitate well-grounded decision-making among citizens (Nunziato, 2023). Fiss contends that the State ought to act as a fair-minded parliamentarian, mandated to intervene when private actors skew the debate in a way that is detrimental to public discourse (Douek, 2022).

2.2 Private Platforms as Modern Public Squares

Society's pivot into a digital age has expanded active individual participation in democracy and political discourse. However, digital technology can simultaneously function as a tool to control or limit public opinion. Some scholars have argued that social media platforms operating as public forums of communication and expression are essentially public forums even if not legally classified as such (De Abreu Duarte, 2025; Lubin, 2023). A classic example is Facebook's admission in its legal brief, stating it has a vital interest in ensuring that speech on the platform receives the same level of protection as speech in public town squares (Klonick, 2021). The court's decision in *Packingham v. North Carolina* describing media platforms as modern public squares and an avenue for private citizens to make their voices heard remains a crucial reminder of the modern perspective of such platforms (Nunziato, 2023). Critics have argued that media companies are quick to adopt a capitalist view of the First Amendment regarding government interference with telecommunication infrastructure. They contend that government regulation of communication networks violates their freedom of speech. However, the same entities advocate for stronger intellectual property laws that shrink the public domain and restrict the public's access to fair use of

digital content (Balkin, 2021). Other scholars have contended that by their nature and operation, private media platforms function as self-regulating systems of governance that dictate the rules of online speech (Gorwa, 2024). Thus, it would be insufficient to perceive these platforms through the lens of the First Amendment but rather as New Governors who have built a private infrastructure for free speech (Roberts, 2021). These New Governors are the facilitators of the modern triadic model of speech, intersecting the state, speakers, and the publishing platforms (Klonick, 2021). Additionally, these platforms exhibit what can be considered a nation-state mimesis, acting as legislators of the digital community by drafting community standards, deciding cases based on these standards just like traditional courts, and enforcing decisions taken. This is synonymous with sovereign states producing rules and exercising monopoly of enforcement. These entities are formally private corporations largely exempt from constitutional checks and balances that limit government power, although they essentially function as modern public squares with increased power over speech (De Gregorio, 2022; Gillespie, 2022).

2.3 Content Moderation and De-platforming

A major way these platforms implement their governance mechanism over speech is through the tool of content moderation. Content moderation is broadly considered the system and rules used by social media platforms to determine how user-generated content is treated (Roberts, 2021). The mechanisms of content moderation largely include community guidelines, algorithmic moderation, and de-platforming, among others. For most media platforms, voluntary content removal is largely dependent on their community guidelines. Based on these guidelines, platforms prohibit a broad range of lawful speech beyond what is legally considered a violation of law (Douek, 2022; Nunziato, 2023). Just like community guidelines, algorithmic moderation suppresses speech. Only this approach tends to be more industrial through automated tools (Helberger & Samuelson, 2024). Algorithmic moderation tends to filter, suppress, or highlight information based on personal significance to the user rather than social impact (De Gregorio & Stremlau, 2023). Critics also highlight the growing trend of content moderation moving from a simple take-down or leave-up approach towards deplatforming and demonetization. Permanent account bans are often criticized, as it prevents the user from interacting with a vast realm of thoughts. A classic example of this was the

banning of political figures from Twitter following alleged incitement, raising concerns about the control media platforms have over political speech. Platforms are also notorious for using the techniques of content demonization, warning labels, or shadow banning to limit the reach of borderline content without necessarily taking it down. The perceptions around freedom of speech not being equated to freedom of reach have increasingly gained attention as algorithmic suppression continually influences the digital marketplace of ideas by demoting content and limiting visibility of user speech through obscure moderation practices (Keller, 2021).

2.4 Constitutional Regulation of Private Content Moderation

A common ground of the foundational theories is that they infamously serve as the basis of free speech jurisprudence incorporated into national constitutions. In the context of evolving digital regulation of private content moderation, more modern approaches have been adopted, distinguishing between foundational precedents and contemporary cases directly challenging platform decisions. Modern courts' views on freedom of expression reflect a nuanced debate between libertarian protections and regulatory mandates (Nunziato, 2023; Orozco Y Villa, 2025). This is evident in the landmark case of *Moody v. NetChoice*, where the Supreme Court ruled that private social media companies retain First Amendment rights regarding editorial discretion and could not be constitutionally compelled to platform speech against their will. Platforms frequently rely on traditional press analogies to argue in favor of their right to take down content as a form of editorial discretion (Gillespie, 2022). Contrary to the dominant position in the United States, India's digital governance framework tends to impose regulatory obligations on platforms. A prominent feature is India's regulation requiring platform "Terms of Service" to meet government-imposed standards. The framework also requires large platforms to appoint resident grievance officers and contact persons within India who can be subjected to personal and commercial pressures directed at the platforms. In the case of Ghana, scholars have raised concerns about a lack of coherent policy for platform governance and have called for digital constitutionalism (Akim, 2023; De Gregorio & Stremlau, 2023). Major regulatory efforts include the Broadcasting Bill and the Cybersecurity Act. Although yet to be passed, the Broadcasting Bill seeks to articulate the values for quality content on media platforms, and the Cybersecurity Act provides a guide for online behavior.

3. Research Methodology and Literature Foundations

3.1 Theoretical Foundations of Free Speech

To evaluate private content moderation, constitutional analysis must be grounded in three foundational frameworks: John Stuart Mill's libertarian utilitarianism, which advocates for expansive expression based on the harm principle to foster truth within a free marketplace of ideas; Alexander Meiklejohn's theory of democratic self-governance, which prioritizes hearing all viewpoints relevant to public decision-making; and Owen Fiss's concept of state facilitation and affirmative equality ("Fissian Equality"), which positions the state as a fair-minded parliamentarian duty-bound to intervene when unregulated private power distorts democratic debate. A critical analytical requirement within this framework is distinguishing three distinct modes of speech control: state action, involving direct regulatory mandates or censorship governed by constitutional bills of rights; private regulation, encompassing discretionary enforcement of Terms of Service and Community Guidelines by platform operators (Roberts, 2021); and indirect state influence ("jawboning" or coerced moderation), where informal executive pressure or regulatory threats push platforms to remove lawful speech (*Kunal Kamra v. Union of India*; Akim, 2023). Crucially, even where platforms are not legally classified as traditional state actors, private moderation generates severe constitutional concerns because platforms function as market-dominant public squares with no viable alternatives, exercise cloudy algorithmic power to shape global information flow, and risk inducing broad chilling effects and self-censorship among civic participants.

3.2 Private Platforms as the Modern Public Square

Modern digital platforms challenge traditional legal categories by functioning as public spaces while remaining privately owned. Judicial rulings and academic consensus recognise

social media platforms as the primary modern public squares for civic engagement (Balkin, 2021; De Gregorio, 2022; Iglesias Keller & De Gregorio, 2025). This dual nature has introduced structural tensions within platform governance:

- **Capitalist First Amendment Contradiction:** Tech platforms frequently invoke First Amendment protections to resist state regulation of their telecommunication networks while simultaneously seeking robust intellectual property laws that restrict public access and shrink fair use.
- **The "New Governors":** Modern platforms act as self-regulating governance bodies (termed the "New Governors") who manage a triadic model of speech involving state authorities, individual speakers, and private publishing platforms (Klonick, 2021; Gorwa, 2024).
- **Nation-State Mimesis:** By issuing Terms of Service, operating automated and human moderation systems, and establishing internal appellate bodies, private platforms exhibit a form of nation-state mimesis. They legislate community standards, adjudicate disputes, and enforce penalties (such as shadowbanning, demotion, demonetisation, and permanent de-platforming) without being subject to standard constitutional checks and balances (Husovec, 2024).

3.3 Methodological Design and Literature Gaps

To analyze these dynamics, each research strand adopts a targeted methodology within a shared comparative framework:

1. **Comparative Regulatory & Policy Analysis:** Evaluates regulatory regimes across the European Union (Digital Services Act), the United States (Section 230 / First Amendment jurisprudence), India (IT Rules), and Ghana (emerging digital framework) (Husovec, 2024; Keller, 2021).
2. **Comparative Constitutional & Doctrinal Legal Research:** Maps how constitutional models (US strict scrutiny, Indian Article 19 qualified rights, European

Drittwirkung, and Ghanaian Article 21 guarantees) approach speech limits and the horizontal application of fundamental rights (*Kaushal Kishor v. State of Uttar Pradesh*; De Abreu Duarte, 2025).

3. Institutional & Socio-Legal Case-Study Analysis: Examines institutional mechanisms such as the Meta Oversight Board and statutory procedural safeguards to evaluate how notice, appeal, and reasoned decision-making operate in practice (Klonick, 2021; Douek, 2022; Zornetta, 2024).

This study bridges three primary gaps in platform-governance literature:

- **Over-reliance on Western Regulatory Models (Global South Gap):** Most academic literature focuses on US First Amendment law or EU regulations. This paper integrates constitutional perspectives from India, Ghana, and South America alongside US and European regimes (De Gregorio & Stremlau, 2023; STF RE 1037396).
- **Uncertainty Surrounding Horizontal Effect (*Drittwirkung*):** It addresses the lack of legal consensus on how fundamental rights apply to private corporations without breaking apart private property rights and corporate self-governance (De Abreu Duarte, 2025; Iglesias Keller & De Gregorio, 2025).
- **Theory versus Practical Enforcement:** It provides concrete procedural frameworks to enforce transparency and due process without converting private platforms into state-controlled entities (Douek, 2022; Zornetta, 2024).

4. Constitutional Norms Shaping Freedom of Expression: An Analysis of the Horizontal Effect in Democratic Jurisprudence

4.1 Doctrinal Evolution and the Conflict Between Speech and Dignity

Historically, the enforcement of fundamental rights under Part III of the Constitution of India was governed by strict "vertical" applicability, operating exclusively against the

State and its instrumentalities under Article 12. Article 19(1)(a) protects freedom of speech and expression subject to the grounds of restriction outlined under Article 19(2). A central constitutional friction arises when the exercise of free speech under Article 19(1)(a) collides with individual dignity and reputation under Article 21. Traditional legal frameworks required aggrieved citizens to seek common-law tort remedies for private infringements, as constitutional writ petitions under Article 32 or Article 226 restricted to state action.

4.2 The Paradigm Shift: *Kaushal Kishor v. State of Uttar Pradesh*

The Supreme Court of India addressed this divide in *Kaushal Kishor v. State of Uttar Pradesh* (2023). A 5-Judge Constitution Bench held that fundamental rights under Articles 19 and 21 possess direct horizontal effect (*Drittwirkung*), making them enforceable against private individuals and non-state entities. The Court affirmed that certain legal interests exist simultaneously as common-law rights and fundamental rights. However, the Court maintained that restrictions under Article 19(2) remain strictly exhaustive and cannot be expanded implicitly. Legal scholars note that direct horizontal enforceability without statutory mediation risks collapsing the boundary between public constitutional law and private common law, creating legal uncertainty for contract rights, private property, and corporate governance (De Abreu Duarte, 2025).

An Independent Youth-Led Think Tank

Legal Framework	Scope of Enforcement	Mechanistic Basis	Primary Critiques
Traditional Vertical Model	Enforceable only against State & Article 12 instrumentalities.	Constitutional writs reserved for state action; common law for private disputes.	Fails to protect individuals from powerful non-state entities.
Kaushal Kishor Horizontal Model	Enforceable against State, instrumentalities, and private entities/individuals.	Dual-nature rights; actionable via constitutional writs and civil remedies.	Risks overwhelming writ courts; destabilises common-law and IP frameworks.

Source: Kaushal Kishor v. State of Uttar Pradesh, (2023) 4 S.C.C. 1; Constitution of India, Arts. 19, 21.

4.3 Socio-Legal Ramifications: Privacy, Surveillance, and Minority Rights

The horizontal application of fundamental rights influences data privacy, employment conditions, and minority protections. Applying Article 21 horizontally provides a constitutional basis to challenge corporate data harvesting, unauthorised algorithmic surveillance, and restrictive employment social-media policies (Lubin, 2023). Furthermore, horizontal applicability offers legal grounding to challenge private social and institutional discrimination against marginalised communities where state action is absent.

4.4 Digital Constitutionalism and the Algorithmic Public Sphere

Applying horizontal constitutional norms to transnational digital platforms remains complex. Content moderation relies on automated systems programmed according to private corporate policies, often operating without public transparency (Helberger & Samuelson, 2024). Over-reliance on automated filtering can result in over-blocking and prior restraint, impacting both the speaker's right to impart information and the audience's right to receive it (De Gregorio & Stremlau, 2023). Regulatory approaches to digital constitutionalism vary:

- **State Federalism:** Direct government dictation of platform moderation policies through statutory mandates, licensing, and compliance requirements, which can risk state overreach.
- **Civil Society Federalism:** Statutory frameworks that empower civil society groups and users to co-regulate content, balancing public interests with local cultural preservation (Palladino, 2025).

4.5 Comparative Insights: Ghana and Brazil

Comparative jurisprudence across the Global South highlights shifting balances between platform liability and constitutional rights. In Ghana, freedom of expression under Article 21(1)(a) protects civic debate from state interference rather than extending horizontal rights to reduce private speech (Akim, 2023). Meanwhile, in Brazil, the Supreme Federal Court evaluated platform obligations under the Marco Civil da Internet in *Recurso Extraordinário 1037396 /Tema 987* (2024), examining whether platforms should be held directly liable for user-generated harm lacking a specific judicial order.

5. Transparency, Accountability, and Due Process Mechanisms in Private Platform Governance

5.1 Comparative Due Process Models

To operationalise procedural fairness without converting private platforms into state utilities, constitutional democracies utilize distinct models:

- **Direct Horizontality (India):** Extends constitutional writ remedies directly to non-state entities via judicial decision (*Kaushal Kishor*), providing broad protection but introducing legal uncertainty for private governance.
- **Statutory Co-Regulation (EU Digital Services Act):** Enforces indirect horizontal effect by requiring platforms to embed procedural due process into their Terms of Service (DSA, 2022; Husovec, 2024; Nunziato, 2023). The DSA mandates structured notice-and-action systems, internal human-reviewed appeal workflows, and out-of-court dispute settlement mechanisms (Zornetta, 2024).
- **Quasi-Judicial Self-Regulation (Meta Oversight Board):** Applies international human rights standards to evaluate content decisions independently from corporate management (Klonick, 2021; Douek, 2022).

5.2 Operationalising Procedural Pillars

Rather than serving merely as administrative governance principles, procedural safeguards function as substantive shields that directly protect fundamental free-speech rights against the chilling effects generated by vague and arbitrary content moderation. Connecting these procedural protections directly to the core free-speech challenge requires five essential pillars:

1. **Immediate, Specific Notice:** Ensuring users receive immediate, specific notice detailing the precise policy breached alongside the flagged content (DSA, 2022).
2. **Individualized Written Explanations:** Requiring platforms to furnish individualized, written explanations distinguishing automated flags from human evaluations (Zornetta, 2024).
3. **Accessible Internal Appeals:** Establishing accessible internal appeal workflows that guarantees contextual human review (Husovec, 2024).
4. **Searchable Precedent Repositories:** Maintaining searchable, anonymized precedent databases to drive public consistency and uniform enforcement (Helberger

& Samuelson, 2024).

5. **Independent Out-of-Court Settlement:** Granting users access to independent out-of-court dispute settlement mechanisms, such as DSA bodies or quasi-judicial entities like the Meta Oversight Board, to resolve persistent grievances without resort to costly litigation (Klonick, 2021; Douek, 2022).

6. Regulating Content Moderation Without Harming Free Speech or Platform Autonomy

6.1 Seven-Dimension Comparative Matrix

A systematic comparison across major jurisdictions reveals markedly different legal approaches to private platform moderation, balancing state oversight, platform obligations, user rights, and constitutional standards. The United States relies on the First Amendment and Section 230 of the Communications Decency Act (CDA), granting platforms broad statutory immunity and discretionary moderation authority while offering minimal statutory user rights beyond private Terms of Service and contract law, anchored by a strict state action doctrine that protects editorial discretion (*Moody v. NetChoice*; Gillespie, 2022). Conversely, the European Union utilizes the Digital Services Act (DSA, 2022; Husovec, 2024) to enforce an indirect horizontal effect (*Drittwirkung*) through structural co-regulation, imposing tiered due diligence, mandatory risk assessments, and structured notice-and-action obligations while guaranteeing explicit user rights to reasoned statements, human appeals, and independent out-of-court dispute settlements. India takes a direct horizontal approach (*Kaushal Kishor*) under Articles 19(1)(a) and 21 alongside the IT Act and IT Rules, subjecting platform safe harbor protections (*Shreya Singhal*) to strict removal timelines while facing judicial boundaries on state fact-checking powers (*Kunal Kamra v. Union of India*). Finally, Ghana operates under Articles 21, 162, and 164 of its 1992 Constitution and the Cybersecurity Act, featuring basic compliance duties alongside emerging constitutional protections against administrative prior restraint (Akim, 2023).

An Independent Youth-Led Think Tank

Dimension	United States (Sec. 230)	European Union (DSA)	India (IT Rules)	Ghana (Emerging/M DHI)
1. Locus of Obligation	Minimal / Voluntary Discretionary	Systemic / Process Tiered / Risk-adjusted	Case-by-case / Compressed Compliance (Hours)	Emerging / Content-based Unspecified
2. Timelines	Market-driven	Tiered / Risk-adjusted	Compressed (24-72 hours)	Unspecified
3. Liability Architecture	Near-blanket immunity	Conditional safe harbour (Systemic audits)	Conditional safe harbour (Licensing threats)	Penal sanctions (Fact-Check Units)
4. Executive Proximity to Content	Low	Low	High	High
5. Redress & Appeal Architecture	Market-driven	Mandatory external route	Grievance Officer model	Thin / Unstructured

An Independent Youth-Led Think Tank

Dimension	United States (Sec. 230)	European Union (DSA)	India (IT Rules)	Ghana (Emerging/M DHI)
6. Market-Access Leverage	High	High (Global turnover %)	High	Low (Relies on blocking)
7. Capacity vs. Textual Constraint	High judicial capacity	High administrative capacity	Active judicial testing	Strong text / Low capacity

Source: Comparative regulatory synthesis of 47 U.S.C. § 230, Regulation (EU) 2022/2065 (DSA), Information Technology Rules 2021/2023 (India), and Constitution of Ghana (1992).

6.2 Proposed Policy Framework and Regulatory Principles

To avoid the pitfall of treating all private moderation decisions as constitutional speech restrictions, which would destroy platform autonomy and flood courts, policy frameworks must adopt a balanced approach centered on five key principles:

- 1. Process-Focused Regulation Over Outcome Dictation:** Regulators should prioritize process-focused regulation over outcome dictation by mandating strict procedural due process, transparency, and auditability rather than dictating which specific posts or accounts stay up or come down (Douek, 2022; Zornetta, 2024).
- 2. Institutional Independence of Flagging Bodies:** Government flagging agencies or fact-checking units must maintain institutional separation and operate with complete independence from executive political bodies to prevent "jawboning" and political censorship (*Kunal Kamra v. Union of India*; Akim, 2023).
- 3. Severity-Tiered Compliance Windows:** Regulatory frameworks should enforce

severity-tiered compliance windows, strictly limiting short compliance timelines (e.g., 24 hours) to illegal content posing imminent physical harm, such as child exploitation or violent extremism, while preserving longer review periods for contestable political speech.

4. Proportionality in Sanctions: Platforms must observe proportionality in sanctions by favoring targeted remedies, such as warning labels, reduced algorithmic reach, or temporary suspensions, over permanent de-platforming, reserving complete account termination solely for persistent, severe offenders (Keller, 2021).

5. Safe Harbors for Platform Autonomy: Policy regimes should establish safe harbors for platform autonomy that protect companies' ability to curate distinct online communities and enforce Terms of Service standards, provided those standards are clearly disclosed, consistently applied, and procedurally fair (Gorwa, 2024).

7. Constitutional Norms Shaping Freedom of Speech and Expression: Insights from Western States

7.1 The Historical Context and Social Vulnerability

Historically, many Western societies operated as strict police states or relied on economies built on slavery, where everyday citizens possessed absolutely no meaningful freedom of expression under totalitarian authorities. The catastrophic fallout of the Second World War triggered a massive shift, fueling aggressive global advocacy to legally recognize free expression as a fundamental human right. The failure of absolute nation-state sovereignty led directly to international human rights declarations, forcing states to draft laws securing internal and external rights. Today, free expression is the mandatory entry ticket for political liberalism. However, multiple factors are currently driving up social vulnerability and accelerating a decay in the actual practice of free speech (Balkin, 2021; Lubin, 2023):

- Rampant political intolerance.
- Hidden mechanics of social-media algorithms manipulating speech visibility (Helberger & Samuelson, 2024).
- Physical and legal attacks on journalists.
- Deeply entrenched institutional media bias.
- Aggressive social policing and public stigma aimed at unpopular opinions.
- Increasingly restrictive, heavy-handed legal frameworks.

7.2 The United States First Amendment Framework

In the United States, the First Amendment dictates that Congress shall make no law abridging freedom of speech; however, this right is not absolute and is subject to restrictions protecting state safety and integrity. Over decades of rulings, the U.S. Supreme Court has carved out specific exceptions while fiercely defending editorial discretion in digital environments (*Moody v. NetChoice, LLC*). Even when state action targets protected speech, courts reject limited carriage mandates that infringe upon a platform's constitutional entitlement to curate content (Nunziato, 2023; Gillespie, 2022).

7.3 International Frameworks and Global Governance

On the international stage, international bodies and multilateral treaties serve as referees for cross-border speech conflicts. Global standards emphasize that fighting discrimination, hate speech, and intolerance must be balanced alongside protecting core freedoms of thought and expression (De Gregorio, 2022). Modern regulatory interventions aim to build capacity, encourage transparency, and assist states in upholding human-rights compliance without establishing statutory censorship mechanisms (Palladino, 2025; Iglesias Keller & De Gregorio, 2025).

8. Discussion

The comparative findings highlight the core tension facing modern democratic governance: traditional constitutional models designed to constrain state overreach are poorly equipped to govern private digital monopolies without undermining free expression or corporate autonomy (De Gregorio, 2022; Gorwa, 2024). Across the jurisdictions analyzed, regulatory strategies fall along a spectrum between non-interventionist editorial protection and direct horizontal fundamental-rights enforcement. Under the American libertarian paradigm established in cases like *Moody v. NetChoice, LLC* and Section 230, United States jurisprudence prioritizes platform editorial autonomy over public access by relying on free-market dynamics (Gillespie, 2022). However, treating major tech platforms solely as private editors ignores their function as primary public squares and allows corporate entities (termed the "New Governors") to set the boundaries of civic discourse without constitutional accountability (Klonick, 2021; Nunziato, 2023; Roberts, 2021).

In contrast, the Supreme Court of India established in *Kaushal Kishor v. State of Uttar Pradesh* (2023) that fundamental rights under Articles 19 and 21 apply horizontally against private actors. While this direct horizontality protects users against arbitrary corporate decisions, enforcing constitutional rights without statutory mediation risks transforming writ courts into primary content moderators and destabilizing common-law frameworks (De Abreu Duarte, 2025). Furthermore, judicial decisions striking down executive Fact-Check Units in *Kunal Kamra v. Union of India* (2024) demonstrates how statutory oversight can easily shift into political censorship when state-run entities lack institutional independence (Akim, 2023). The comparative findings highlight the core tension facing modern democratic governance: traditional constitutional models designed to constrain state overreach are poorly equipped to govern private digital monopolies without undermining free expression or corporate autonomy (De Gregorio, 2022; Gorwa, 2024). Across the jurisdictions analyzed, regulatory strategies fall along a spectrum between non-interventionist editorial protection and direct horizontal fundamental-rights enforcement. Under the American libertarian paradigm established in cases like *Moody v. NetChoice, LLC* and Section

230, United States jurisprudence prioritizes platform editorial autonomy over public access by relying on free-market dynamics (Gillespie, 2022). However, treating major tech platforms solely as private editors ignores their function as primary public squares and allows corporate entities (termed the "New Governors") to set the boundaries of civic discourse without constitutional accountability (Klonick, 2021; Nunziato, 2023; Roberts, 2021).

The European Union's Digital Services Act (DSA, 2022) offers a middle path through indirect horizontal effect by mandating structural co-regulation (Husovec, 2024). By enforcing transparent notice-and-action mechanisms, mandatory human appeals, and independent out-of-court dispute settlement, the DSA regulates the process of moderation rather than dictating speech outcomes (Douek, 2022; Nunziato, 2023; Zornetta, 2024). This procedural approach, alongside quasi-judicial bodies like the Meta Oversight Board, provides accountability without turning private platforms into state utilities (Klonick, 2021).

Finally, in emerging democracies across the Global South, constitutional guarantees robustly protect media from state censorship, but the absence of comprehensive domestic platform frameworks leaves local civic discourse vulnerable to both external corporate policies and proposed statutory overreach (De Gregorio & Stremlau, 2023). Regulatory design in these regions must account for institutional capacity constraints to avoid relying on crude blocking mechanisms or total deference to foreign platforms.

9. Conclusion

The governance of private content moderation presents a structural constitutional challenge for modern democracies. As digital communication platforms have displaced traditional public forums, private technology corporations now exercise unprecedented authority over public expression, political debate, and fundamental human rights (Gorwa, 2024). This comparative analysis demonstrates that neither total non-intervention nor direct state dictation of online speech provides a viable solution.

Pure libertarian frameworks insulate private platforms from accountability, allowing arbitrary de-platforming and obscure algorithmic suppression to restrict public discourse (Keller, 2021). Conversely, direct application of constitutional rights to private entities (*Drittwirkung*) without clear statutory limits creates legal uncertainty and threatens platform autonomy (De Abreu Duarte, 2025; Iglesias Keller & De Gregorio, 2025). Furthermore, state-dominated regulatory models frequently risk executive overreach and political censorship under the guise of public safety (*Kunal Kamra v. Union of India*; Akim, 2023).

To safeguard democratic discourse in the digital age, legal systems must transition toward a framework of procedural digital constitutionalism (De Gregorio, 2022; Douek, 2022). Regulators should refrain from dictating specific speech outcomes or mandating the removal of lawful content, focusing instead on enforcing auditable due-process requirements such as clear notice, mandatory contextual human review for account suspensions, and published written decisions (Husovec, 2024; Zornetta, 2024).

Furthermore, content-flagging and regulatory oversight bodies must remain institutionally separate from government entities to prevent incumbent political regimes from controlling online speech. Finally, multi-layered redress mechanisms, including independent external dispute resolution and transparent oversight bodies, are necessary to provide accessible remedies for users without overwhelming civil courts (DSA, 2022; Palladino, 2025). By embedding structural procedural protections into platform governance, constitutional democracies can hold private tech platforms accountable while preserving both freedom of expression and corporate autonomy.

References

1. Akim, M. (2023). *Human rights in the (social) media contract: Digital constitutionalism's response to government censorship online* (International Human Rights Internship Program Working Paper Series, Vol. 12, No. 1, pp. 1–46).
2. Balkin, J. M. (2021). To repair our public square, engage the people. *Harvard Law Review*, 134, 1–42.
3. De Abreu Duarte, F. M. (2025). Digital constitutionalism and online content moderation: Three models for the future of online speech. In G. De Gregorio, O. Pollicino, & P. Valcke (Eds.), *The Oxford handbook of digital constitutionalism*. Oxford University Press.
4. De Gregorio, G. (2022). *Digital constitutionalism in Europe: Reframing rights and powers in the algorithmic society*. Cambridge University Press.
5. De Gregorio, G., & Stremlau, N. (2023). Internet shutdowns and AI moderation: Global perspectives on freedom of expression. *International Journal of Communication*, 17, 1845–1864.
6. *Digital Services Act (DSA)*, Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC, 2022 O.J. (L 277) 1.
7. Douek, E. (2022). Governing online speech: From content moderation to procedural justice. *Columbia Law Review*, 122(1), 1–82.
8. Gillespie, T. (2022). Do social media companies have an editorial voice? *Information, Communication & Society*, 25(10), 1388–1403.
9. Gillespie, T., Partin, W. C., & Myers, M. (2020). The complexities of moderation: Governance beyond the state-firm binary. *New Media & Society*, 22(8), 1401–1421.

10. Gorwa, R. (2024). *The politics of platform governance: Accountability, regulation, and the global information infrastructure*. Oxford University Press.
11. Helberger, N., & Samuelson, P. (2024). Algorithmic governance under the Digital Services Act: Building a global transparency regime. *Computer Law & Security Review*, 52, Article 105920.
12. Husovec, M. (2024). *Principles of the Digital Services Act: Platform regulation in the European Union*. Oxford University Press.
13. Iglesias Keller, M., & De Gregorio, G. (2025). Digital constitutionalism as an online speech governance framework: Potentials and pitfalls. *European Constitutional Law Review*, 21(1), 45–72.
14. *Kaushal Kishor v. State of Uttar Pradesh*, (2023) 4 S.C.C. 1 (India).
15. Keller, D. (2021). Systemic safeguards: Structural approaches to freedom of expression in platform regulation. *Harvard Journal of Law & Technology*, 34(2), 523–578.
16. Klonick, K. (2021). The Facebook Oversight Board: Creating an independent institution to adjudicate online free expression. *Yale Law Journal*, 130(8), 2418–2499.
17. *Kunal Kamra v. Union of India*, Writ Petition (L) No. 9792 of 2023 (Bom. H.C. Sept. 26, 2024).
18. Lubin, A. (2023). The rights of digital citizens and the limits of platform power. *Harvard International Law Journal*, 64(2), 311–368.
19. *Moody v. NetChoice, LLC*, 603 U.S. ____ (2024).
20. Nunziato, D. C. (2023). The Digital Services Act and the First Amendment: A comparative structural analysis. *Vanderbilt Journal of Entertainment and*

Technology Law, 25(3), 489–532.

21. Orozco Y Villa, L. H. (2025). *Constitutionalising online content moderation: The role of the CJEU under digital constitutionalism* [Doctoral dissertation, University of Oxford]. Oxford University Press / University of Oxford Thesis Repository.

22. Palladino, N. (2025). Civil society's role in constitutionalising global content governance. *Internet Policy Review*, 14(1), 1–28.

23. Roberts, S. T. (2021). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press. 24. *Shreya Singhal v. Union of India*, AIR 2015 SC 1523 (India).

24. Supreme Federal Court of Brazil [STF]. (2024). *Recurso Extraordinário 1037396 / Tema 987* [Extraordinary Appeal 1037396 / Topic 987].

25. Zornetta, G. (2024). The Digital Services Act as a transparency machine: Systemic risk assessment and the due process of moderation. *European Journal of Risk Regulation*, 15(2), 310–335.