RCMOS – Revista Científica Multidisciplinar O Saber. ISSN: 2675-9128. São Paulo - SP.
Ano VII, v.1 2026 | submissão: 24/07/2026 | aceito: 27/07/2026 | publicação: 30/07/2026

Mechanisms for Containing Online Harms and Their Effects on the Exercise of Freedom and Democracy on the Internet

Os Mecanismos de Contenção e Seus Reflexos sobre o Exercício da Liberdade e da Democracia na Internet

Mecanismos de Contención de Daños en Línea y Sus Efectos sobre el Ejercicio de la Libertad y la Democracia en Internet

Luciana Maia Parente Linhares Bucar – Universidad del Museo Social Argentino (UMSA), linhares.luciana88@gmail.com – PhD candidate in Legal Sciences at Universidad del Museo Social Argentino, Buenos Aires, Argentina. Lattes: http://lattes.cnpq.br/3102260715190863

Abstract: This article examines the mechanisms devised to contain excesses in the Brazilian digital environment and their impact on the exercise of freedom of expression and democracy, focusing on three axes: private content moderation carried out by digital platforms, judicial activism exercised by the Brazilian Supreme Federal Court in the face of regulatory gaps, and, more recently, the jurisprudential redefinition of the civil liability regime applicable to internet application providers. Drawing on a deductive method, bibliographical review, and comparative case-law analysis (United States and European Union), the study examines private content moderation in light of the opacity of the criteria adopted and the absence of democratic oversight; the judgment of Inquérito n.º 4.781/DF, in light of the substantialist, proceduralist, and minimalist conceptions of constitutional adjudication; and the partial and progressive unconstitutionality of Article 19 of Law n.º 12.965/2014, as declared by the Supreme Federal Court in general repercussion Themes n.º 987 and 533. It concludes that, despite the legitimacy of the stated purposes underlying each of these mechanisms, their operation absent clear and specific legal parameters undermines the principle of legality and favors the privatization and judicialization of the assessment of the lawfulness of speech, supporting the case for urgent, specific legislative regulation equipped with procedural safeguards as the appropriate path to simultaneously protect freedom of expression and democracy in the Brazilian digital environment.

Keywords: Freedom of expression. Content moderation. Judicial activism. Marco Civil da Internet. Principle of legality.

Resumo: O presente artigo analisa os mecanismos de contenção dos excessos no ambiente digital brasileiro e seus reflexos sobre o exercício da liberdade de expressão e da democracia, a partir de três eixos: a moderação privada exercida pelas plataformas digitais, o ativismo judicial exercido pelo Supremo Tribunal Federal diante de lacunas normativas e, mais recentemente, a redefinição jurisprudencial do regime de responsabilidade civil dos provedores de aplicações de internet. Utilizando-se de metodologia dedutiva, de revisão bibliográfica e de análise jurisprudencial comparada (Estados Unidos e União Europeia), o estudo examina a moderação privada de conteúdos sob a ótica da opacidade dos critérios adotados e da ausência de controle democrático; o julgamento do Inquérito n.º 4.781/DF, à luz das concepções substancialista, procedimentalista e minimalista de jurisdição constitucional; e a declaração de inconstitucionalidade parcial e progressiva do artigo 19 da Lei n.º 12.965/2014, fixada pelo Supremo Tribunal Federal nos Temas de repercussão geral n.º 987 e 533. Conclui-se que, não obstante a legitimidade dos objetivos declarados por cada um desses mecanismos, sua operação sem parâmetros legais claros e taxativos compromete a reserva legal e favorece a privatização e a judicialização do juízo de licitude do discurso, razão pela qual se defende a urgência de disciplina legislativa específica, dotada de garantias procedimentais, como via adequada para a proteção simultânea da liberdade de expressão e da democracia no ambiente digital brasileiro.

Palavras-chave: Liberdade de expressão. Moderação de conteúdo. Ativismo judicial. Marco Civil da Internet. Reserva legal.

Resumen: El presente artículo analiza los mecanismos de contención de los excesos en el ambiente digital brasileño y sus reflejos sobre el ejercicio de la libertad de expresión y la democracia, a partir de tres ejes: la moderación privada ejercida por las plataformas digitales, el activismo judicial ejercido por el Supremo Tribunal Federal ante lagunas normativas y, más recientemente, la redefinición jurisprudencial del régimen de responsabilidad civil de los proveedores de aplicaciones de internet. Mediante metodología deductiva, revisión bibliográfica y análisis jurisprudencial comparado (Estados Unidos y Unión Europea), el estudio examina la moderación privada de contenidos desde la óptica de la opacidad de los criterios adoptados y la ausencia de control democrático; el juzgamiento del Inquérito n.º 4.781/DF, a la luz de las concepciones sustancialista, procedimentalista y minimalista de jurisdicción constitucional; y la declaración de inconstitucionalidad parcial y progresiva del artículo 19 de la Ley n.º 12.965/2014, fijada por el Supremo Tribunal Federal en los Temas de repercusión general n.º 987 y 533. Se concluye que, no obstante la legitimidad de los objetivos declarados por cada uno de estos mecanismos, su operación sin parámetros legales claros y taxativos compromete la reserva legal y favorece la privatización y la judicialización del juicio de licitud del discurso, razón por la cual se defiende la urgencia de una disciplina legislativa específica, dotada de garantías procedimentales, como vía adecuada para la protección simultánea de la libertad de expresión y la democracia en el ambiente digital brasileño.

Palabras clave: Libertad de expresión. Moderación de contenidos. Activismo judicial. Marco Civil de Internet. Reserva legal.

1 Introduction

The exponential expansion of social networks and other internet applications has, within a few decades, transformed the structure of the public sphere of debate, shifting a substantial share of opinion formation and political discourse to private environments operated by economic agents subject to their own operating logic. In this new setting, the dissemination of disinformation, hate speech, and manifestly unlawful content has come to coexist, often without clear boundaries, with the legitimate exercise of freedom of expression and political criticism, posing to the Brazilian legal system the challenge of curbing abuses in the digital environment without compromising, in the process, the very foundations of the Democratic State governed by the Rule of Law that it seeks to protect.

Given the Legislature's failure to regulate, in a clear, precise, and exhaustive manner, the limits of freedom of expression in the digital environment, this regulatory vacuum has been filled by three distinct mechanisms, whose democratic legitimacy and compatibility with the principle of legality are the central subject of this article: private content moderation carried out by the platforms themselves, on their own initiative or under public and governmental pressure;

judicial activism exercised by the Brazilian Supreme Federal Court in proceedings such as Inquiry No. 4,781/DF, in which measures restricting rights were adopted in the absence of a specific legal standard; and, more recently, the jurisprudential redefinition of the civil-liability regime for internet application providers, embodied in the declaration of the partial and progressive unconstitutionality of Article 19 of Law No. 12,965/2014 (the Marco Civil da Internet), established by the Supreme Federal Court in its judgment of General-Repercussion Themes Nos. 987 and 533.

The question guiding the investigation may be formulated as follows: to what extent are these three containment mechanisms - private moderation, judicial activism, and the new jurisprudential interpretation of Article 19 of the Marco Civil da Internet - capable of curbing abuses in the digital environment without themselves becoming forms of censorship incompatible with freedom of expression and the democratic principle that legitimizes it? The hypothesis advanced is that, notwithstanding the legitimacy of the objectives declared by each mechanism, their operation outside clear and exhaustive legal parameters undermines the principle of legality enshrined in Article 5, II, of the Federal Constitution and encourages the privatization and judicialization of judgments concerning the lawfulness of the speech of others - a judgment that, in a Democratic State governed by the Rule of Law, should fall primarily to the Legislature through a lawmaking process endowed with democratic legitimacy.

To develop this hypothesis, the study adopts a deductive methodology of a bibliographic and jurisprudential nature and draws on comparative law - particularly the United States experience, in light of the doctrine of jawboning and Section 230 of the Communications Decency Act, and the European experience, in light of the Digital Services Act - as a contrasting standard for assessing the institutional adequacy of the Brazilian model. As a further methodological premise, it adopts a predominantly proceduralist orientation, enriched by Cass Sunstein's contributions on judicial minimalism, while recognizing the consistency of the substantive argument, whose relevance is considered throughout the analysis.

The article is structured around three main axes. The first examines private moderation by digital platforms, focusing on the opacity of the criteria adopted, dependence on fact-checking agencies, and the gray area of censorship induced by informal state pressure. The second analyzes judicial activism by the Supreme Federal Court, taking Inquiry No. 4,781/DF as an emblematic case in light of substantive, proceduralist, and minimalist conceptions of constitutional adjudication. The third examines the jurisprudential shift embodied in Themes Nos. 987 and 533, which redefined the civil-liability regime for internet application providers, and compares it with the regulatory models in force in the United States and the European

Union. Finally, the conclusions are synthesized to demonstrate the urgent need for specific and exhaustive legislation surrounded by procedural safeguards as the appropriate means of simultaneously protecting freedom of expression and democracy in Brazil's digital environment.

2 Private Moderation by Platforms

Fearing advertising boycotts, sanctions, and growing regulatory pressure, combined with the absence of clear legal parameters, application providers began, on their own initiative, to engage in intensive content moderation. For this purpose, they use artificial-intelligence systems, automated detection mechanisms, and partnerships with fact-checking agencies. This moderation takes the form of suspending or deleting accounts, removing posts, reducing the reach of content, and other measures intended, according to the platforms, to combat the dissemination of disinformation as well as hate speech, racism, and incitement to violence.

Notwithstanding the legitimacy of the stated objectives, this form of private control raises important questions from the perspective of the Democratic State governed by the Rule of Law. Granting private companies the power to determine, on the basis of criteria that are often opaque and lack democratic oversight, which statements may remain in the digital public sphere transfers a significant share of the governance of public debate to economic agents without observing the safeguards inherent in due process and the principle of legality. In this context, Mariana Valente (2018) warns:

> Anyone today who discusses or advocates on issues related to freedom of expression can think of an example in which a major platform has censored, on its own initiative (and possibly as a result of reports by other users), content that appears lawful and desirable, whether individual or political expression. Yet it seems to me that these same people could also easily imagine content that they believe platforms should prohibit on their own initiative, from non-consensual nudity involving women to hateful actions against minorities. It is practically impossible to remain strictly at either extreme. What happens, however, is that rules on expression are being constructed for these spaces, and these rules are not subject to democratic scrutiny. At most, they are subject to public pressure.
> (VALENTE, 2018, p. 129)

This problem is compounded by the issue of the independence and impartiality of fact-checking agencies. How can it be ensured that such agencies are not subordinated to a particular political bias or line of thought? And how can truth and falsehood be defined while exempting that definition from subjective analysis? The answers are discouraging: there is no purely

objective criterion of truth applicable to opinions, and classifying content as false or hateful inevitably entails a degree of subjectivity, which becomes more serious when entrusted to private entities not subject to democratic oversight.

The issue currently occupies a central position in United States constitutional debate. The doctrine of jawboning - understood as informal coercion or pressure exerted by public authorities on private intermediaries to restrict the circulation of certain speech - dates back to Bantam Books v. Sullivan (1963), in which the United States Supreme Court held that the guarantee of freedom of expression cannot be emptied of meaning by transferring censorship to private parties. This understanding was unanimously reaffirmed in National Rifle Association of America v. Vullo (2024), when the Court recognized that the State may not use the threat of regulatory or economic sanctions to coerce private actors into suppressing expressions of thought with which it disagrees.

Later, in Murthy v. Missouri (2024), when considering allegations that federal officials pressured digital platforms to remove content related to the COVID-19 pandemic and elections, the Supreme Court declined to reach the merits because it found that the plaintiffs lacked standing. In dissent, however, Justice Samuel Alito warned that sufficiently sophisticated campaigns of state pressure may escape judicial review precisely because they operate indirectly and informally.

The comparative experience is particularly relevant to the Brazilian context. The distinction between legitimate state action aimed at protecting fundamental rights and unconstitutional coercion intended to induce the suppression of speech is, in many cases, extremely narrow. In this context, the private-moderation regime examined in this section frequently develops under incentives, guidance, or pressure originating from public authorities themselves, a circumstance that calls for special caution to prevent the constitutional prohibition of censorship from being circumvented through private intermediaries.

The contrary view argues that describing private moderation as censorship is a conceptual error: the constitutional guarantee of freedom of expression is directed at the State, not private parties, so a platform that removes content is itself exercising a freedom - the freedom to define the editorial line of the space it maintains. This is essentially the position adopted by the United States Supreme Court in Moody v. NetChoice (2024), which recognized platforms as holders of freedom of expression with respect to their curatorial decisions. This article takes that objection seriously but disagrees on two points. First, contemporary moderation does not occur in a vacuum of pure private autonomy, but under the influence of fear of liability and public and governmental pressure - precisely the gray area of induced

censorship. Second, the scale and infrastructural position of large platforms, which have become the principal forum for public debate, counsel against transferring to them, without adversarial proceedings or legal authorization, the judgment of the lawfulness of others' speech. The problem, therefore, is not the existence of moderation, but its conversion into a decision-making authority on lawfulness that is removed from democratic and judicial oversight.

3 Judicial Activism and the Collision of Rights

When the conflict between freedom of expression and other rights reaches the courts, judges, faced with regulatory gaps, have exercised broad discretion. Although they employ balancing criteria, their interpretations remain subject to personal convictions and may violate the principles of proportionality and legality and, at times, usurp the authority of the Legislature, which is composed of elected representatives.

Addressing this activism first requires explaining the competing conceptions of constitutional adjudication. For substantivism, of which Ronald Dworkin is a leading exponent, the Constitution incorporates moral principles whose protection is entrusted to the Judiciary; the so-called moral reading authorizes judges to give concrete effect to constitutional values such as dignity and fundamental rights, even against occasional legislative majorities, because rights operate as trumps over majority will. In Brazil, this strand has an influential advocate in Lenio Streck. For proceduralism, whose principal reference is Jürgen Habermas, the legitimacy of law derives from the democratic procedure through which collective will is formed; the Judiciary's role is above all to safeguard the conditions of public debate and the prerequisites of deliberation, not to replace the legislator in selecting substantive values, under penalty of paternalistic guardianship that erodes citizens' autonomy. This approach also includes John Hart Ely's democracy-reinforcing review, aimed at protecting the representative process rather than imposing a value agenda.

To these schools is added Cass Sunstein's judicial minimalism, according to which judicial decisions should be deliberately narrow and shallow, limited to resolving the specific case and avoiding broad holdings that reduce the space for democratic deliberation. From this perspective, issues of great moral, political, or social complexity should, whenever possible, be resolved through the legislative process, preserving the openness of public debate and the legitimacy of democratic choices.

When applied to Inquiry No. 4,781 and to the jurisprudential reconfiguration of the liability regime for application providers under Article 19 of the Marco Civil da Internet,

examined below, the three conceptions lead to different conclusions. The substantive view tends to legitimize more intensive constitutional adjudication whenever a concrete threat to fundamental rights or the democratic regime itself is identified. By contrast, both proceduralism and judicial minimalism recommend greater institutional self-restraint and deference to the legislator, especially in the absence of specific legal rules, thereby reinforcing this article's central objection concerning observance of the principle of legality.

Without disregarding the consistency of the substantive argument - whose relevance will be recognized at the appropriate point - this research adopts, as a methodological premise, a predominantly proceduralist orientation enriched by the contributions of judicial minimalism. From this perspective, where no formal statute clearly, precisely, and exhaustively establishes the criteria for restricting statements in the digital environment, defining the limits of freedom of expression should fall primarily to the Legislature, the constitutional body endowed with democratic legitimacy to regulate the matter, rather than being constructed through judicial discretion.

It is useful to recall the content of these two principles, which operate as standards of review. Regarding the principle of proportionality, Ingo Sarlet, Luiz Guilherme Marinoni, and Daniel Mitidiero explain:

> [...] it unfolds in three stages: (a) suitability, according to which the state measure must be capable of securing the result intended by restricting the individual's right; (b) necessity (the least sacrifice or interference), which requires that, when more than one suitable measure is available, the one that interferes least with the legal sphere be chosen; and (c) so-called proportionality in the strict sense, which, if the first two questions are answered affirmatively, requires a balancing of means and ends in the specific case, which is why most critical observations concerning the principle arise at this level. It is important to recall that, in the sense indicated, proportionality operates as a criterion for assessing the constitutional legitimacy of measures by public authorities that interfere with the protective scope of fundamental rights as defensive rights (negative rights).
> (SARLET; MARINONI; MITIDIERO, 2015, pp. 455-456)

As for the principle of legality, Article 5, II, of the Constitution provides that no one shall be compelled to do or refrain from doing anything except by virtue of law: for conduct to be punishable, a clear, express, and prior statute must define it as such. Consequently, sanctioning someone for disseminating false news would require an express and exhaustive statute on the subject, which, as shown, does not exist in the Brazilian legal system.

The most emblematic example of this activism is Inquiry No. 4,781/DF, known as the Fake News Inquiry, opened in the Supreme Federal Court under Justice Alexandre de Moraes

as reporting justice. During the investigation, the reporting justice ordered the seizure of devices belonging to persons under investigation and the blocking of social-media accounts on the ground that it was necessary to interrupt the dissemination of fake news and speech containing hatred, subversion of order, and encouragement of the breakdown of institutional and democratic normality. In the decision, the Justice stated that the investigated persons were seriously suspected of participating in a scheme to disseminate false news through social networks, placing the independence of the branches of government and the Rule of Law at risk of harm.

Regarding this judgment, José Claudio Monteiro de Brito Filho and Peterson Pedro Souza e Sousa (2020) observed:

> The great majority of those censored in Inquiry No. 4,781/DF expressed political opinions (such as defending dictatorships or the impeachment of a Justice), sharp criticism of judicial activism, and some opinions that were even offensive to the Supreme Court and its Justices. Yet none of the statements listed in the individual decision justifies muzzling the persons under investigation.
> Even if the association of persons disseminating unpleasant opinions on the internet creates some danger to the independence of the branches of government and the Rule of Law, one must ask: what actual harm was caused to others? None. There is only a fear laden with subjectivity.
> It is counterproductive to invoke the defense of the independence of the branches of government when the Court itself not infrequently engages in activism that usurps functions typical of the Legislative and Executive Branches. Likewise, it is fair to consider that the anarchist thought present in the opinions of some of those censored is a consequence of discredit in the legal system to which they are subject.
> In this setting of social and technological change, citizens naturally express themselves according to their understanding. The major difference is that opinion is no longer confined to the threshold of one's home or a circle of friends. With the advent of the internet as a channel of communication, individual expression has come to address the world, and that is healthy for democracy.
> (SOUSA; BRITO FILHO, 2020, p. 50)

Conversely, a significant part of legal scholarship and the courts themselves maintains that the context of coordinated attacks on institutions - culminating in the events of January 8, 2023 - justified exceptional measures to defend the Democratic State governed by the Rule of Law, and that legislative inaction imposed on the Judiciary a duty to protect constitutional interests under concrete threat.

Regarding this dilemma involving freedom of expression and democracy, Michelman's reflection (2007) is of paramount importance. He argues that democracy and freedom of

expression mutually condition each other: although more democracy may often mean more freedom of expression and vice versa, it must also be recognized that freedom of expression may create risks for democracy, which, once weakened, may compromise freedom of expression itself (MICHELMAN, 2007, p. 49 et seq.).

Another delicate aspect of the free expression of opinion is the difficult balancing with other values and rights protected by the Constitution, such as human dignity, without creating censorship. In this regard, Jónatas Machado (2002) states:

> Valuing and proscribing any of the opposing opinions would imply the existence of a system of censorship, which would naturally tend to expand. Therefore, a doctrine restricting speech on the basis of hatred [...] in the name of a morally correct politics of love must be approached with the greatest caution, lest the "new freedom of expression" become confused with the "old censorship." It should be recalled that one of the functions of freedom of expression is to provide an alternative to physical violence in releasing accumulated feelings and tensions, including hatred, or especially hatred. At most, the balance in question may entail limiting extreme forms of overtly produced speech whose form and content are intended to stigmatize, insult, and humiliate a particular group, whether a minority or a majority, beyond any serious purpose of confronting facts, ideas, and opinions. [...] Nevertheless, it [the principle of equal concern and respect] cannot be used to remove from discussion issues as important as the identity of social groups and their relationships with individuals, which are not always free from coercion and discrimination.
> (MACHADO, 2002, pp. 847-848)

Within the Electoral Courts, decisions in the same direction have been recorded. In a decision authored by Justice Benedito Gonçalves (Case No. 0601522-38.2022.6.00.0000), the Superior Electoral Court ordered, until the end of the 2022 elections, the suspension of the monetization of YouTube channels and the screening of a particular documentary. In another proceeding (No. 0601483-41.2022.6.00.0000), a broadcaster was ordered to pay a fine if it disseminated facts known to be false or decontextualized against a candidate, with rights of reply granted - a measure that, in practice, led the outlet to instruct its commentators internally to avoid certain expressions. Such decisions demonstrate the expansion of judicial control over speech, with direct effects on freedom of journalistic information.

4 The Jurisprudential Shift: The (Un)constitutionality of Article 19 of the Marco Civil

The tension accumulated throughout the period examined culminated in a jurisprudential outcome of extraordinary relevance, giving current meaning to the very title of this article. In a judgment that began in November 2024 and concluded on June 26, 2025, the

Supreme Federal Court, in deciding Extraordinary Appeals Nos. 1,037,396 (General-Repercussion Theme 987, reporting Justice Dias Toffoli) and 1,057,258 (Theme 533, reporting Justice Luiz Fux), declared Article 19 of the Marco Civil da Internet partially and progressively unconstitutional.

The Court's rationale was that, by requiring a specific judicial order as a condition for providers' civil liability, the provision failed to afford sufficient protection to constitutional legal interests of great importance - fundamental rights and democracy - thereby creating a state of partial omission. The thesis of progressive unconstitutionality was adopted: Article 19 was not unconstitutional from its inception but became unconstitutional over time as the scale, sophistication, and economic power of platforms and the magnitude of the harm they amplify changed.

The new interpretation moves the Brazilian regime closer to an extrajudicial notice-and-takedown model for a broad range of unlawful content, dispensing with a prior judicial order and, in various circumstances, making notice sufficient to create a duty to remove and the possibility of liability. Distinct regimes were nevertheless preserved: for crimes against honor - slander, defamation, and insult - a judicial decision remained necessary for removal before liability, while removal upon extrajudicial notice was permitted; private messaging, email, and closed-meeting services remained protected by the original text of Article 19 because they involve constitutionally protected private communications. The holding applies from June 26, 2025, until the National Congress regulates the matter.

> Article 19 of Law No. 12,965/2014 (the Marco Civil da Internet), which requires a specific judicial order for the civil liability of an internet application provider for harm arising from third-party content, is partially unconstitutional. There is a state of partial omission because the general rule in Article 19 does not afford sufficient protection to constitutional legal interests of great importance (the protection of fundamental rights and democracy).
> (BRAZIL. Supreme Federal Court. Holding under General-Repercussion Theme 987, established on June 26, 2025)

Three sets of reasons support the decision. The first is the inadequacy of the former model: requiring a prior judicial order for liability made the removal of manifestly unlawful content slow and episodic, allowing harm to spread at the scale and speed characteristic of digital networks. The second is victim protection: the impossibility of liability without a prior judicial order shifted to the injured party the burden of litigating individually against each unlawful item of content, placing that person at a clear disadvantage vis-à-vis globally operating

platforms. The third lies in confronting new threats to the information environment and the democratic regime, which the model established by the Marco Civil da Internet, conceived in 2014, was not designed to address. These threats include the mass dissemination of disinformation amplified by recommendation algorithms and echo chambers; the use of bots, inauthentic accounts, and automated messaging to manipulate public debate artificially; coordinated influence operations aimed at delegitimizing the electoral process and democratic institutions; the propagation of hate speech, racism, misogyny, and incitement to violence; and the systemic risks revealed by attacks on democratic institutions, culminating in the events of January 8, 2023. It is precisely this interpretation - based on the insufficiency of the protection afforded by the prior regime in the face of these new forms of harm to fundamental rights and the Democratic State governed by the Rule of Law itself - that permeates the Supreme Federal Court's reasoning.

Conversely, objections of considerable constitutional weight arise. The first is the so-called chilling effect: faced with the risk of liability, platforms tend to adopt an excessively cautious posture, preventively removing potentially controversial content and, out of prudence, suppressing entirely lawful statements - a phenomenon known as overblocking. Added to this is the privatization of the judgment of lawfulness, whereby private entities lacking a democratic mandate, institutional impartiality, and guarantees of independence are entrusted with determining what may or may not remain in the digital public sphere, a task that, in a State governed by the Rule of Law, belongs primarily to the Judiciary through judicial review and observance of due process.

As a consequence, the guarantees of adversarial proceedings, a full defense, and reasoned decisions are weakened because removal based on a simple extrajudicial notice generally dispenses with prior judicial review. The tension with the principle of legality is also aggravated because, in the absence of clear, precise, and exhaustive legislation on the circumstances in which freedom of expression may be restricted, the will of the democratically elected legislator is replaced by standards of lawfulness formulated case by case by the platforms themselves or derived from jurisprudential constructions. This shift in decision-making undermines one of the foundations of the Democratic State governed by the Rule of Law: in a democracy, normative choices that limit fundamental rights must emanate from the people through their legitimately elected representatives, not from private agents or judicial decisions that substitute for the legislative function. Paradoxically, in attempting to preserve democracy, there is a risk of weakening precisely the democratic principle that gives the solution legitimacy.

Comparison with foreign law highlights the distinctiveness of the Brazilian path. In the United States, Section 230 of the Communications Decency Act (1996) grants providers broad immunity regarding third-party content and protects good-faith moderation, placing it at the pole of maximum freedom - and minimum liability - for intermediaries. In the European Union, the Digital Services Act (EU Regulation 2022/2065) adopts an intermediate solution: it conditions removal on a notice-and-action procedure, prohibits the imposition of a general monitoring duty, and, above all, surrounds moderation with procedural safeguards - a duty to state reasons, internal complaint mechanisms, out-of-court dispute settlement, and transparency duties, with enhanced obligations for very large platforms. The Supreme Federal Court's decision moves Brazil closer to the notice-and-removal model, but does so jurisprudentially and without the architecture of safeguards that accompanies the transfer of responsibility under the European model. It is precisely this mismatch - shifting the burden of judging lawfulness to platforms without establishing due process to mitigate it - that reinforces this article's conclusion: the matter requires specific and exhaustive legislation surrounded by safeguards, not the tacit delegation of speech control to private parties.

From the perspective of this article, the decision is ambivalent and deserves an impartial analysis. On one hand, it may be understood as an institutional response to the legislative vacuum described in the preceding section and as an attempt to protect fundamental rights repeatedly violated in the digital environment - the interpretation defended by those who saw Article 19 as an obstacle to effective redress for victims. On the other hand, it raises precisely the central concern of the hypothesis advanced here: by transferring to platforms the duty to assess the unlawfulness of content under penalty of liability, without prior judicial screening and in the absence of an exhaustive statute, the new regime tends to encourage preventive and excessive removal of speech - the so-called chilling effect - shifting to private entities, stimulated by fear of liability, a judgment of lawfulness that the constitutional logic defended here reserved for subsequent judicial review. The very description of the judgment as an examination of the (un)constitutionality of Article 19 reflects the indeterminacy running through the issue that this article seeks to illuminate.

Final Considerations

The analysis undertaken throughout this article supports the conclusion that curbing abuses in Brazil's digital environment is not organized around a single decision-making center but results from the simultaneous and sometimes overlapping actions of three distinct actors -

private platforms, the Judiciary, and, more recently, the Supreme Federal Court itself as interpreter of the civil-liability regime for providers. Far from constituting a coherent system for protecting freedom of expression and democracy, this multiplicity of actors reveals regulatory fragmentation in which far-reaching decisions about what may or may not circulate in the digital public sphere are made outside clear, precise, and exhaustive legal parameters.

Regarding private moderation, it has been demonstrated that the absence of specific regulation grants digital platforms a curatorial power over public debate whose democratic legitimacy is, at the very least, debatable. The opacity of the criteria adopted, reliance on fact-checking agencies whose impartiality is not institutionally guaranteed, and the permeability of this process to informal state pressure - to which the doctrine of jawboning gives conceptual shape - reveal that the line between private self-regulation and induced censorship is narrow and, in many cases, deliberately obscured.

As for judicial activism, the examination of Inquiry No. 4,781/DF confirmed the hypothesis that, in the absence of specific legislation, the Judiciary tends to fill the regulatory vacuum by resorting to substantive grounds. Although understandable given the urgent need to protect important constitutional interests, this strains the principles of proportionality and legality and brings constitutional adjudication closer to a normative function that, strictly speaking, belongs to the Legislature. The proceduralist and minimalist conceptions adopted as methodological premises of this research do not deny the seriousness of the threats confronted, but warn that a response to those threats, if it is to be democratically legitimate, must emanate from the legislative process rather than judicial discretion.

The jurisprudential shift embodied in Supreme Federal Court Themes Nos. 987 and 533, in turn, synthesizes the tensions identified in the two previous axes. Although the progressive unconstitutionality declared with respect to Article 19 of the Marco Civil da Internet responds to a historically observed inadequacy of protection and benefits victims of repeated harm in the digital environment, the new system moves Brazil closer to a notice-and-takedown model without the procedural safeguards that confer legitimacy on similar solutions in comparative law, as seen in both Section 230 of the United States Communications Decency Act and the European Union's Digital Services Act. By transferring to the platforms themselves, under the influence of fear of liability, the judgment of the lawfulness of third-party content, the new Brazilian regime risks aggravating the very problem it sought to solve: shifting the definition of the limits of speech to bodies not subject to democratic oversight, now reinforced by incentives for preventive removal and overblocking.

The central conclusion of this article follows from this situation: the three mechanisms analyzed, whether separately or together, do not replace - and should not replace - legislation on the matter. Private moderation lacks democratic oversight; judicial activism lacks legitimacy to innovate primarily in defining rights; and the new jurisprudential interpretation of Article 19 lacks procedural safeguards accompanying the transfer of responsibility to platforms. The urgency of the issue does not authorize dispensing with the principle of legality, lest the effort to protect democracy and the fundamental rights threatened by abuses in the digital environment compromise precisely the democratic principle that gives any lasting solution its legitimacy. It is therefore imperative to enact a specific and exhaustive statute surrounded by procedural safeguards - inspired, where appropriate, by the European experience - capable of reconciling the effective protection of fundamental rights violated in the digital environment with the preservation of freedom of expression and the democratic regime that legitimizes it.

References

BRAZIL. Constitution of the Federative Republic of Brazil of 1988. Brasília, DF: Federal Senate, 1988.

BRAZIL. Law No. 12,965 of April 23, 2014 (Marco Civil da Internet). Brasília, DF: Presidency of the Republic. Available at: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Accessed: July 21, 2026.

BRAZIL. Supreme Federal Court. Inquiry No. 4,781/DF. Reporting Justice Alexandre de Moraes. Brasília, 2020.

BRAZIL. Supreme Federal Court. Extraordinary Appeal No. 1,037,396 (General-Repercussion Theme 987). Reporting Justice Dias Toffoli. Judgment on June 26, 2025.

BRAZIL. Supreme Federal Court. Extraordinary Appeal No. 1,057,258 (General-Repercussion Theme 533). Reporting Justice Luiz Fux. Judgment on June 26, 2025.

BRAZIL. Superior Electoral Court. Case No. 0601483-41.2022.6.00.0000. Brasília, 2022.

BRAZIL. Superior Electoral Court. Case No. 0601522-38.2022.6.00.0000. Reporting Justice Benedito Gonçalves. Brasília, 2022.

DWORKIN, Ronald. Freedom's Law: The Moral Reading of the American Constitution. Translated by Marcelo Brandão Cipolla. São Paulo: Martins Fontes, 2006.

ELY, John Hart. Democracy and Distrust: A Theory of Judicial Review. Cambridge: Harvard University Press, 1980.

UNITED STATES. Communications Decency Act of 1996, 47 U.S.C. § 230.

UNITED STATES. Supreme Court. Bantam Books, Inc. v. Sullivan, 372 U.S. 58 (1963).

UNITED STATES. Supreme Court. Moody v. NetChoice, LLC, 603 U.S. ___ (2024).

UNITED STATES. Supreme Court. Murthy v. Missouri, 603 U.S. ___ (2024).

UNITED STATES. Supreme Court. National Rifle Association of America v. Vullo, 602 U.S. ___ (2024).

HABERMAS, Jürgen. Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy. Rio de Janeiro: Tempo Brasileiro, 1997.

MACHADO, Jónatas E. M. Freedom of Expression: Constitutional Dimensions of the Public Sphere in the Social System. Coimbra: Coimbra Editora, 2002.

MICHELMAN, Frank I. Relations between Democracy and Freedom of Expression: A Discussion of Some Arguments. In: SARLET, Ingo Wolfgang (ed.). Fundamental Rights, Information Technology and Communication. Porto Alegre: Livraria do Advogado, 2007.

SARLET, Ingo Wolfgang; MARINONI, Luiz Guilherme; MITIDIERO, Daniel. Constitutional Law Course. 4th ed. São Paulo: Saraiva, 2015.

SOUSA, Peterson Pedro Souza e; BRITO FILHO, José Claudio Monteiro de. Freedom of Expression and Judicial Censorship: An Analysis of the Internet. Brazilian Journal of Philosophy of Law, v. 6, no. 2, pp. 38-54, 2020.

STRECK, Lenio Luiz. Constitutional Jurisdiction and Legal Decision-Making. 4th ed. São Paulo: Revista dos Tribunais, 2014.

SUNSTEIN, Cass R. One Case at a Time: Judicial Minimalism on the Supreme Court. Cambridge: Harvard University Press, 1999.

EUROPEAN UNION. Regulation (EU) 2022/2065 of the European Parliament and of the Council of October 19, 2022 (Digital Services Act).

VALENTE, Mariana. Internet and Censorship: Who Speaks, Who Listens, and Who Defines Truth in the Digital Age? Revista Concinnitas, v. 2, no. 33, pp. 123-133, 2018.

NOTE: This work was carried out by the author as a contribution to her PhD in Legal Sciences, as a Research Student, within the Research Project of Universidad del Museo Social Argentino (UMSA), entitled: "THE (UN)CONSTITUTIONALITY OF THE CONTROL CURRENTLY EXERCISED IN BRAZIL OVER THE EXERCISE OF FREEDOM OF EXPRESSION ON SOCIAL NETWORKS (2018-2024)".