The challenges of regulating social media platforms: A Brazilian scenario analysis

Ingo Wolfgang Sarlet, Andressa de Bittencourt Siqueira

This paper aims to describe and analyse the challenges of platform regulation, especially in the Brazilian context, in order to find out how social media providers have evolved to play a key role in the management of online content. Based on this evolutionary approach, the regulatory schemes that have been considered in the legal literature in this regard are described, allowing to identify the most appropriate model. In sequence, the goal is to explore the recent developments in the Brazilian scenario to regulate social media platforms, in order to access the quality of the regulatory scenario in Brazil, especially in its constitutional dimension. In the Brazilian case, it is noteworthy that the three branches of power at the federal level are involved, to a greater or lesser extent, in the regulation of social networks, with special emphasis on the developments in the matter carried out by the Superior Electoral Court (TSE) and, recently, by the Supreme Federal Court (STF)..

A. Introduction

The extensive power exercised by digital platforms is now evident, especially in terms of their control over discourse, impacting how fundamental rights are exercised in online environments. Especially regarding the social media platform providers[1], it can be seen that they have come to wield greater influence over fundamental rights, mainly freedom of expression

1 Note on terminology: the terms "social media provider" and "social media platform provider" are adopted to refer to companies that manage social media platforms (a.k.a. social networks), while the terms "social media platforms", "social media" and "social networks" are used to describe the online environment enabled by such providers. We use the term "platform provider" to better describe those actors that the Brazilian Federal Statute *Marco Civil da Internet* (MCI – Civil Rights Framework of the Internet, in a free translation) entitles "application provider" (*provedor de aplicações*, in Portuguese). From this perspective, a social media provider is a species of the genus platform provider.

and information, data protection and personality rights,[2] as a result of the mass use of information and communication technologies, intensified by the exponential increase in the number of social media users.

Due to the changing role played by social media platforms in the technological society, intense debates are emerging about how to regulate these platforms, especially in Brazil. In view of the proposed scenario, the following research problem is established: how can the challenges of platform regulation be analysed in the context of constitutional law? Based on the relevant doctrine, legislation, and case law, the focus lies above all on the Brazilian scenario, without neglecting some comparative elements, given the transnational reach of platforms, developments in foreign literature on the subject as well as the common problems faced worldwide, so that the view of platform regulation, to a certain extent, needs to remain open.

Therefore, the first step is to analyse how social networks have changed from mere intermediaries to protagonists. After this examination, we move on to explore the ways in which social media platforms can be managed, through self-regulation, external regulation, or hybrid regulation, so that the most appropriate model is identified. Finally, we explore the latest developments in the Brazilian scenario to regulate those environments and their providers in order to assess the quality of the regulatory panorama in Brazil.

B. The changing role of social media platforms

The understanding of how digital platforms, considered here in a broad sense, should be regulated has changed over time, given the gradual evolution of the way these players act, especially social networks platforms. Due to this change and their broad control over fundamental rights exercised in

2 Among others, see Marion Albers and Ingo Wolfgang Sarlet (eds) 96 *Personality and Data Protection Rights on the Internet: Brazilian and German Approaches* (Springer, Ius Gentium: Comparative Perspectives on Law and Justice, 2022); Indra Spiecker gen. Döhmann, Michael Westland, Ricardo Campos (eds) 64 *Demokratie und Öffentlichkeit im 21. Jahrhundert: zur Macht des Digitalen* (Nomos, Frankfurter Studien zum Datenschutz, 2022); Gilmar F Mendes, Peter Häberle, Ingo W Sarlet, Francisco Ballaguer Callejón *et al* (eds) *Direitos fundamentais, desenvolvimento e crise do constitucionalismo multinível* (Fundação Fênix, 2020); Indra Spiecker gen. Döhmann, 'The difference between online and offline communication as a factor in the balancing of interests with freedom of speech' in Clive Walker and Russel L Weaver (eds) *Free Speech in an Internet Era* (Carolina Academic Press 2013).

online environments, platforms have come to be known as "custodians"[3] or "gatekeepers"[4] and as the "new governors" [5].

It is possible to analyse in two distinct periods the evolution of the role of social network providers, which have gone from being mere intermediaries to effective protagonists, especially in terms of choosing the design of the platform (the tools that will be available to users, e.g., comments, shares, posts), editing their own rules (e.g., Terms of Use and Community Guidelines), carrying out content moderation procedures (e.g., blocking and restoring content and banning users), as well as content curation, i.e. the distribution of different content to different users by means of algorithms, made possible by the collection of users' personal data[6].

The first moment in the evolution of social media platforms, connected to the development of the internet, is characterized by the publication of rules protecting digital platforms, in a broad sense. A landmark in this context is the Declaration of Independence of Cyberspace, published in 1996 by John Perry Barlow at the World Economic Forum in Davos, which defended the free functioning of online environments, that would not be bound by state regulations[7]. Additionally, the exponential growth of providers in the 1990s removed closer state regulation, causing these companies to self-regulate[8].

As for state initiatives, mention should be made of Section 230 of the US Communications Decency Act (CDA) of 1996, which establishes a safe harbour for platform providers that do not edit content, extending this

3 Tarleton Gillespie, *Custodians of the internet – Platforms, content moderation, and the hidden decisions that shape social media* (Yale University Press, 2018) 209.

4 Edoardo Celeste, 'Digital Constitutionalism: A new systematic theorization' (2019) 33 *International Review of Law, Computers and Technology* 76, 79.

5 Kate Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (2018) 131 *Harvard Law Review* 1598, 1662.

6 Jörn Reinhardt, „Fake News", „Infox", Trollfabriken: Über den Umgang mit Desinformationen in den sozialen Medien. Meinungsfreiheit in Zeiten der Internetkommunikation' (2019) 97(107) *Vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik* 97, 99–100; Ivar Hartmann, 'A new framework for online content moderation' (2020) 36 *Computer Law & Security Review* 4 <https://doi.org/10.1016/j.clsr.2019.105376> accessed 30 March 2022.

7 John Perry Barlow, A Declaration of the Independence of Cyberspace (1996) <https://www.eff.org/pt-br/cyberspace-independence> accessed 10 May 2022.

8 Giovanni De Gregorio, 'The rise of digital constitutionalism in the European Union' (2021) 19(1) *International Journal of Constitutional Law* 41, 47; Wolfgang Hoffmann-Riem, 'Autorregulamentação regulamentada no contexto digital' 2019 46(146) *Revista da AJURIS* 529, 540.

exemption from civil liability to those platform providers that establish self-regulatory measures to curb harmful content. In the Brazilian context, Federal Statute n. 12.965/2014, popularly known as the Marco Civil da Internet (Civil Rights Framework of the Internet, in a free translation, hereinafter MCI), although close to the CDA, is not identical to it[9]. MCI was initially designed to allow the liability of application providers, including social networks, only after non-compliance with a court decision, and, exceptionally, application providers can be held liable after the reporting of users in cases of non-consensual disclosure of intimate images[10]. Nowadays – since the STF's decision on the partial (and progressive) unconstitutionality of the Art. 19 MCI enacted in June 2025 –, the framework on liability of platform providers by content generated by their users set forth by the MCI is not applied as originally proposed (see Section D., below)[11].

In this context, it is also possible to refer to the European Directive 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the internal market (hereafter referred to as the E-Commerce Directive). The Directive, among other aspects, defines the impossibility of establishing a general obligation to monitor towards service providers, including social networks, with regard to content generated on their platforms by their users (art. 15 (1), and recitals 46 and 48 of the E-Commerce Directive).

---

9 Anderson Schreiber, 'Civil Rights Framework of the Internet: Advance or Setback? Civil Liability for Damage Derived from Content Generated by Third Party' in Marion Albers and Ingo Wolfgang Sarlet (eds) 96 *Personality and Data Protection Rights on the Internet Brazilian and German Approaches* (Springer, Ius Gentium: Comparative Perspectives on Law and Justice, 2022) 250.

10 "Art. 19. In order to ensure freedom of expression and prevent censorship, providers of Internet applications can only be civilly liable for damages resulting from content generated by third parties if, after specific court order, they do not make arrangements to, in the scope and technical limits of their service and within the indicated time, make unavailable the content identified as infringing, otherwise subject to the applicable legal provisions.
[...]
Art. 21. Providers of Internet applications who make available the content generated by third parties shall be held subsidiarily responsible for the breach of privacy resulting from the disclosure, without the participants' permission of images, videos or other materials containing nudity or sexual acts of private character when, upon receipt of notification by the participant or their legal representative, fails to diligently promote, within the technical limits of their service, the unavailability of that content" (Free translation. MCI 2014)".

11 STF, RE 1.037.396 (Theme 987) e 1.057.258 (Theme 533) (2025) Full Court, judgment on June 26th, (Justices Dias Toffoli and Luiz Fux, respectively).

However, platforms no longer resemble those on which the protective laws were based[12], either those that establish total immunity or those that create more difficult procedures for their liability. Precisely for this reason, Ana Frazão states that it is not possible to use innovation as a pretext for the regulatory vacuum of platforms, in a broad sense, including social networks, given that it is possible to have a regulation that, on the contrary, can encourage technological innovation[13], once identified areas destined for external regulation and others likely to remain within the scope of self-regulation[14].

This new paradigm shift in the protection afforded to platforms can be best observed from the second half of the 2010s when the neutrality of platforms came into question. After the scandal involving the leak and misuse of data from Facebook users, led by Cambridge Analytica, both in the context of the Brexit referendum in 2016 and the US elections in the same year, social media came to be seen not just as a leisure tool, but as a powerful instrument capable of influencing public opinion, whether platform users or not, as well as potentially violating democracies[15].

Thus, in this second stage of the evolution of social media platforms, unlike the first, we can see the adoption of various external counteractions, mainly adopted by national states (e.g., the German Net Enforcement Act – *Netzwerkdurchsetzungsgesetz*, henceforth just the German NetzDG, adopted in 2017 and in full in force between 2018 and May 2024, when partially revoked by the DSA) and by civil society organizations (e.g., the Manila Principles on Intermediary Liability) for greater control over the actions of platform providers.

An inescapable characteristic of this second moment, and a factor of particular concern in terms of the powers of social media providers, is the vast number of users, which has reached the scale of billions worldwide. In Brazil, for example, according to a report published in January 2024 by Data Reportal, there are 187.9 million social media users, representing 86.6%

12 Frank Pasquale, 'Platform neutrality: Enhancing freedom of expression in spheres of private power' (2016) 17(2) *Theoretical Inquiries in Law* 487, 488.

13 Oreste Pollicino and Giovanni De Gregorio, 'Constitutional Law in the Algorithmic Society' in Hans W Micklitz and Oreste Pollicino *et al* (eds) *Constitutional Challenges in the Algorithmic Society* (Cambridge University Press, 2022) 10.

14 Ana Frazão, 'Plataformas digitais e os desafios para a regulação jurídica', in Leonardo Parentoni *et al.* (eds) *Direito, tecnologia e inovação* (D'Plácido, 2018) 656–657.

15 Francisco Balaguer Callejón, 'Redes sociais, companhias tecnológicas e Democracia' (2020) 14(42) *Revista Brasileira de Direitos Fundamentais & Justiça* 25.

of the Brazilian population[16]. Among the most widely used social networks today, Facebook is the most popular in the world, with 3.03 billion monthly active users and 2.02 billion daily active users, according to Meta Investor Report based on data from the second quarter of 2023[17]. Thus, although there is disagreement about the extent of the power and function of social media platforms, it is indisputable that they enable communication and public debate[18].

In this regard, and once the influence of social networks on everyday life has become increasingly evident, the debate on the need to regulate these online environments is gaining ground. For example, the aforementioned German NetzDG was a pioneer in this regard, establishing a regulatory framework for social media platforms in order to curb the spread of hate speech and fake news[19], including, among other specifications, the establishment of heavy fines in the case of non-compliance[20]. As a result, similar laws have been drafted in other countries, such as the French law against hate speech (*Loi Avia*)[21], although it was declared partially unconstitutional

16 Data Reportal, Digital 2024: Brazil (2025) <https://datareportal.com/reports/digital-2024-brazil> accessed 08 March 2025.

17 Meta, Meta Reports Second Quarter – Results (2023) <https://investor.fb.com/investor-news/press-release-details/2023/Meta-Reports-Second-Quarter-2023-Results/default.aspx> accessed 1 September 2023.

18 In 2019, the German Constitutional Court, in its ruling on the so-called *III Weg* case in a preliminary injunction, recognized not only the market dominance of the social network Facebook, but also the key role the platform plays in public debate, with more than 30 million people in Germany accessing it every month at the time, according to data cited in the ruling itself (BVerfGE, 1 BvQ 42/19, 6, 19).

19 William Echikson and Olivia Knodt, 'Germany's NetzDG: A key test for combatting online hate' (2018) CEPS Research Report Thinking ahead for Europe 2018/9, i; Wolfgang Schulz, 'Regulating Intermediaries to Protect Privacy Online – the Case of the German NetzDG' (2018) HIIG Discussion Paper Series 2018-01 5 <https://www.hiig.de/wp-content/uploads/2018/07/SSRN-id3216572.pdf> accessed 29 July 2020; From a Brazilian perspective, see Ingo Wolfgang Sarlet and Gabrielle Bezerra Sales Sarlet, "Liberdade de expressão e discurso do ódio nas mídias sociais – uma análise à luz da jurisprudência da Corte Europeia de Direitos Humanos e da Lei Alemã sobre a Efetividade do Direito na Internet" in Mendes, Häberle, Sarlet, Ballaguer Callejón *et al* (eds) (n. 2) 109.

20 2017 Network Enforcement Act (*Netzwerkdurchsetzungsgesetz*, NetzDG) <https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html> accessed 3 May 2022.

21 Loi n. 2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur internet (*Loi Avia*), <https://www.legifrance.gouv.fr/dossierlegislatif/JORFDOLE000038745184> accessed 3 May 2022.

by the French Constitutional Council[22]. The Charter of Human Rights in the Digital Age was also enacted in Portugal in May 2021 and a debate on its provisions has also emerged[23].

Besides, in the European Union, the Digital Services Act (henceforth DSA) was adopted on October 19th, 2022, based on systemic-risk assessment regulation, while in the United Kingdom, the Online Safety Act (henceforth OSA), approved on October 26th, 2023, has an approach based on duties of care. Besides, in the German context, on May 16th, 2024, the *Digitale-Dienste-Gesetze* (DDG) came into force as domestic legislation to implement and complement the provisions of the DSA, which, despite its relevance, will not be here further developed. From a Brazilian perspective, the Brazilian Fake News Draft Bill (PL n. 2630/2020, henceforth also *PL das fake news*)[24], which combines the approach on systemic-risk assessment and the creation of duties of care, had one of its different versions scheduled for deliberation in the Brazilian House of Representatives (*Câmara dos Deputados*) at the beginning of May 2023, but did not proceed due to a lack of political support. At the time of writing this text, the Brazilian Fake News Draft Bill was still under legislative consideration, and its future remained uncertain. During the course of the Congress' deliberations, a number of other draft pieces of legislation underwent discussion, including PL 4691/2024, proposed in December 2024, which is connected to PL 2120/2023, proposed in April 2023. However, none of these drafts have gained force so far.

In June 2025, in Brazil, the STF has taken a further step towards a regulatory approach of social media platforms[25]. By a majority of 8 to 3 Justices, art. 19 MCI was considered partially unconstitutional. Even though art. 19 MCI was intended to rule on liability of providers due to content generated

22 França, Conselho Constitucional, Decisão n. 2004-496 DC, judgment on June 10th <https://www.conseil-constitutionnel.fr/decision/2004/2004496DC.htm> accessed 3 May 2022.

23 Interview with José Carlos Vieira de Andrade, "Carta dos Direitos na Era Digital entra em vigor na 6.ª feira com falta de consenso sobre artigo polémico" Lusa (15 julho 2021) <https://www.publico.pt/> accessed 15 November 2023.

24 The Fake News Draft Bill, which has been under discussion in the Brazilian National Congress since 2020, was initially intended to curb the spread of fake news online. Today, the Bill's wording goes beyond the disinformation agenda, aiming to regulate, in several aspects, the social media platforms used in Brazil, see (Fake News Draft Bill, Brazilian Federal Senate (2630/2020) <https://www25.senado.leg.br/web/atividade/materias/-/materia/141944> accessed 3 May 2022.

25 STF (n. 11).

by their users, the legal thesis (*tese de repercussão geral*, in Portuguese) established by the STF goes even further. It stated a proposal on the adoption of duties of care (*deveres de cuidado*, in Portuguese) by providers if a systemic risk is identified. It also recommended the presumption of liability if a piece of content is boosted/recommended by the provider or through the use of illegal artificial distribution of content through robots, as well as the adoption of complaint/report channels, publication of internal rules and transparency reports, and called for a regulation that demands legal representation of providers in Brazil[26].

It is precisely in this context, which highlights the broad power of platform providers, characterized above all by the imbalance in relation to the user in isolation[27] and the measures adopted to develop and maintain its business model[28], that a deep analysis on the regulatory panorama arises.

C. Management and regulation of social media platforms

I. Regulatory scope and categorization

With regard to the dynamic between technology and the law, Lawrence Lessig argues that the regulation of platforms is the result of the tension and interaction of four regulatory forces: (i) the code structure of the platforms, (ii) social norms, (iii) market orientations of an economic nature, in addition to the vertical influence of (iv) legal norms, notably state norms, on the online environment[29]. For this author, although regulation by means of the platforms' code structure is the most effective in terms of determining and encouraging behaviour[30], which, depending on the way that it is applied, becomes a factor of particular concern. On the other hand, it is essential to balance these regulatory modalities, above all through the intervention of legal norms, in order to achieve regulation that protects users by balancing these forces[31].

26 For more details on the judgment, see Section D.
27 Ivar Hartmann and Ingo Wolfgang Sarlet, 'Direitos fundamentais e direito privado: a proteção da liberdade de expressão nas mídias sociais' (2019) 16(90) *Direito Público* 85, 99.
28 Balaguer Callejón (n. 15) 585.
29 Lawrence Lessig, *Code: Version 2.0* (Basic Books 2006) 233.
30 ibid 123–130.
31 ibid 233–234.

In the digital environment, given that threats to fundamental rights mainly come from private actors[32], the aim of external regulation, i.e. regulation that is not developed by the platform itself but by an external actor, such as state regulations, is to ensure that the activity carried out by the platforms complies with a series of parameters and achieves certain objectives, defined based on the balance of interests at stake: the platform itself, its users, civil society in a broad sense, as well as state interests.

From a regulatory perspective, Wolfgang Hoffmann-Riem proposes categorizing the possibilities for managing technologies, understood in a broad sense, as follows: self-regulation, state regulation, hybrid regulation, techno regulation, social self-regulation, and state-regulated self-regulation, as well as the possibility of replacing legal norms with extra-legal ethical standards[33].

In order to adapt the analysis to the reality of social network platforms, considering actors with the capacity to establish a regulatory body for the management of such environments, the range of regulatory possibilities is gathered into three main axes: (i) *self-regulation*, which includes social self-regulation, conforming self-regulation or self-conforming, as well as decentralized models; (ii) *external or vertical regulation*, in which state intervention stands out, by a single state or in partnership; as well as (iii) *hybrid regulations*, which include regulated self-regulation and multi-stakeholder regulation.

II. A brief overview of regulatory models

1. Self-regulation

Self-regulated standards, in the context of social media platforms, are considered to be measures that have the scope of the organization and internal management of a given platform (self-compliance) and those aimed at creating social standards (social self-regulation), as well as other measures that are, in some way, outsourced to other subjects, but which, because they are the initiative and guided by the interests of the providers themselves, are in-

32 De Gregorio (n. 8) 46.

33 Wolfgang Hoffmann-Riem, 'Inteligência Artificial Como Oportunidade para a Regulação Jurídica' 2019 16(90) *Revista Direito Público* 11, 31–38; See also Wolfgang Hoffmann-Riem, *Teoria geral do direito digital – Transformação digital e Desafios para o Direito* (Forense, 2021).

cluded here in the scope of self-regulation. Among these measures executed by other agents or technologies are *decentralization*, carried out voluntarily by the users themselves; *techno regulation*, applied by algorithms and the architecture of the network; and *supervision*, carried out by a body or entity set up by the social network provider itself, but independent of it, in order to monitor the decisions taken in the management of the platform.

Self-compliance refers to business decisions for the organization of platforms, which do not require vertical intervention by the state, and is close to the notion of compliance, as it consists of "guidelines for companies' own behaviour"[34]. In the context of social media platforms, examples are the guiding documents for content moderation carried out by humans.

As soon as norms are adopted not only by those who drafted them but also by individuals who did not take part in the decision-making process, such as social media users, Wolfgang Hoffmann-Riem calls this modality social self-regulation (*gesellschaftliche Selbstregulierung*)[35]. The main feature of these norms is the possibility of making them legally binding, or at least of maintaining the expectation of compliance with these norms[36]. In the context of social media platforms, examples include Codes of Conduct, Community Standards and Guidelines, as well as Terms of Service, which define the rights and obligations of platforms and users[37].

Still on the subject of self-regulation, it is worth highlighting *decentralized models* which, although situated within the scope of private governance exercised by platform providers, are characterized by the fragmentation of power among users[38]. An example of decentralization is flagging,

34 Free translation. Hoffmann-Riem, 'Inteligência artificial como oportunidade para a regulação jurídica' (n. 33) 32.

35 Ibid. See also Wolfgang Hoffmann-Riem, *Teoria geral do direito digital* (n. 33) 531.

36 Wolfgang Hoffmann-Riem, *Teoria geral do direito digital* (n. 33) 136-137.

37 Gerald Spindler, 'Löschung und Sperrung von Inhalten aufgrund von Teilnahmebedingungen sozialer Netzwerke – Eine Untersuchung der zivil- und verfassungsrechtlichen Grundlagen' (2019) 35(4) *Computer und Recht* 238, 240; Nicolas Suzor, 'A constitutional moment: How we might reimagine platform governance' 2020 (36) *Computer Law & Security Review* 1, 3 <https://doi.org/10.1016/j.clsr.2019.105381> accessed 18 August 2021.

38 Thomas E Kadri and Kate Klonick, 'Facebook v. Sullivan: Public Figures and Newsworthiness in Online Speech' (2019) 93 *Southern California Law Review* 39, 94; Christian Djeffal, "Soziale Medien und Kuratierung von Inhalten. Regulative Antworten auf eine demokratische Schlüsselfrage" in Indra Spiecker gen. Döhmann, Michael Westland, Ricardo Campos (eds) 64 *Demokratie und Öffentlichkeit im 21. Jahrhundert: zur Macht des Digitalen* (Nomos, Frankfurter Studien zum Datenschutz, 2022).

in which providers allow users to report content that they believe to be infringing[39]. For its part, *technoregulation* (*Technoregulierung*), permeated with concerns and risks, involves governance by algorithms through the use of artificial intelligence on a large scale, replacing human decision-making[40]. In fact, this perspective of platform management is intrinsically connected to the aforementioned concept of regulation by code, proposed by Lawrence Lessig, but which is of particular concern given the high potential for side effects to occur through the indiscriminate use of algorithms and artificial intelligence, such as algorithmic discrimination[41] and misidentification of infringing content[42].

Self-regulation through oversight, on the other hand, is related to the development of mechanisms to monitor the measures applied by social media providers by bodies or entities in a neutral position. Popularly known as the "Facebook Supreme Court," the Meta's Oversight Board is the pioneering and paradigmatic example to this day. Notwithstanding its precursory nature,[43] it cannot be perceived as a Constitutional Court and cannot, in a Constitutional State, occupy its functions. It is important to recognize the Board's limitations, mainly in terms of its action in a few representative cases and its difficulties in setting minimum standards of protection to be applied by Meta at a global level, especially in relation to freedom of expression[44].

39 Kate Crawford and Tarleton Gillespie, 'What is a flag for? Social media reporting tools and the vocabulary of complaint' (2016) 18(3) *New Media & Society* 410, 411.

40 Hoffmann-Riem, 'Inteligência artificial como oportunidade para a regulação jurídica' (n. 33) 36.

41 Laura Schertel Mendes and Marcela Mattiuzzo, 'Discriminação Algorítmica: conceito, fundamento legal e tipologia' (2019) 16(90) *Revista Direito Público* 39, 47–52; Christopher E Peterson, User-generated censorship: manipulating the maps of social media (B.A. Legal Studies, Thesis, MIT 2013) 16, 40.

42 Thiago D Oliva, Dennys M Antonialli and Alessandra Gomes, 'Fighting Hate Speech, Silencing Drag Queens? Artificial Intelligence in Content Moderation and Risks to LGBTQ Voices Online' (2021) 25(2) *Sexuality & Culture* 700, 712.

43 Evelyn Douek, 'Facebook's Oversight Board: Move fast with stable infrastructure and humility' (2019) 21 *North Carolina Journal of Law & Technology* 2 <https://scholarship.law.unc.edu/ncjolt/vol21/iss1/2> accessed 4 March 2022.

44 Ibid 74.

2. External regulation

As mentioned in section B. The changing role of social media platforms, the dynamic and disruptive nature of social networks initially pushed aside state regulation, encouraging self-regulation of such environments[45]. However, as Ana Frazão rightly points out, technological innovation cannot be used as an excuse to avoid regulation,[46] and, above all, with the consolidation of a given technology, it is becoming increasingly difficult to justify exceptions from state regulations[47].

Wolfgang Hoffmann-Riem notes that the term "regulation" (*Regulierung*) is now commonly used to refer to state interventions "in social processes which, with a specific objective, establish general guidelines for behaviour"[48]. It should be noted that external regulation, despite being easily linked to state intervention, can be drawn up and implemented by economic and political blocs, such as the Digital Services Act, as well as multilateral initiatives, such as the Christchurch Call Initiative, drawn up by the governments of France and New Zealand in 2019, and international treaties, such as the Budapest Convention on Cybercrime and its protocol on Xenophobia and Racism.

Thus, regulatory interventions arise when self-regulation triggers undesirable results, as well as from the need to protect fundamental rights and social and cultural values in the online environment[49]. It is recognized, however, that if applied widely and in isolation, this model triggers side effects in areas that require greater flexibility, creativity, and cooperation[50]. However, we do not want to say that this model is outdated, as it remains relevant and up-to-date for specific matters. Ana Frazão even points out

45 Hoffmann-Riem, 'Autorregulamentação regulamentada no contexto digital' (n. 8) 540.

46 Frazão (n. 14) 654.

47 Thomas Wischmeyer, 'The role and practices of online stakeholders' in Mart Susi (ed) *Human Rights, Digital Society and the Law* (Routledge Handbook, 2019) 5.

48 Free translation. Hoffmann-Riem, 'Autorregulamentação regulamentada no contexto digital' (n. 8) 532.

49 Patrícia Baptista and Clara I Keller, 'Por que, quando e como regular as novas tecnologias? Os desafios trazidos pelas inovações disruptivas' (2016) 273 *Revista de Direito Administrativo* 123, 140; Hoffmann-Riem, 'Inteligência artificial como oportunidade para a regulação jurídica' (n. 33) 36.

50 Hoffmann-Riem, 'Inteligência artificial como oportunidade para a regulação jurídica' (n. 33) 36.

that the reformulation of regulatory techniques is necessary in order to allow the rules and procedures implemented to be flexible to new situations[51].

It is therefore not surprising that regulation faces a number of setbacks in order to be sufficiently effective. The code structure of the platforms is one of the main challenges for a regulation that pretends to be solely state-driven[52], especially because the state alone does not have the technical knowledge to implement a sufficiently adequate regulation. Nevertheless, removing the state from the regulatory scheme is not a wise move.

Since the state is the paradigmatic regulator, state regulation faces difficulties in terms of the timing of its actions[53]. If the state regulates a certain segment prematurely, i.e. social networks, there is a risk that the approved regulation will quickly become outdated or will be applied without the slightest legal certainty[54]. However, if there is an excessive delay in regulation, either due to excessive caution or inability to create a set of rules applicable to reality, there is a state protection deficit – or, in more serious cases, a state protection omission. In addition, the regulatory state is challenged in terms of the scope of its intervention, running the risk of drawing up a regulatory framework that is either too comprehensive, with the aim of including as many management possibilities as possible, or too restrained, resulting in excessively limited regulation[55].

In view of this, there is a need to be aware of the separation between areas that are likely to be subject to external authorities, especially state authorities, and those that work better within the scope of self-regulation or based on hybrid models[56], which will be developed further below.

3. Hybrid models of regulation

Hybrid models have emerged as an alternative to the *a priori* uncontrolled management of self-regulation and the state's rigid regulatory capacity and, for this reason, involve two or more actors with the capacity to manage

51 Frazão (n. 14) 651.

52 Patrícia Pinheiro, *Direito digital* (5th edn, São Paulo, 2013) 50; Hoffmann-Riem, 'Autorregulamentação regulamentada no contexto digital' (n. 8) 530–531.

53 Baptista and Keller (n. 49) 145.

54 Hoffmann-Riem, 'Inteligência artificial como oportunidade para a regulação jurídica' (n. 33) 26; Baptista and Keller (n. 49) 145.

55 Baptista and Keller (n. 49) 145.

56 Hoffmann-Riem, 'Autorregulamentação regulamentada no contexto digital' (n. 8) 537; Patrícia Baptista; Baptista and Keller (n. 49) 145.

a given technological segment. With regard to the regulation of social media platforms, there are two possible classifications in terms of hybrid nature: (i) hybridity in the process of creating and drafting regulations, i.e., multistakeholder regulation; (ii) hybridity in terms of supervision and enforcement of standards, i.e., regulated self-regulation.

A terminological nuance needs to be mentioned beforehand, given that the research carried out here is based on the studies of Wolfgang Hoffmann-Riem. According to the author, hybrid regulatory models are characterized by the development of rules in which the state and private actors participate, which emerge from social self-regulation and in which the State participates in the development of the rules[57], thus approaching the notion presented here of multi-stakeholder regulation. The concept of regulated self-regulation presented by Wolfgang Hoffmann-Riem is adopted here in its entirety. According to the proposed categorization presented here, therefore, the difference lies in the fact that hybrid regulation is a genus of which multi-stakeholder regulation and regulated self-regulation are species.

The multistakeholder regulation model is also gaining ground as an alternative, especially as it involves standardization through a regulatory governance triangle[58]. In this model, each actor – states, NGOs, and private companies – is divided into zones of action, maintaining their autonomy while fostering cooperation between them[59]. There are a number of benefits to including the widest range of players in the debate on social media platforms, especially platform providers, otherwise regulation will be ineffective[60].

In this regard, Ronaldo Lemos goes further and points out that the multi-stakeholder model is not a point of arrival, but rather a starting point. Not only do we need to understand the most diverse perspectives on a given technological segment, but in order for regulation to be able

57 Hoffmann-Riem, 'Inteligência artificial como oportunidade para a regulação jurídica' (n. 33) 35.

58 Kenneth Abbott and Duncan Snidal, 'Strengthening international regulation through transmittal new governance: Overcoming the orchestration deficit' 2009 (42) *Vanderbilt Journal of Transnational Law* 501, 513.

59 ibid 501; Robert Gorwa, 'The platform governance triangle: conceptualizing the informal regulation of online content' (2019) 8(2) *Internet Policy Review* 1, 7ff.

60 Wischmeyer (n. 47) 14. At the same time, Thomas Wischmeyer also draws attention to the limitations of the multisectoral regulation model, since there is not enough consensus on illegal content on the Internet, especially hate speech and fake news, which could lead to any regulation coming from this model being hampered, ibid 12–13.

to achieve minimally satisfactory results, it is necessary to avoid radical positions and to compromise[61] – one of the main lessons learned from the Brazilian MCI decision-making process.

The Brazilian MCI, although is legislation enacted by the state, originated from a debate with various sectors of society, which is why it is pointed out as an example of a multistakeholder regulatory model[62]. In the Brazilian scenario, the consulting process with users, companies, civil society organizations, government sectors, and universities, in which all those interested were able to provide public comments, lasted eighteen months, so that the stakeholders collaborated in the preparation of a draft which, in the end, was submitted to the Brazilian National Congress, resulting in advanced legislation for the time, and is considered an achievement in terms of protecting rights on the Internet[63].

The second model of hybrid regulation discussed here – regulated self-regulation – emerges as a procedural model, focusing on autonomy and cooperation between agents with management capacity[64], in order to acquire technical knowledge and overcome the complexities of the digital environment[65]. For this reason, Dan Wielsch mentions that the development of digital services, including social networks, must be linked to legal norms that ensure the fundamental rights of users, especially those linked to communicative freedoms, and the autonomy of the institutions that provide the space for the exercise of these rights[66].

Based mainly on the studies of Wolfgang Hoffmann-Riem, it is understood that the model based on regulated self-regulation, in general terms, involves a hard normative core set by the state, while a margin of choice is provided to the platforms, outside the scope of this state core, in which private entities themselves can define internal rules allied to technological

61 Ronaldo Lemos, 'The Internet Bill of Rights as an Example of Multistakeholderism' in Carlos Affonso Souza, Mario Viola *et al* (eds) *Brazil's Internet Bill of Rights: A Closer Look* (2nd ed, ITS Rio, 2017) 48.

62 Ibid 42.

63 Ibid 42–43.

64 Ingo Wolfgang Sarlet, 'Liberdade de expressão e o problema da regulação do discurso do ódio nas mídias sociais' (2019) 5(3) *Revista Estudos Institucionais* 1207, 1230; Pollicino and De Gregorio (n. 13) 16.

65 Wolfgang Hoffmann-Riem, *Teoria geral do direito digital* (n. 33) 137.

66 Dan Wielsch, 'Die Ordnungen der Netzwerke. AGB – Code – Community Standards' in Martin Eifert and Tobias Gostomzyk (eds) *Netzwerkrecht – Die Zukunft des NetzDG und seine Folgen für die Netzwerkkommunikation* (Nomos, 2018) 73; Spiecker genannt Döhmann (n. 2).

development and innovation[67]. It is therefore an action by the state, based on cooperation and trust, as the guardian of individual and collective interests, to prevent reckless self-regulation, for example by creating a regulatory framework and/or state incentives for social media platforms, which are subsequently managed by their providers with relative autonomy[68].

The paradigmatic example of regulated self-regulation of social media platforms is the pioneering proposal contained in the German NetzDG, full in force between 2018 and 2024, which listed a series of duties to be fulfilled by providers of such online environments. Under the terms of the legislation, providers of social media platforms with at least two million users in Germany are obliged, among other things, to monitor and block content deemed to be infringing on the basis of the crimes set out in the German Criminal Code, which is generally removed from the platform within 24 hours of the provider becoming aware of it, or within seven days in borderline cases where the definition of illegality is difficult to assess, under penalty of a fine. Platforms providers, based on NetzDG, had also to publish transparency reports regarding content removal. In other words, NetzDG aimed to create commitments with private entities that have large market power[69] and the capacity to manage the environments at hand, based on state legislation itself (i.e. the Criminal Code), which fits in with the concept of regulated self-regulation presented here.

The NetzDG, despite the pioneering spirit and progress represented by the legislation in terms of platform regulation, is not immune to criticism from specialized literature, which highlights not only the potential for the law to violate European rules[70] but also shows concern about the privatization of decision-making power over illegal content, as well as the possibility of overblocking content in order to avoid the payment of fines[71], which,

67 Wolfgang Hoffmann-Riem, *Teoria geral do direito digital* (n. 33) 137.

68 Ibid 136–137.

69 As recognized by the German Constitutional Court in the *III Weg* case, highlighted above (see n. 18), v. BVerG, 1 BvQ 42/19, paras. 6 e 19.

70 Gerald Spindler, 'Internet Intermediary Liability Reloaded. The New German Act on Responsibility of Social Networks and its (In-) Compatibility with European Law' (2017) 8 *Journal of Intellectual Property*, Information Technology and Electronic Commerce Law 166, 175.

71 Ibid 166; Schulz (n. 19); Sandra Schmitz; Christian Berndt, 'The German Act on Improving Law Enforcement on Social Networks (NetzDG): A Blunt Sword?' (2018) <https://ssrn.com/abstract=3306964> accessed on 15 July 2020; Matthias Cornils, 'Behördliche Kontrolle sozialer Netzwerke: Netzkommunikation und das Gebot der Staatsferne' in Martin Eifert and Tobias Gostomzyk (eds) *Netzwerkrecht – Die*

despite the relevance of the criticisms, will not be explored in depth. As already stated above, the German NetzDG was partially revoked by the DSA and the domestic main legislation on the matter is the German DDG since May 2024.

In addition, it is important to emphasize that regulated self-regulation implies the protection of fundamental rights in online environments through procedure[72]. As Jörn Reinhardt rightly demonstrates, once their leading role has been recognized, social media platforms are gradually improving their community terms and standards, as well as developing structured procedures for moderating user-generated content[73], such as the aforementioned implementation of the Meta's Oversight Board. Due to the flawed nature of human interpretation and limited artificial intelligence mechanisms, there is a growing demand for decisions to be re-examined through appeals within the platform itself, giving rise to the need to implement procedural safeguards, to guarantee impartiality in decisions, as well as limited discretion on the part of the decision-maker[74].

With regard to regulated self-regulation with an emphasis on procedures, the NetzDG is also a leading example – even though Art. 3 NetzDG is no longer into force. According to the terms of the law, the social media platform provider can offer the user with a right of reply before deciding on the unlawfulness of the content (art. 3, (3)(a) of the NetzDG) in cases of non-manifest unlawfulness, as well as informing the user of its decision and consequent justification (art. 3, (2)(5)(a) of the NetzDG).

In addition to legislative provisions, in 2021, a new chapter in the evolution of regulated self-regulation of social networks was opened by the German Federal Court of Justice (*Bundesgerichtshof*, hereinafter BGH), in which compliance with specific parameters in content moderation procedures was required, namely: the obligation of social network providers to notify the user at least after the removal of content, in addition to the prior notification required in the case of partial or total blocking of

---

Zukunft des NetzDG und seine Folgen für die Netzwerkkommunikation (Nomos, 2018).

72 Wielsch (n. 66) 90.

73 Jörn Reinhardt, 'Algorithmizität und Sichtbarkeit – Konflikte um Bilder in den sozialen Medien' in Eva Schürmann and Levno von Plato (eds) 4 *Rechtsästhetik in rechtsphilosophischer Absicht* (Nomos, 2020) 254.

74 Nicolas Suzor, 'Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms' (2018) 4(3) *Social Media + Society* <https://doi.org/10.1177/2056305118787812> accessed 8 September 2021.

user accounts[75]. These decisions are in line with what Gerald Spindler has already pointed out, regarding the need to develop judicial mechanisms to deal with disputes arising from content platforms[76].

In fact, the BGH's decisions show that regulated self-regulation of social media platforms can take place through the Judiciary, based on procedural regulation and the moderation structure of a given platform. Augusto Aguilar Calahorro highlights the indispensability of access to the courts in the digital society, especially with regard to supranational and international courts, from the perspective of multilevel constitutionalism[77], which, despite its relevance, will not be explored in depth here.

Finally, it can be seen that the current configuration of the regulatory landscape allows for multipolar regulation, in which the models briefly discussed so far are applied simultaneously[78]. There is both self-regulation by the platforms, which adopt self-managing rules, as well as making use of artificial intelligence, distributing the ability to identify infringing content among users and building supervision mechanisms, and there are also certain sectors of the platforms subject to external regulation and others subject to hybrid regulation, depending on the country or supranational structure from which any analysis is carried out.

D. Recent developments on social media regulation in Brazil

As for the Brazilian regulatory structure for social media platforms, recent developments show the need to move forward with the process of building a regulatory model that is consistent and effective, but above all constitutionally adequate. The issue here is highly controversial, complex and involves not only political and economic interests of almost incalculable magnitude, but also interests of the members of society as a whole. Sig-

75 As decided by the German Federal Supreme Court (*Bundesgerichtshof*) on BGH, III ZR 179/20, paras. 87–88, as well as on BGH, III ZR 192/20, para. 99.

76 Spindler (n. 70) 175. See also Amélie Heldt, 'Content Moderation by Social Media Platforms: The Importance of Judicial Review' in Edoardo Celeste, Amélie Heldt and Clara Iglesias Keller (eds) *Constitutionalizing Social Media* (Hart Publishing 2022) 264.

77 Augusto Aguilar Calahorro, 'Direitos fundamentais, desenvolvimento e crise do constitucionalismo multinível', in Mendes, Häberle, Sarlet, Ballaguer Callejón *et al* (eds) (n. 2) 708–709.

78 Lessig (n. 29) 233; Zulmar Fachin, 'Desafios da regulação do ciberespaço e a proteção dos direitos da personalidade' (2021) 25(56) *Revista Jurídica FURB* 1, 16.

nificant developments have been identified at least since the last election campaign in 2022, mostly related to the need to tackle disinformation on social media platforms[79].

For instance, in the final stage of the 2022 electoral process, the Brazilian Superior Electoral Court (*Tribunal Superior Eleitoral* – TSE) issued the Resolution n. 23.714/2022 – within the scope of its functions during election periods – with the aim of more effectively combating the so-called *fake news*, which is defined as content "that is intended to undermine the integrity of the electoral process, including the processes of voting, collecting and counting of votes", given the dynamism of digital environments and their widespread use during the electoral campaign period[80].

Among other provisions, the mentioned Resolution, whose effects ended alongside the 2022 election period, established a one-hour deadline to remove disinformation content against the integrity of the electoral process between the election's eve and the third subsequent day (Art. 2º, § 2º of the Resolution). It also provided that it was unnecessary to file isolated lawsuits to remove disinformation electoral content that has been replicated on other websites. In cases where disinformation content that undermines the integrity of the electoral process, for which a decision to remove has already been issued by the TSE Full Court and which has reappeared on social media in an identical format, it is possible for the TSE Presidency to determine, by order, the extension of a collegiate decision already issued (art. 3, *caput* of the Resolution), indicating, in the same act, the URLs, URIs or URNs with identical content that should be removed by the provider.

Not surprisingly, the constitutionality of this Resolution was analysed by the STF in the Direct Action for the Declaration of Unconstitutionality (*Ação Direta de Inconstitucionalidade* – ADI) 7261. The collective decision upheld, by a majority, that by issuing the Resolution the TSE acted legitimately within the scope of its prerogatives, reinforcing that the reaction time, in the electoral disinformation scenario, if it is short, can impose immeasurable damage to the legitimacy of the election, since it is recog-

79 Ingo Wolfgang Sarlet and Andressa de Bittencourt Siqueira, 'Direitos fundamentais e regulação de plataformas digitais no Brasil' (2023) Consultor Jurídico, *Coluna Observatório Constitucional* <https://www.conjur.com.br/2023-jun-03/observatorio-constitucional-direitos-fundamentais-regulacao-plataformas-digitais/> accessed 15 July 2023.

80 TSE, Resolução n. 23.714, de 20 de outubro de 2022 <https://www.tse.jus.br/legislacao/compilada/res/2022/resolucao-no-23-714-de-20-de-outubro-de-2022> accessed 8 September 2023.

nized the abuse of those who disseminate disinformation on their profiles, accounts and channels on the Internet. It also stated that the democratic Institutions must take action to guarantee the isonomy and legitimacy of the electoral election[81].

In the first semester of 2023, the regulatory agenda reached another level. The debate on the aforementioned Fake News Draft Bill was reignited after the attacks on democratic institutions in Brasília, which took place on January 8th, 2023, through invasions of the STF, the National Congress and the Planalto Palace (where the Office of the President of the Republic is located). Very similar to the attack that occurred against the Capitol on January 6th, 2021, in the United States, the key difference between both events is that, differently than the Capitol Attack, in which only the Legislative Branch was invaded, in the January 8th Attack in Brasília all Branches of Power were attacked.

In addition, in mid-April there was a tragic wave of violence in schools across Brazil, especially given the inertia of platform providers in removing content posted by users that aimed to promote or encourage those acts of violence. In order to remove this type of unlawful content, the Ministry of Justice and Public Security, part of the Executive Branch, issued the Decree (*Portaria*) n. 351/2023, to preventing the dissemination of blatantly unlawful content online[82]. It, among other clauses, provides for need of social media platform providers to mitigate measures for systemic risks, including the use algorithms, as well as the need for the Public National Security Department (*Secretaria Nacional de Segurança Pública* – SENASP) to create a *hash* data-base on illegal content.

As already stated above, in this scenario, the Fake News Draft Bill, that was discussed since 2020, regained its strength[83]. Differently from its original text, that was intended to curb the spread of disinformation content, by May 2023, around 40% its text was modified to broaden its scope and propose a regulation in several aspects of the social media platforms used in

81 ADI 7261 MC-Ref (2022) STF Full Court, judgment on October 26th, (Justice Fachin) [7]–[9].

82 Ministério da Justiça e Segurança Pública, Portaria do Ministro n. 351/2023, de 12 de abril de 2023 <https://www.gov.br/mj/pt-br/centrais-de-conteudo/publicacoes/categorias-de-publicacoes/portarias/portaria-do-ministro_plataformas.pdf/view> accessed 4 May 2023.

83 See n. 24.

Brazil, but it was not further discussed openly in society and due to lack of political support was not enacted[84].

Among its many provisions, besides combining the concepts of systemic-risk assessment and the creation of duties of care, approached in the DSA and in the OSA, respectively, the main discussion regarding the Fake News Draft Bill lies in the creation of an agency for the enforcement and oversight of the statute once enacted. It shall be stated that the creation of an agency for this purpose, according to the Brazilian Federal Constitution, shall be made only by the President of the Republic and not by the Congress, that, in turn, can set forth how this agency is organized (art. 61, §1º, n. II, letters "a" and "e", Brazilian Federal Constitution)[85].

Therefore, it is worth mentioning the proposal of the Special Commission on Digital Law of the Federal Council of the Brazilian Bar Association (CFOAB). The Commission suggests the creation of a Brazilian Digital Platform Regulation System, structured and organized on a tripartite basis, through the creation of a Digital Policy Council (CPD), a deliberative body made up of people appointed by the Legislative, Executive and Judiciary branches at federal level, as well as those appointed by the Brazilian National Telecommunications Agency (ANATEL), Administrative Council for Economic Defense (CADE), Brazilian Data Protection Authority (ANPD) and the CFOAB. In addition to the CPD, the System would also include the Brazilian Internet Steering Committee (CGI.br), as the body responsible for carrying out studies and issuing recommendations, as well as Self-Regulation Entities, in charge of analysing practical cases involving content moderation on platforms.

---

84 PL das Fake News: 44% do seu texto foi alterado desde sua primeira versão em 2022 in (2023) Mobile Time <https://www.mobiletime.com.br/noticias/26/04/2023/pl-das-fake-news-44-do-seu-texto-foi-alterado-desde-sua-primeira-versao-em-2022/> accessed 23 October 2023.

85 "Art. 61
(...)
§ 1º The President of the Republic shall have exclusive power to initiate the following laws:
(…)
II - laws that deal with:
a) creation of public offices, positions or jobs in the direct administration and autarchies, or an increase in their remuneration;
(…)
e) creation and abolition of Ministries and agencies of public administration, observing the provisions of art. 84, VI" (Free translation. Brazilian Federal Constitution 1988)".

As for progress in the Judiciary, mention should be made regarding the decisions enacted by the higher courts, STF the Brazilian Superior Court of Justice (*Superior Tribunal de Justiça* – STJ).

In 2021, STJ has issued a decision in a Special Appeal (*Recurso Especial* – REsp) that upheld the liability of a social media provider to be established based on a report of a user within the platform, not applying Art. 19 MCI[86]. Concerning the reasoning, STJ stated that a platform provider that, after being notified, refuses to delete an offensive publication involving a minor, should be held liable, based on the principle of full protection of children and adolescents and from the perspective of their social vulnerability[87].

Finally, in June 2025, as already anticipated above on Section B., STF ruled Art. 19 MCI as partially unconstitutional. Between April 2014 and May 2023 (moment in which Art. 19 MCI was full into force), as a general rule, the liability of platform providers, including social media, was established only after non-compliance with a court decision (Art. 19 MCI). Exceptionally, platform providers could be held liable after the reporting of users in cases of non-consensual disclosure of intimate images (Art. 21 MCI)[88].

In general, the constitutional interpretation adopted by the STF points out to the application of the "notice and take down" approach to the liability of platform providers – beforehand only applied under the Brazilian legislation for content related to nonconsensual distribution of intimate images (NCDII) on Art. 21 MCI. Providers must demonstrate that acted diligently and in a reasonable timeframe against unlawful and unauthorized content. Art. 19 MCI continues in force if the content is related to crimes against honour (slander, defamation, and libel), i.e., the liability of platforms will continue to require a court order as already stated by the legal

86 The Brazilian Supreme Federal Court (STF) exercises both the function of abstract constitutionality control and the concrete constitutionality control of norms as the final instance, as it was originally based on the model of the US Supreme Court and has shown a growing tendency towards the European model of constitutional jurisdiction, especially since 1988 (with the current federal constitution). The Brazilian Superior Court of Justice (Superior Tribunal de Justiça – STJ) is the highest instance of ordinary jurisdiction for controlling the uniformity and authority of federal laws and can exercise the concrete control of constitutionality of norms, subject to the STF's review. In principle, STJ can be equated with the German Federal Court of Justice (BGH).

87 STJ, REsp n. 1.783.269 (2021) 4th Panel, judgment on December 14th, (Justice Ferreira) [16].

88 See n. 10.

provision before the STF's decision, but the shared content can still be removed by platform internal rules, or by user notification.

The thesis of general repercussion (*Tese de repercussão geral*) is remarkable, as it is quite extensive compared to what Art. 19 MCI aimed to address[89]. Among its 14 paragraphs (21, if sub-paragraphs are also taken separately into account), the following additional aspects deserve particular attention, in addition to those already highlighted:

(i) the presumption of liability if a piece of content is boosted/recommended by the provider or through the use of illegal artificial distribution of content through robots;
(ii) the adoption of duties of care (*deveres de cuidado*, in Portuguese) by providers if a systemic risk is identified, as well as the adoption "additional duties", such as the implementation of complaint/report channels, publication of internal rules and transparency reports, including the obligation to point out a legal representative in Brazil;
(iii) email services (e.g., Gmail), instant messaging services (e.g., WhatsApp) and closed meetings platforms (e.g., Zoom) are considered as "neutral providers", i.e., STF considers that that those providers do not interfere with content, as long as it concerns interpersonal communications[90].

The thesis of general repercussion set forth by the STF is applied to cases after June 26th, 2025. Besides, STF called upon the Brazilian National Congress "to elaborate a legislation that is able to remedy the shortcomings of the current regulatory scenario with regard to the protection of fundamental rights"[91].

By taking the lead in regulating the matter, STF made clear that judicial intervention (though exceptional!) was necessary due to the congressional inertia. The expectation, now, is that the Brazilian National Congress

89 STF, Informação à Sociedade, RE 1.037.396 (Tema 987) e 1.057.258 (Tema 533) - Responsabilidade de plataformas digitais por conteúdo de terceiros, 27 de junho de 2025 <https://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/Informac807a771oa768SociedadeArt19MCI_vRev.pdf> accessed 28 June 2025.

90 At the time of finalizing this text, the full Justices' opinions had not yet been made publicly available. Nevertheless, based on the trial report (n. 89) and the general repercussion thesis, which have already been published, it is possible to infer that, despite the thesis's broad normative reach, *instant messaging groups* continue to occupy a regulatory gray area. This persisting regulatory gap underscores the need for targeted legislative action by the Brazilian National Congress on the matter.

91 Free translation, see (11), item 13 of the thesis of general repercussion.

will respond accordingly, taking up the task of enacting comprehensive and democratically debated legislation to ensure a more consistent and legitimate framework for the protection of fundamental rights in the digital environment.

E. Final remarks

From an evolutionary perspective, it can be seen that the regulation of social media platforms is still far from a consensus, especially due to the complexity of the matter and the substantial change in the way digital platforms operate. In fact, the models for management and regulation – self-regulation, external regulation, and hybrid regulation – apply simultaneously in the context of social networks, contributing to the multiplicity of proposals for changing the regulatory landscape currently shaped.

Numerous regulatory models have emerged to reconcile the protection of users compatible with the promotion of innovation. These models illustrate the tensions that exist in this context, namely between users, platforms, states, civil society organizations, and supranational bodies, such as the European Union. In addition to the debate on Internet governance, the discussion on the regulation of social media platforms also includes doubts about the legitimacy of the actions adopted by providers and the resulting impact on fundamental rights.

It is clear that the dynamic nature of the Internet makes it particularly conducive to self-regulation, which, however, as we have seen, cannot rule out the possibility of external or hybrid regulation without creating deficits in the protection of rights online. Although there are various proposals for regulating online environments, it is argued that the most appropriate position is the one based on cooperation and interaction between entities with the capacity to manage digital environments, i.e., regulated self-regulation of social media platforms.

Regarding the Brazilian scenario, despite recent developments in the regulatory framework, we are still far from a regulated self-regulation of social media platforms. Recently the debate has reached a new level as the discussion has been guided by a clearly emotional and even passionate dimension, involving the persistence of polarization in society. Social pressure put on the National Congress, but also on the Executive Branch and the Judiciary, has been intense. What is certain is that the regulatory debate has already gone beyond the definition of the civil liability of platform

providers, and is now discussing matters regarding transparency, oversight agency, mitigation of systemic risks and the creation of duties of care.

The recent declaration of partial (and progressive) unconstitutionality of Article 19 MCI by the STF reveals the elevation of the debate on the responsibility of digital platforms to a new level of regulatory density and constitutional relevance, in contrast to the persistent omission of the Brazilian National Congress to comprehensively regulate the matter. As stated beforehand, the thesis of general repercussion established in the judgment goes beyond the limits of civil liability for third-party content, traditionally attributed to providers in Brazil through Art. 19 MCI, and inaugurates a broader understanding, which incorporates duties of care, assumptions of liability, and normative parameters applicable, not only liability of providers due to content generated by their users, but also liability due to the platforms' own actions and omissions.

The STF ruling, thus, has an evident regulatory vocation, which seeks to fulfil, albeit provisionally, the existing legislative void as Art. 19 MCI no longer no curb the challenges related to liability of the providers in isolation in the digital realm. The breadth of the thesis highlights not only the centrality of the issue in the current stage of digital constitutionalism, but also the urgency of legislative action that consolidates, with democratic legitimacy, the parameters of platform regulation in Brazil.

Regardless of what is yet to come, it is clear that situations of social upheaval tend to speed up discussions in Brazil, and this has been no different when it comes to regulating digital platforms. Although this phenomenon drives the search for solutions, it also tends to lead to a problematic reduction of complexity, a deficit in democratic-deliberative legitimacy, contradictions and regulatory gaps, as well as the undermining the multi-stakeholder nature of the debate, among other worrying aspects.