Revista Percurso

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

FREEDOM OF EXPRESSION AND CONTENT MODERATION: AN ANALYSIS FROM THE "BRAZILIAN GENERAL SPEECH" CASE

LIBERDADE DE EXPRESSÃO E MODERAÇÃO DE CONTEÚDOS: UMA ANÁLISE A PARTIR DO CASO "DISCURSO DE GENERAL BRASILEIRO"

FERNANDO RISTER DE SOUSA LIMA

Professor in the Postgraduate Program in Political and Economic Law at Universidade Presbiteriana Mackenzie (930, Consolação St., 41, São Paulo/SP – Brazil, 01302-000) and in the Department of General Theory of Law at Pontifícia Universidade Católica de São Paulo. Address: 930, Consolação St., 41, São Paulo/SP - Brazil, 01302-000. Lattes: http://lattes.cnpq.br/0125281695257124; e-mail: frsl.sociologyoflaw@gmail.com

CAMILLA TAVARES DE ALBUQUERQUE

PhD student in Political and Economic Law at Universidade Presbiteriana Mackenzie (930, Consolação St., 41, São Paulo/SP – Brazil, 01302-000). Master's degree in Information Society Law (FMU). Major in Law (UFPE). Researcher. Professor (FMU). Police Chief of São Paulo State. http://lattes.cnpq.br/9570276997068133; e-mail: albuquerque.camilla33@gmail.com

ABSTRACT

The aim of this article is to offer the reader a perspective on freedom of expression from the Brazilian political scenario, with the case of the "Brazilian General Speech" (2003-001-FB-UA) as a background: a Facebook post with content that called for the siege of Brasília, which in fact happened a few days later. The case reached Facebook's Oversight Board, which decided that the content should be removed. To answer the question of possible violation of freedom of expression, three

Revista Percurso

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

fundamental hypotheses are offered: a) the publication violates freedom of expression, here considered both EU rules and soft or hard law; b) the publication does not violate any rule applicable to freedom of expression, or c) the publication has no relation to the right to freedom of expression. The results were achieved following content-based North-American doctrine, an epistemological choice that, perhaps, may help us contribute to the discussion around platform regulation. The research object was analyzed empirically and confronted with the theoretical study. Methodologically, the legal-exploratory line was adopted, using a combination of inductive-deductive reasoning, given the limitations of both. In terms of technique, documentary review and bibliographical research were particularly employed. In conclusion, no evidence was found that the post to which the "Brazilian General Speech" case refers could be included in the sphere of protection of the right of freedom of speech.

Keywords: Freedom of Expression; Meta's Oversight Board; Content Moderation; Platform regulation; Bill of Law nº 2630/00; Content-based regulation.

RESUMO

O objetivo deste artigo é oferecer ao leitor uma perspectiva sobre a liberdade de expressão a partir do cenário político brasileiro, usando como plano de fundo o caso do "Discurso de General Brasileiro": uma postagem no Facebook com conteúdo em que se incentivava sitiar Brasília, o que de fato aconteceu alguns dias depois. O caso chegou ao Conselho de Supervisão do Facebook, que decidiu que o conteúdo deveria ser removido. Para responder à pergunta acerca de eventual violação à liberdade de expressão, são oferecidas três hipóteses fundamentais: a) a publicação viola a liberdade de expressão, aqui consideradas tanto as normas comunitárias quanto a soft ou hard law; b) a publicação não viola nenhuma norma aplicável à liberdade de expressão, ou c) a publicação não tem relação com o direito à liberdade de expressão. Os resultados foram alcançados a partir da doutrina norte-americana, uma escolha que talvez possa contribuir para a discussão acerca da regulação das plataformas. O objeto de pesquisa foi analisado empiricamente e confrontado com o estudo teórico. Metodologicamente, foi adotada a linha jurídico-exploratória,

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

utilizando-se a combinação do raciocínio indutivo-dedutivo, dadas as limitações de ambos. Em se tratando de técnica, foram empregados, especialmente, levantamento documental e pesquisa bibliográfica. Em conclusão, não foi encontrada qualquer evidência de que o cargo a que se refere o processo "Brazilian General Speech" pudesse ser incluído na esfera de proteção do direito à liberdade de expressão.

***Palavras-chave**: Liberdade de expressão; Comitê de Supervisão da Meta; Projeto de Lei nº 2630/00; Regulação de plataformas; Regulação com base no conteúdo.*

1 INTRODUCTION

The central question of this paper is: does the subject matter of the case “Brazilian General Speech”, i.e. a post with content in which a general claims to besiege Brasília, violate the right to freedom of expression? The hypotheses seem simply dichotomous; however, it does not beat to conclude the affirmative or negative answer of our working question. Therefore, we offer three fundamental hypotheses: a) the publication violates freedom of expression, here considered both community standards and soft or hard law; b) the publication does not violate any applicable standard on freedom of expression, or c) the publication has no relation with the right to freedom of expression. All these possibilities will all be assessed during the review.

Our aim is to offer the reader a perspective on freedom of expression from the Brazilian political scenario, especially in view of the events of January 8, 2023, in which the *Praça dos Três Poderes* in Brasília was besieged. More specifically, our aim is to map the debate on platform regulation in content moderation and raise some relevant theoretical questions about the issue.

Recognizing the limitations of a single paper, a selection of what we consider the most relevant normative and doctrinal points for the treatment of the theme was made. So, for our purposes, the case of the “Brazilian General Speech” reached by the Oversight Board will be examined. Additionally, the Bill of Law 2630, some international standards of freedom of expression and the North-American construction about the interpretation of this right, mostly about the paradigm of

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

neutral-based and content-based moderation and the progress of this theory, will also be examined.

The research object, the case reached by the Board, was analyzed empirically and, then, confronted with the theoretical study. The legal-exploratory methodological line was used, with a combination of inductive-deductive reasoning, given the limitations of both. Technically, documentary review and bibliographical research were particularly employed.

It is expected that the study will bring some explanations about how the discussion around the regulation of content moderation on platforms has been developing and, perhaps, offer some possibilities for such a task, preventing violent acts such as the one on January 8. These achievements may justify our efforts in this review.

Firstly, brief notes of a Brazilian overview were done, presenting the "Brazilian General Speech" case and its final decision. Afterwards, it was essential to draw an overview of the protection of the right to freedom of expression, observing the international criteria applicable to the interpretation of the case. In a third moment, the efforts focused on the task of describing the dynamics of the regulation of content moderation by platforms in Brazil, giving special emphasis to the controversial points. Finally, the Brazilian case was analyzed in the light of the North American doctrine, which allowed a final step, consisting of gathering the results collected in order to extract fruitful conclusions from them, which may be useful for the debate about the regulation of platforms.

In conclusion, we have found no evidence that the post to which the "Brazilian General Speech" case refers could be included in the sphere of protection of the right of freedom of speech. Even applying standards of interpretation from the US, no other answer could be found.

2 BRAZILIAN OVERVIEW

The Brazilian political background has seen an escalation of polarization during the last few years. The ultimate presidential elections were permeated by high

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

degrees of violent speeches and every kind of threat. As expected, the result of run-off elections, with tiny differences in the percentage of votes, created an environment of uncertainty for political transition activities.

On January 3, 2023, a few days after the investiture of the new president, we saw the invasion of the *Praça dos Três Poderes* in Brasília and the most representative buildings, the palaces of executive, legislative and judiciary powers. The occupation was followed by all kinds of violations, such as deliberate damages and destruction of historical pieces. The most surprising thing was that the action was not just broadcasted live by the press, but also registered and spread by those who were supposedly protesting, through lives, photographs, videos, posts, reels etc. A register of an act that will go down in history just because of the violence perpetrated. Although our purpose is not to evaluate these from a partial perspective, there is an undeniable consensus: violence, especially when perpetrated against instruments so dear to democracy cannot be tolerated or, at the very least, evaluated as something good.

Considering that, our objective is to raise some relevant questions that might help readers understand why and how some kind of action happens and to map out ways to prevent democracy threats like these. Of course, such a large-scale occupation cannot be carried out without a minimum of planning and unity of purpose. How did this come out?

There is evidence to suggest that there is some connection between the January 8 event and the way in which the discourse of regulation of platforms is developing at the present time, mainly when the topic of content moderation is raised. In other words, it is necessary to talk about freedom of expression and its boundaries in a democratic space.

The reason for this first conclusion lies on the fact that, just two days after the investiture of the new president, a Facebook user made a post in which he calls to besiege Brazil's Congress and shows an uniformed general, with the words "hit the streets" and "go to the National Congress… [and the] Supreme Court" (OVERSIGHT BOARD, 2023, p. 2). In addition, it shows a video with the following caption: "Come to Brasília! Let's storm it! Let's besiege the three powers" (OVERSIGHT BOARD, 2023, p. 2).

Revista Percurso

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

On the same day of the post, on January 3, 2023, another Facebook user reported the content on the grounds that it violated Meta policies on violence and incitement. Six similar reports were received between January 3 and 4, but all of them were ineffective, because the platform moderation considered the content to not violate its own rules. So, after the Brazilian riot, on January 9, Meta issued a statement pledging to remove posts supporting it, as well as framing Brazil as a high-risk location given the repercussions of recent events. Therefore, on January 20, the case called "Brazilian General Speech" was selected by the Oversight Board,[1] whose role would be to decide about Facebook content moderation. Just after the selection, Facebook admitted that it was a wrong decision to keep the post available and finally removed it.

The Board announced the case and established the major points it would appreciate:

> The Board would appreciate public comments that address:● The political situation in Brazil in advance of October's election and how it shifted between October 2022 and January 8, 2023. ● The relationship between political violence, election denialism, and calls for offline mobilization on social media. ● When Meta's election integrity efforts should begin and end, and what criteria should guide decisions about those timeframes, particularly as they relate to transitions of power. ● How Meta should distinguish between legitimate political organizing and harmful coordinated action. ● How Meta should treat content attacking or delegitimizing democratic institutions and processes. (OVERSIGHT BOARD, 2023a, p. 2)

The case received 19 public comments all around the world and the first Oversight Board notes were quite incisive, saying that "Meta's initial decision to leave this content up during a time of heightened political violence represented a clear departure from its own rules" (OVERSIGHT BOARD, 2023, p. 3). Moreover, it was stated that Meta "failed to escalate [content] for further review" (OVERSIGHT BOARD, 2023, p. 3) and was unable of recording and systematizing data election-related specifically, because, in the platform's own words, "it does not adopt any particular metrics for measuring the success of its election integrity efforts generally"

[1] The Oversight Board - which today operates on all of Meta's networks, i.e. both Facebook and Instagram - has a very similar structure to that of a court, bringing a commission of "lay" judges together, whose function is to decide on the removal or maintenance of content in paradigmatic cases and to offer opinions with recommendations addressed to the platform.

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

(OVERSIGHT BOARD, 2023, p. 3). So, the Board decided to overturn Meta's original decision to leave up the aforementioned content.

The Board's analysis raised some problematic points in the conduct of Meta. Firstly, despite content rules being clear enough about violence and incitement, the two first degrees of content moderators neither found the violation, nor were they even able to escalate the post for further revision. Secondly, policies about high risk location – a category in which both Brasília and the three power's buildings were already included at the time for Facebook standards – didn't seem to be satisfactory, specially because the platform overlooked the political context in Brazil and didn't consider the corresponding increased risk.

In addition, there was no data on specific claims, nor even to measure the effectiveness of the efforts to safeguard the integrity of the 2022 Brazil elections. Even though there was sufficient data, the board pointed out the lack of an effective strategy, given that "the review and potential removal of individual pieces of content from Meta's platforms is insufficient and relatively ineffective when such content is part of an organized and coordinated action aimed at disrupting democratic processes." (OVERSIGHT BOARD, 2023, p. 19).

Last but not least, in light of all this, the Board clearly stated that Meta's decision was not in line with its human rights responsibilities:

Given the above, the Board finds that the removal of the content is consistent with its human rights responsibilities. Removing the content is a necessary and proportionate response to protect the right to life of people, including public officials, and public order in Brazil. The removal of this similar pieces of content is also necessary and proportionate to protect Brazilian's right to vote and participate in public affairs, in a context where attempts to undetermined a democratic transition of power were underway. (OVERSIGHT BOARD, 2023 p. 19)

On the substance, it can be seen that the Board's verdict revolves around the right to freedom of expression[2]. Not by chance, the normative framework on which it was based consisted, precisely, of international standards contained in the Covenant

[2] About North American context, Evelyn Douek says: "…most discourse about it [systems and processes of content moderation] remains stuck in the register of individual ex post constitutional rights litigation. Perhaps this should be unsurprising: content moderation decisions are decisions about *speech* after all. And speech, of course, is special" (2022, p. 31-32).

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

on Civil and Political Rights, the General Comment nº 34 and, pointing out the most relevant ones, the Rabat Plan of Action. The Board approaches the issue in a topic called “Compliance with Meta’s human rights responsibilities” and exposes its concern about the harmful potential of the company's conduct. In fact, an online harm led to an offline one: the *Praça dos Três Poderes* was besieged just five days after the post that originated the “Brazilian General Speech” case was published.

The case is usually compared to the Capitol insurrection in the United States, notably because of the common election background and the impacts of online platforms. It is worth remembering that it was just after that attack on the Capitol that Facebook suspended the North American president’s account, precisely because of his support for the event. Moreover, the case was also selected by the Oversight Board (2021-001-FB-FBR), whose decision “has upheld Facebook’s decision on January 7, 2021, to restrict then-President Donald Trump’s access to posting content on his Facebook page and Instagram account” (2021, p. 1), also considering that

However, it was not appropriate for Facebook to impose the indeterminate and standardless penalty of indefinite suspension. Facebook’s normal penalties include removing the violating content, imposing a time-bound period of suspension, or permanently disabling the page and account. (OVERSIGHT BOARD, 2021, p. 1).

At the Capitol episode, despite the Board showing disagreement on some points, it was vigorous in exposing its concerns about the risk of violence and harm, which justified the restrictions imposed for the United States then-President:

> The Board found that, in maintaining an unfounded narrative of electoral fraud and persistent calls to action, Mr. Trump created an environment where a serious risk of violence was possible. At the time of Mr. Trump’s posts, there was a clear, immediate risk of harm and his words of support for those involved in the riots legitimized their violent actions. As president, Mr. Trump had a high level of influence. The reach of his posts was large, with 35 million followers on Facebook and 24 million on Instagram. (OVERSIGHT BOARD, 2021, p. 3)

By the way, as it could be seen, the reasoning shares similarities with the “Brazilian General Speech” case (2023-001-FB-UA), mostly because of the electoral context and the fact that the online behavior was able to cause concrete violence on

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

the offline field. As in the Brazilian case, the Capitol has been the subject of wide debate and fierce disagreement. People were divided between a more liberal position, claiming a violation of the First Amendment by unduly restricting freedom of expression, and those who, under the same legal argument, claimed that the restriction was proper, if not overly permissive:

> Many critics among the political right argue that social media platforms are abusing their power and removing an egregious amount of conservative-based material. Claims of election fraud following the 2020 Presidential Election led Twitter and Facebook, among others, to flag posts making such accusations and include disclaimers warning other users of potentially false information. Content associated with or spreading dangerous theories from the popular conspiracy group, QAnon have been targeted and removed from most social media platforms following the election and subsequent insurrection of the U.S. Capitol. Among many conservatives, these actions by social media platforms feel like an invasion of their beliefs and an infringement on their First Amendment rights. (…) On the opposite end of the spectrum, many believe that social media companies are not removing nearly enough content. While some people may applaud social media platforms for taking a stand and flagging or removing posts regarding things such as election fraud or QAnon conspiracies, those same people blame social media for allowing those users and their messages to flourish in the first place. However, somewhat surprisingly, both sides agree that change needs to occur—whether that includes modifying platform content policies or an interventional mandate by federal law. (YOUNG, 2021, p. 4-5)

In Brazil, the controversy was even greater, if not of the same magnitude, with hostilities being even more evident. Even after the elections, left and right continued to fight each other and the electorate still showed no signs of reconciliation. Even in a quiet political context, the question of the limits of freedom of expression is a tormenting one and, considering the dichotomous positions in the country at the time, there was no expectation of a consensus. However, although there is no position that can be said to be peaceful today, a few observations can be useful for the development of this review.

If, on one hand, the phenomenon cannot be reduced to a generic causal analysis, on the other there is no doubt that mobilization through social networks played a relevant role in the success of the January 8 events. The question, then, arises as to why Meta considered the content referred to fall within the scope of freedom of expression in the “Brazilian General Speech” case. To this end, a brief

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

overview of what is considered freedom of expression is necessary, as will be done below.

3 WHAT DOES FREEDOM OF EXPRESSION MEAN?

Any reader, no matter how alien to the legal field, does not risk denying that concerns about freedom of expression are neither new, nor a simple dilemma. In a globalized world, since, at least, the Universal Declaration of Human Rights, there have been substantial efforts to regulate rights and agree on standards to promote and guarantee the value. Legal attempts to achieve this goal in the American Continent – which are most relevant to this research – go through International Covenant on Civil and Political Rights, American Declaration of the Rights and Duties of Man, American Convention of Human Rights till Rabat Plan of Action, aside from domestic laws.

Rules are numerous and, certainly, it would be counterproductive – or, worse, tiring to the reader – to appreciate all of them in their singularities. Despite that, relevant questions to this research have been discussed in some of them, so it is not possible to escape their assessment.

One of the most important rules[3] about the issue is contained in articles 19 and 20[4] from International Covenant on Civil and Political Rights (ICCPR). Specifically, article 19 establishes standards to comprehend freedom of expression and indicates methods which can help the task of protecting them in concrete terms. Briefly, the article states that the right to freedom of expression "shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers,

[3] In the context of the Americas, it is worth mentioning furthermore: American Convention on Human Rights (art. 13); American Declaration of the Rights and Duties of Man (art. IV); Inter-American Democratic Charter (art. 4), and the Office of the Special Rapporteur for Freedom of Expression.

[4] "Article 19 1. Everyone shall have the right to hold opinions without interference. 2. Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice. 3. The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary: (a) For respect of the rights or reputations of others; (b) For the protection of national security or of public order (ordre public), or of public health or morals. Article 20 1. Any propaganda for war shall be prohibited by law. 2. Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law" (UNITED NATIONS, 1966, s.p.).

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

either orally, in writing or in print, in the form of art, or through any other media of his choice." Then, it continues with the requirements to restrict the right, often referred to as a "three-part test", covering legality, legitimacy and necessity and proportionality. This third demand is linked to "respect of the rights or reputations of others" and "the protection of national security or of public order (ordre public), or of public health or morals".

Likewise, in 2013, the Rabat Plan of Action, the annual report of United Nations High Commissioner for Human Rights refers to articles 19 and 20 of ICCPR, ratifying and detailing its requirements and distinguishing three kinds of freedom of expression in a useful ranking:

> […] expression that constitutes a criminal offence; expression that is not criminally punishable, but may justify a civil suit or administrative sanctions; expression that does not give rise to criminal, civil or administrative sanctions, but still raises concern in terms of tolerance, civility and respect for the rights of others. (2013, p. 9)

Moreover, the plan added some needs of the three-part test, expressively corroborated, including that "restrictions are clearly and narrowly defined and respond to a pressing social need; (…) the least intrusive measure available; (…) not overly broad (…); and proportionate so that benefit to protect interest outweighs the harm to freedom of expression" (UNITED NATIONS, 2013, p. 9).

The Rabat Plan establishes that States should use the guidance provided by the General Comment 34 (UNITED NATION, 2013, p. 10) of the nineteenth section of the High Commissioner, which details the ways to apply the articles 19 and 20 of the ICCPR. It divided freedom of expression in two parts: freedom of opinion, which admits no restriction, and freedom of speech (UNITED NATIONS, 2011, p. 2-3), that could be restricted, unless these do not put in jeopardy the right itself (UNITED NATIOS, 2011, p. 5).

In addition to these, it would be possible to mention a countless number of other standards and protocols that explore freedom of expression. It is more important, however, to investigate how the parameters in question can be useful in understanding the case presented.

Despite the merits of the attempts to systematize the application of freedom of expression, rights like this are unavoidably abstract. The numerous guidelines put

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

in place do not cope with the complexity of concrete cases, especially when there are other rights on the table. The task of "getting the right measure" of freedom of expression can be a demanding one. To say that it is a relative right and that it must be balanced, eventually giving up space to other rights, is just to insist on the commonplace. Actually, the abstraction of relative rights says very little about the decision-making practice, which could have in it only a starting point. In this regard, Ronaldo P. Macedo Junior believes that "the use of balancing techniques has been naturalized in Brazilian doctrine and practice" (2017, p. 282), so "jurists who most often disagree about how to balance or about evaluation of the rights, and decide on a case-by-case basis" (MACEDO JUNIOR, 2017, p. 282).

Not even the guidelines of the paradigmatic international concrete cases seem to lead up to a satisfactory argumentation. In this sense, Macedo Junior points to a lack of in-depth study of the right to freedom of expression in Brazil, especially because of the lack of philosophical grounding:

> There is a necessary connection between freedom of speech and its philosophical foundations. Political philosophy is clearly rooted in American free speech scholarship. This is not the case in Brazil where a philosophical treatment of the subject – especially in terms of the presupposed theory of justice embedded in it – is usually absent in most of legal practice and constitutional doctrine (legal dogmatism). Brazilian mainstream doctrine tend to accept as an almost sufficient (and sometimes exclusive) starting point the unavoidable and irreconcilable conflicting positive rights and values stated in the constitution. (2017, p. 284).

For the author, the American experience shows that "any legal analysis of free speech should start with a theory of what this expression should mean" (2002, p. 284). Although this fact seems to be obvious, a few countries take the theorization of freedom of expression as seriously as the United States (US), a country from which, according to Macedo Junior, Brazilian decision-making tradition has a lot to learn (MACEDO JUNIOR, 2002, P. 284).

In a similar critique, Evelyn Douek approaches the theoretical debate on the concept of freedom of expression from another angle. The researcher brings to light the problem arising from the fact that the Board intends to pacify a single benchmark of freedom of expression, which, given the global scale of the platforms, would be

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

unreasonable (2022, p. 569). Despite the relocation of the problem, Douek seems to reach similar conclusions to those already set out: although it was created in order to improve and give legitimacy to the moderation of the platform (DOUEK, 2019, p. 46), the Board's decisions would thus tend to disregard fundamental cultural and contextual issues, which would lead to a conflict between context sensitivity and legal consistency (DOEUK, 2022, p. 569).

These points suggest that freedom of expression is not a matter of laws, but mostly of enforcement. Indeed, this remains a problem in various legal cultures and contexts. Even though there is no doubt about the indispensability of building a well-developed and predictable enough doctrine, Marin Scordato and Paula Monopoli argue that, even in the US, the First Amendment fills very few interpretative gaps (2002, p. 188):

> "…almost no scholar who was seriously examined First Amendment jurisprudence has ultimately concluded that it is currently characterized by clarity, coherence or predictability. One scholar after another, in one article after another, decries unruly nature of free speech doctrine.".

Despite acknowledging deficiencies in the law, as Macedo Junior argues, it is necessary to accept that there is still a lot to be learned from North-American doctrine on the right to freedom of expression (2017, p. 284). With the First Amendment as just a starting point, the right has been subject of not only disagreements, but also of fruitful juridical discussions.

Before delving into the topic, it is pertinent to assess how the legal scenario for the regulation of content moderation has developed in Brazil. The legislative attempts to regulate platforms result precisely in the interpretation and conception of the right to freedom of expression, which leads us to the next topic.

4 RULES OF CONTENT MODERATION

Approximately three years after the proposal of the Bill of Law nº 2630 (*Projeto de Lei nº 2630*) establishing the Brazilian Law on Internet Freedom, Accountability and Transparency, commonly called the "Fake News Bill" or simply "PL 2630", the procedure strongly returned to the political agenda in early 2023. Along with it, the General Repercussion Themes nº 533 and 987, in progress since

Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

2017 at the Federal Supreme Court, were finally unearthed. Both of them approach the topic of responsibility of platforms' liability for third-party content and have regained ground in view of the presidential election in 2022, mostly after the January 8, 2023 event.

In particular, with regard to PL 2630, there was a decisive influence of the European *Digital Services Act* (DSA), which ended up inspiring much of the textual changes promoted so far. According to Kenneth A. Bamberger and Deirdre K. Mulligan, the DSA creates an asymmetric model of regulation in order to improve the applying of principle of proportionality (2021, p. 56).

> Generally speaking, the DSA makes three contributions to the regulatory framework for content moderation by: (1) setting out general rules of liability for providers of intermediary services; (2) establishing a regime of due diligence obligations, with a special focus on online platforms including social media; and (3) strengthening the cooperation between national authorities in charge of the public enforcement of online regulation. (BAMBERGER, MULLIGAN, 2021, p. 53)

Despite the similarities, Brazilian law grew up in a particular environment. The undeniable mess of the political game during months (or years) before the scrutiny betrays a high risk of violence acts. Nevertheless, people apparently did not foresee something like the January 8 event. The addition of what happened to the problem of electoral fake news and irregular propaganda – both of them with large impact of platforms – created an even more unstable environment, in which people demanded that public institutions take vigorous action.

Moreover, the recent attacks on schools, which were planned by children on social networks, were one more reason to hurry. By the way, as an "emergency response" for these events, on April 12, 2023, the Ministry of Justice and Public Safety, an Executive body, published the Ordinance nº 351/2023, that provides for administrative measures to be adopted for the purpose of preventing the dissemination of blatantly unlawful, harmful or damaging content by social networking platforms. This standard anticipates some of the concepts specific to PL 2630 and, more often than not, from the DSA. These include systemic risks, duty of care and crisis protocols, specifically targeting "illegal, harmful and damaging content on social media platforms referring to violent extremism that encourages attacks on the school

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

environment or condones and incites such crimes or their perpetrators" (BRASIL, 2023, s.p.).

All of this led up to a race to platform regulation, resulting in the renewal of the PL 2630 agenda. The bill was completely disfigured from the original and dozens of other law procedures with different issues were attached to it. The more themes were included in the bill, the bigger the controversies became. In addition, during the procedure, the platforms demonstrated their displeasure with the terms placed in the project, even protesting in the form of hostile warnings. The companies insinuated that the new legislative proposal would end freedom of expression and create a scenario of censorship.

Some points of PL 2630 should be highlighted in order to provide a brief overview of the most controversial issues on freedom of speech. The bill stated a list of situations (mostly criminal) that give rise to the characterization of the "platform's duty of care" (art. 11). Any damage caused by non-compliance with duty of care obligations may lead to joint and several civil liabilities of application providers (Art. 6º). In this sense, there seems to be a few concerns, especially because platforms used to include duties referring to those criminal themes in their community standards. On the other hand, the liability could be damaged if it depends on a peremptory criminal subsumption.

Section II deals with systemic risks, briefly stating that providers should identify them and take proportional measures to mitigate them. Paragraph 1º of article 8 states that, when such measures involve the use of computerized systems, "they shall include safeguards that are appropriate and effective, in particular through human supervision with a view to ensuring foresight, proportionality and non-discrimination that is unlawful or abusive". Nowadays, most of the regulatory activities of the platforms take place through computerized systems without the need for human intervention, which is justified by the volume of activities. If this circumstance is analyzed with the obligation of algorithmic "accountability" for transparency, it is necessary to take into account possible losses for the business activity, depending on how much the provider's duties are extended.

In this regard, platforms' concerns also turn to transparency requirements, which must not be such as to frustrate competitive gains. There are trade and

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

industrial secrets that must be secured (which is even provided for in Art. 21), but there is no consensus on what they are and to what extent they can be exposed. Measures of transparency and opacity are still a big controversy.

It should also be noted that the risks also served as a criterion for measuring the obligations regarding the duties of care (art. 11, §2º, II). Non-compliance with duty of care obligations may also give rise to the establishment of a security protocol (art. 12) that must be strictly observed by the provider, under penalty of incursion into another hypothesis of joint and several civil liability (art. 6º).

For better or for worse, the process, still at the House of Representatives, was suspended after the withdrawal of the agenda by the rapporteur on May 2, 2023. Much of the press reported it as a move to prevent the government's defeat. The rapporteur's argument, however, was reasonable in the sense that the various proposals about the PL 2630 were not sufficiently evaluated and debated.

Another possible explanation for the suspension lies on the General Repercussion Comments nº. 533 and 987 which, in summary, address the topic of platforms' liability for third-party content. The theme is, precisely, the constitutionalism of the art. 19 from Law nº 12.965/14 (*Marco Civil da Internet*, MCI), a standard that establishes something like an internet civil framework and, among other topics, deals with the responsibility of the server for third-party content. Despite the fact that the PL 2630 expressed the exceptional nature of its terms in relation with the MCI (art. 55[5]), it is clear that the position of the Constitutional Court will impact the acceptance and interpretation of the bill.

[5] According to André Faustino e Jorge Shiguemitsu Fujita, “There are two types of content removal systems from Internet application providers, the first of which is the system known as Notice and Takedown (NAT) and the second system is through judicial determination. The NAT system is one in which the person interested in removing the infringing content makes the request directly to the application provider and this one, making an analysis of the pertinence and adequacy, performs the removal of the requested content, this provision is through existing legislation or through judicial precedents that allow such legal possibility, the judicial determination system is one in which the interested person has to resort to the Judiciary which, by specific court order, determines the application provider to remove the infringing content.”. Original: “Existem dois tipos de sistemas de remoção de conteúdo dos provedores de aplicação da internet, o primeiro deles é o sistema conhecido como *Notice and Takedown* (NAT) e o segundo sistema é por meio de determinação judicial. O sistema NAT é aquele em que a própria pessoa interessada na remoção do conteúdo infringente faz o pedido diretamente ao provedor de aplicações e esse fazendo uma análise da pertinência e adequação, realiza a remoção do conteúdo solicitado, essa previsão é por meio de uma legislação existente ou por meio de precedentes judiciais que permitam tal possibilidade jurídica, já o sistema de determinação judicial é aquele em que a pessoa interessada tem que recorrer ao Poder Judiciário que, por ordem judicial específica, determina ao provedor de aplicações a remoção do

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

The crucial issue about MCI is the immunity offered to the server, assigning precedence for freedom of expression, as exposed by Chiara Teffé and Maria Celina Bodin Moraes (2017, p. 113-114):

> Several interpreters maintain that the MCI legislator would have placed freedom of expression in a preferential position over other rights, due to certain options in the wording of the law, notably the mentions of the right to freedom of expression. In fact, the Marco Civil valued freedom of expression, and this legislative option is in line with recent positions of the Supreme Court. However, this does not mean that the interpreter should attribute to freedom of expression the condition of an absolute right, immune to any limit, nor even that he should establish a kind of prior hierarchy between constitutional norms.[6]

In this sense, art. 19 removes the "notice and take down"[7] system for content removal, forcing the user to resort to a court order. Some scholars support the user's hyposufficiency as one of the arguments against the judicial system that, perhaps, causes an excessive burden on the less privileged party of the legal relationship in question:

> In Brazil, article 19 of the Brazilian Civil Rights Framework for the Internet established the need for a judicial subpoena for the service provider to remove inappropriate material from the air. This requirement ended up overly burdening the user, because, as already discussed, he is a vulnerable and hyposufficient party in the relationship with the provider. The vast majority of relationships occurring on the Internet are governed by the Consumer Protection Code, given that the user always uses the network through providers and as an end user. Consumer vulnerability was already recognized in the offline world and, in the online world, it is also present and, as Martins reports, is treated as informational vulnerability. (CAVALCANTI; LEITE; BARRETO JÚNIOR, 2018, p. 522).

conteúdo infringente" (FAUSTINO; FUJITA, 2017, p. 822).

[6] Free translation of "Diversos intérpretes sustentam que o legislador do MCI teria colocado a liberdade de expressão em posição preferencial frente aos demais direitos, em virtude de determinadas opções da redação da lei, notadamente às menções ao direito à liberdade de expressão. De fato, o Marco Civil realizou uma valorização da liberdade de expressão, estando tal opção legislativa de acordo com recentes posicionamentos do Supremo Tribunal Federal. Todavia, isso não significa que o intérprete deve atribuir à liberdade de expressão a condição de direito absoluto, imune a qualquer limite, nem mesmo que deva estabelecer uma espécie de hierarquia prévia entre as normais constitucionais." (TEFFÉ; MORAES, 2017, p. 113-114).

[7] Original: "No Brasil, o artigo 19 do Marco Civil da Internet determinou como exigência a necessidade da intimação judicial para que o provedor de serviços retire do ar material inapropriado. Essa exigência acabou por onerar sobremaneira o usuário, pois, como já discorrido, ele é parte vulnerável e hipossuficiente na relação com o provedor. A grande maioria das relações ocorridas na internet é regrada pelo Código de Defesa do Consumidor, haja vista que o usuário sempre utiliza a rede através de provedores e como usuário final. A vulnerabilidade do consumidor já era reconhecida no mundo off-line e, no mundo on-line, ela também está presente e, como relata Martins, é tratada como vulnerabilidade informacional." (CAVALCANTI; LEITE; BARRETO JÚNIOR, 2018, p. 522)

Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

At a different perspective, Caio Miachon Tenorio and Thays Bertoncini argue that “the solution found by the *Marco Civil da Internet* [MCI] in its article 19 represented, at the time, a legal, political and legislative consensus on ensuring freedom of expression and preventing censorship, creating a safe harbor for digital platforms” (2023, p. 295). In any case, the objective here is not to go into the question of the constitutionality of that article, not only for the sake of space, but above all because there is no procedural dependence between the rules in question. For now, it is enough to know that the law only adds more wood to the fire already lit.

Briefly, we analyzed the Brazilian case brought before the Board whose controversial point is the right to freedom of expression; then, we went on to a brief digression about the right, especially in the Americas. This was followed by an understanding of how tortuous the development of content moderation regulation in Brazil has been. So far, there have been no answers; however, we hope that in the next section we can offer some.

5 CONTENT-BASED AND NEUTRAL-BASED FREEDOM OF EXPRESSION

There is no doubt that North-America and the European Union have different law cultures and systems. Nonetheless, their doctrines are not necessarily incompatible with each other, which allows for a successful joint analysis. The European DSA has been the inspiration for several legislations around the world (as is the case with PL 2630), but this does not prevent us from using parameters from other regions, as long as they are compatible and, of course, useful for the purpose.

However, borrowing part of the US doctrinal tradition serves other more specific purposes. If the contradictory point of the “Brazilian General's Speech” case is the scope given to the right to freedom of expression, a knowingly more liberal legal culture would be able to carry our argument, roughly speaking, to its ultimate consequences. Because not all jurisprudence or doctrine could fit this review, the

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

paradigm of neutral-based and content-based regulation[8] was chosen and the progress of this case.

Briefly and crudely, the idea is to use the American interpretive rationality to evaluate the Brazilian case, analyzing which guidelines comply with freedom of expression from this perspective. This way, we evaluate Meta's initial decision of maintaining the content and its reversal by the Oversight Board, under arguments from a markedly more liberal law tradition. Of course, legal differences result from cultural distinctions on the comprehension of the right, as affirmed by Doeuk, from whom content moderation demonstrates that the "elevation of speech rights is as much a cultural phenomenon as a legal one" (2022, p. 32). The scholar explains:

> Particularly in the United States, it feels almost sacrilegious to suggest that speech should be administered systematically rather than treated as a sacred individual right deserving of highest protection. There is perhaps no more emblematic and carefully guarded constitutional right than freedom of speech. Its *Firstness* is emphasized (even though it wasn't actually first in the original draft of the Bill of Rights. (DOEUK, 2022, p. 32).

Already in 1983, Richard Schauer questions whether freedom of expression should be seen as something "special" and, recapturing the right through a historical approach, argues that "recent developments have made first amendment considerations applicable to issues that in the recent past were considered well without the boundaries of the first amendment." (1983, p. 1288). The scholar makes use of cases *Brandenburg v. Ohio, New York Times Co. v. Sullivan* and *Cohen v. California* to assert that there is a broadening of the First Amendment when it concerns freedom of expression:

> With this process of broadening, or at least arguing about broadening, has come the reemergence of theory. For although the accepted assumptions, traditional metaphors, and standard platitudes about the value of free speech might have been largely sufficient to deal with the issues of the past, they are clearly inadequate to confront the questions we must ask when trying to determine the extent to which, if at all, the courts should broaden the coverage of the first amendment to encompass a wide range of activities

[8] Macedo Junior reveals that "In Brazil, arguments based on this kind of distinction are very rare and content-based restrictions are not considered illegal since the interests behind it should be balanced with other societal interests" (2017, p. 295).

Revista Percurso

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

> seemingly so far from the comprehension of the classical free speech theorists that the relevance of classical theory has become attenuated. In the place of the classical theories have come new attempts to ask about the "Why?" of the first amendment, in the hope of developing a theory that will explain the values that the concept of free speech is designed to serve. With such a theory in place, of course, it becomes much easier to confront the questions raised by the broadening of the first amendment. For if we know why we have the principles of free speech, then we can determine in the new case whether that class of activities is the type that the first amendment is designed to promote. (SCHAUER, 1983, p. 1288)

In light of this, the scholar asks whether or not freedom of speech, as an activity covered by the first amendment, possesses any more theoretically relevant difference from those activities not so covered (1983, p. 1289). Schauer exposes a critique apparently similar to that put forward by Macedo Junior, who refutes *ad hoc* interpretations, explaining that the approach “to attempt to work out an ideal political theory independent of the particular constitutional provision at issue, such as freedom of speech, and then proceed to apply that clause to the extent that it supports that theory” (1983, 1305) should lead to some distortion:

> To the extent that the text does not fit the preconceived theory, then a little pushing and pulling, huffing and puffing, bending and slicing, and-*voila*-one's preconstructed political theory just happens to be embodied in the Constitution, with nothing left out. (SCHAUER, 1983, p. 1305).

Likewise, when dealing with decision-making rationalities at the constitutional level in regards to the right to freedom of expression, Stone (2008) stresses the relevance of determining whether the right is being balanced from a content-based or a neutral perspective. His purpose consists, precisely, of giving visibility to a theory that is proper to freedom of expression, pointing out grounds that do not lend themselves to simply substantiating an intuitive decision. Regardless of its success, the task is certainly worthwhile.

The author evaluates the construction of the decisions made by the American Constitutional Court of the 20th century on freedom of expression in order to point out some conclusions from which it is possible to develop the theme. In his fourth observation, he assesses the Court’s recognition of the content-based/content-neutral distinction from the case *Schach v. United States*, in which “the Court held

Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

unconstitutional a law prohibiting soldiers from wearing their uniforms in a theatrical productions if those productions held the military in contempt" (2009, p. 278-279). Explaining:

> […] the Court held that a content-neutral law that banned more speech was less problematic under the First Amendment than a content-based law that banned less speech. As the Court put the point, the government cannot constitutionally punish soldiers for wearing their uniforms to protest "the role of… our country Vietnam" while the same time allowing them to wear their uniforms to "praise the war in Vietnam". (STONE, 2008, p. 279)

The Court articulated these arguments to maintain something that not always seems obvious: the First Amendment rationality requires one to consider that "content-based restrictions are more likely to skew public debate for or against particular ideas" and "to be tainted by a constitutionally impermissible motivation" (STONE, 2008, p. 280). On the same line of reasoning, Macedo Junior notes the distinction made by the Court as something that "does not tell us how to evaluate the constitutionality of specific laws that fall on one side of the line or the other" (2017, p. 294). However, the distinction remains important for reasoning as it states that neither content regulation is permitted, nor speech regulation is based on content, when held in the "public square" (MACEDO JUNIOR, 2017, p. 294).

This logic seems to be close to what is set out by the Comment 34, when it stated that freedom of opinion admits no restriction and freedom of speech just admits some restriction, unless it does not put at risk the right itself. Indeed, this rule is well illustrated by Srinkhant Srinivasan, when discussing distinct speeding law and reentry law:

> "One conceivable justification for the differential treatment of the speeding law and the reentry law might focus on speech restrictive effect. The speeding law might often only delay expressive activity, while the reentry law might completely foreclose it. (1995, p. 414)

Although the distinction between content-based and neutral-based regulation is just a first step for argumentation, one thing can be certainly concluded about the "Brazilian General Speech" case: the moderation made by the Board and Facebook (even late) seems to be a content-based one. The Board's decision made it clear that the post violated the Community Standards, specifically about violence and

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

incitement. More than that: it was accurate to say that those standards were clear and comprehensible enough.

Firstly, the conclusion is important for one reason. Although Macedo Junior says that “content-neutral laws can also be a threat to speech” in particular situations (2017, p. 295), he seems to agree with Stone when recognizing that content-based regulation, unlike the neutral one, has a greater potential for unconstitutionality (2008, p. 280). Moreover, if we leave aside the criminal sphere[9], which is not relevant for now, the case fits in the Rabat Plan of Action category of expression that “concern in terms of tolerance, civility and respect for the rights of others” (2013, p. 9). So, it is definitely a matter of content. Let us keep this first finding in mind and continue thinking abstractedly.

Stone follows the logic behind the trials until he reaches a second relevant specificity: the “low value” speeches. According to the scholar, “one obvious problem with a doctrine that presumptively holds all content-based restrictions unconstitutional is that there may be some types of content that do not merit such protection.” (2008, p. 283).

Low value doctrine holds the view that certain content does not further core constitutional value of freedom of expression. It establishes that some categories of speeches do not impact the constitutional right to freedom of speech at all. In these cases, what can be seen is just a reflective and almost irrelevant effect. Stripping away any prejudices and considering that the court's understanding has evolved to rule out any absolute immunity of such subjects, it is possible to make a case-by-case interpretation. Although unpopular[10], Stone's position is not altogether unreasonable on the point that he maintains that

[9] In addition, even if we consider the criminal sphere, there is a content matter, just because the act fits a conduct classified as a crime. This is simply a legislative option about criminal offense. Most important, however, is the offense to national security and to other’s rights.

[10] According to Srinivasan, “in the Court's view, such profound implications for judicial review suggest that incidental restraints simply cannot always trigger First Amendment scrutiny-the resulting deluge of First Amendment claims could overwhelm the courts with constitutional balancing inquiries. (…) This concern with "First Amendment overload" is somewhat unclear, for the balancing inquiry need not be an involved one in every case; the Court could quickly dispense with situations like the arrest of a newscaster for speeding by engaging in only a proforma balancing. (SRINIVASAN, 1995, p. 406).

Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

> […] the low value doctrine is a sensible and pragmatic compromise that serves a salutary function by operating as a useful safety valve, enabling the Court to deal reasonably with somewhat harmful, but relatively insignificant speech. (2008, p, 284)

As said above, this is not an easy position – and we are not willing to agree with it at first sight –, but, if we think about it seriously and casuistically, it may be possible to say that this would be the case of the Brazilian content that reached the Board. It can be explained by the fact that the discourse was not about an opinion, but about an act forbidden by law – and, in our case, by the Community Standards as well –, which is the violence incitement. The benefit of placing the case outside the field of freedom of expression, pointing that it has nothing to do with it, would be not to scale that to a constitutional level.

In spite of this, we have to acknowledge that, in a country with young cultural traditions, the participation of the Supreme Court in some controversial themes should not be seen as just a bad procedure. The judgment of a constitutional court may have juridical positions and, through that, it may allow the doctrine's development in some coherent way. However, it is undeniable that our Supreme Court receives many more demands than it is able to deal with in a reasonable time frame. Time and context are particularly important in a digital framework and the dispute can just lose its object.

Revisiting what the Comment nº 34 stated, the right to speech is not complete freedom and could be restricted unless it does not put at risk the right itself. In the Brazilian case, the taking down of the content is not a measure able to eliminate freedom of expression. It would just eliminate illegal content, without prejudice for the user, who will be able to post other contents expressing their disagreement with political issues. There will be no suspension of the user's count and, if we take serious our compromise with transparency, the platform would have to notify the user about what is wrong with the content withdrawn – the violation of Community Standards.

6 CLUES FOR A DIAGNOSIS

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

The absence of a systematic model may lead to *ad hoc* interpretations, based mostly on the judge's intuition other than on a consistent doctrine. If there will be no consensus at all times, jurists should create some for now and maintain coherence for the future. This is not an easy task, but there are useful foreign experiences which, even if they cannot be transplanted, can probably give us a direction to follow.

Particularly with regard to PL 2630, one of the major controversial points was the provision for collecting personal data by the provider, also requiring linkage to a telephone number. The requirement was disproportionate and dangerous in relation to data protection. In addition, the link to a cell phone, apart from generating unjustifiable discrimination, presented risks to anonymity which, as it is well known, also represents a guarantee for the democratic exercise of freedom of expression. Fortunately, these provisions were deleted in the substitute. In this sense, at the 5th public hearing held on August 19, 2021 this issue was raised by Paulo Renâ, arguing that "anonymity is important for whistleblowers and disinformation is made by messages from known people" (BRASIL, 2023a, s.p.).

As highlighted in the 1st public hearing, a sort of due process of law seems to have been created for content moderation with the provision of a security protocol in art. 12 of the law. In such a manner, it is reasonable that the protocol was established as an *ultima ratio*, being extended only "when the insufficiency of less severe measures to remove the imminent risk is demonstrated" (art. 12, §1, PL 2630).

It would be interesting to apply subsidiarity also when the protocol is initiated, but the law has remained silent on this. We support the subsidiarity of the protocol because, according to art. 13, *caput* and sole paragraph, it is from the establishment of the security protocol and due notification that "providers may be held civilly liable for damages arising from content generated by third parties when prior knowledge is demonstrated". The liability of the provider in case of imminent risk of damage, in this case, was established as joint and several. Thus, given that the establishment of the security protocol may bring considerable penalties for the provider, it should be done cautiously and, above all, in the absence of less burdensome measures.

unicuritiba
Revista Percurso
Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

Still in relation to the due process of law, art. 17 established some parameters for moderation, namely “fairness, consistency and respect for the right of access to information, freedom of expression and free competition” (art. 17, *caput*). In the sole paragraph of the same provision, it is cited “by the principles of necessity, proportionality and non-discrimination, including regarding users' access to the services of providers” (art. 17, sole paragraph). In fact, although the parameters are reasonable, they seem to provide little clarification on practical issues and to repeat well-known constitutional values. For the providers, they may seem like risky and overly vague forecasts, which allows for such different interpretations, as it is the case with freedom of expression.

The same criticism applies to art. 18, whose §1 provides that, if systemic risks are found, the measures promoted by the provider that involve the use of automated systems must include safeguards that prove to be appropriate and effective. The provision goes on to state that such measures must be equipped with “human supervision with a view to ensuring accuracy, proportionality and non-discrimination that is unlawful or abusive”, which may cause problems depending on the demand for moderation, given the volume of publications. Perhaps a more assertive alternative would consist of linking human intervention to cases involving duty of care, giving greater objectivity at least to the most relevant situations.

In the midst of the controversy, the rapporteur's argument for suspending the procedure is justified. In terms of content moderation alone, at least 52 other bills have been added to PL 2630. In relation to transparency, at least 16 more projects were included, and there was no time to analyze the various positions. On the other hand, the urgency of legislating on the subject is recognized, and there is no benefit in prolonging the suspension, even pending Themes nº 533 and 987 at the Federal Supreme Court. In other words, the issue cannot be extended indefinitely. All in all, several other points of PL 2630 could be analyzed, but such a broad approach would not fit in this work. For now, we have presented the issues that we consider most relevant.

With the analysis of the case presented in this paper, it is possible to verify how important it is to regulate platforms. Rules on platform regulation should not empty formalities without compromise. They should reflect society’s dearest values

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

and human rights. Rules represent a procedure that may lead to more or less freedom of expression. Even with the efforts of Meta to comply with the law by creating the Oversight Board, the Brazilian case showed that the measures taken were neither appropriate, nor fast enough. The risk of harm was evident and, despite several complaints, the platform did not recognize that.

Rules on platform regulation should not empty formalities without compromise. They should reflect society's dearest values and human rights. Rules represent a procedure that may lead to more or less freedom of expression. Even with the efforts of Meta to comply with law by creating the Oversight Board, the Brazilian case showed that the measures taken were neither appropriate, nor fast enough. The risk of harm was evident and, despite several complaints, the platform did not recognize that.

Perhaps this kind of negligence explains part of the Brazilian political context – and, possibly, the reverse must also be true at the moment. Of course, as it has been said, what happened on January 8 cannot be explained by means of an ethereal phenomenology. The event could or could not have occurred regardless of any social network. However, the fact that it did occur should raise an alarm, making us question which factors contributed to it and, especially, which could have been avoided. At this point, it is clear that, with a minimum of diligence, content moderators could and maybe should have removed or, at least, escalated the post for higher review.

Some clues of a way to follow could be deduced from the conclusions of the research. Freedom of expression must be treated as a right for itself before being submitted to any kind of balance. It demands a consistent theory that express population deserves and, moreover, that guarantees foster legal certainty and the construction of a coherent doctrine. No magic wand will help this purpose; it is the responsibility of legislators to commit to it before anything else.

The governance model requires joint action by public authorities, private entities and the population. All parts are still learning, including platforms, which indicates the potential of a co-operative model that goes beyond the simple allocation of responsibilities. Neither public authorities will be able to turn content moderation into reality themselves, nor will just rules succeed in promoting it. If there is still any

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

doubt, that is a good reason not to take the platforms as enemies of the state. They are, above all and just like freedom of expression, an instrument of guaranteeing the exercise of other rights. The creation of the Oversight Board itself demonstrates the willingness of Meta to act in accordance with national legal parameters.

These warnings may indicate the need for a public debate that is less focused on political parties or governmental needs and more aligned with community discourse. To say that platforms will be accountable is small-minded compared to the many questions that permeate the causes of such accountability. Moreover, to define a list of criminal legal situations that should be considered by platforms is to say more of the same, because most of these criteria are already found in community standards.

Thus, there is no way to foresee the facts of tomorrow, although it is possible to predict what should be preserved in the last circumstances. This demands an agenda of constant debate, not only centered in any regulatory body, but one that is an initiative of the public authorities and shows the willingness to build public policies, not only to distribute responsibilities.

7 FINAL CONSIDERATIONS

The aim of this article was to analyze the “Brazilian General Speech” case in light of freedom of expression. Firstly, the case was described and allocated as content moderation that, as expected, implies the right of freedom of expression. Then, the right was resignified from the applicable international human rights standards, concluding that there is a need for a systematization of the matter. This task was concluded in the last topic, not before correlating the issue with the debate about the regulation of platforms in Brazil.

In the ultimate topic, it was concluded that, at least considering the North-American content-based/neutral-based doctrine, there is no evidence that the “Brazilian General Speech” case fits in the sphere of freedom of speech. So far, it seems that the first objectives have been achieved with the answer to the question of this review.

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

At the end, efforts were made to carefully link our conclusions with the discussion of regulation of content moderation. A brief diagnosis of what can be done to guarantee freedom of expression on platforms was done, indicating a collaborating governance approach. This may enable commitment for all parts, instead of a simple distribution of responsibilities. Hence the need to build public policies that make clear which are the aims of our actions.

As said before, there are no concrete conclusions, but rather hints of a path to be followed. Giving meaning to freedom of expression is an ongoing job – one that, fortunately, has already begun. Rules and procedures may help it, but are not the core of the solution as our needs will always be bigger than this.

If we want to preserve our democracy, we have to build a democratic model of regulation. Different solutions in every case will lead us to an undemocratic way. Platform regulation will not change Brazilian political background and technological developments will not go into reverse, but even so, it is urgent to regulate platforms. However, it is even more vital to know which are the goals of these rules. What kind of freedom of expression do we want?

REFERENCES

BAMBERGER, Kenneth A., MULLIGAN, Deirdre K., Allocating Responsibility in **Content Moderation, A Functional Framework, *Berkeley Technology Law Journal***, v. 36, n. 3, p. 1091-1172, 2021.

BRASIL, Planalto, ***Lei nº 12.965, de 23 de abril de 2014***, Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm Acesso em: 17.jun.2023.

BRASIL, **Supremo Tribunal Federal, *Tema 533*,** Dever de empresa hospedeira de sítio na internet fiscalizar o conteúdo publicado e de retirá-lo do ar quando considerado ofensivo, sem intervenção do Judiciário, Agravo em recurso extraordinário (RE 1057258) em que se discute, à luz dos artigos 5º, II, IV, IX, XIV, XXXIII e XXXV; e 220, §§ 1º, 2º e 6º, da Constituição Federal, se, à falta de regulamentação legal da matéria, os aludidos princípios constitucionais incidem diretamente, de modo a existir o dever de empresa hospedeira de sítio na rede mundial de computadores de fiscalizar o conteúdo publicado em seus domínios eletrônicos e de retirar do ar informações consideradas ofensivas, sem necessidade de intervenção do Poder Judiciário. Disponível em:

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

https://portal.stf.jus.br/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=5217273&numeroProcesso=1057258&classeProcesso=RE&numeroTema=53 Acesso em: 17.jun.2023.

BRASIL, **Supremo Tribunal Federal, *Tema 987***, Discussão sobre a constitucionalidade do art. 19 da Lei n. 12.965/2014 (Marco Civil da Internet) que determina a necessidade de prévia e específica ordem judicial de exclusão de conteúdo para a responsabilização civil de provedor de internet, websites e gestores de aplicativos de redes sociais por danos decorrentes de atos ilícitos praticados por terceiros, Recurso extraordinário (RE 1037396) em que se discute, à luz dos arts. 5º, incs. II, IV, IX, XIV e XXXVI, e 220, caput, §§ 1º e 2º, da Constituição da República, a constitucionalidade do art. 19 da Lei n. 12.965/2014 (Marco Civil da Internet) que impõe condição para a responsabilização civil de provedor de internet, websites e gestores de aplicativos de redes sociais por danos decorrentes de atos ilícitos de terceiros. Disponível em: https://portal.stf.jus.br/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=5160549&numeroProcesso=1037396&classeProcesso=RE&numeroTema=987 Acesso em: 17.jun.2023.

BRASIL, **Ministério da Justiça e Segurança Pública, *Portaria MJSP nº 351***, *de 12 de abril de 2023*, Dispõe sobre medidas administrativas a serem adotadas no âmbito do Ministério da Justiça e Segurança Pública, para fins de prevenção à disseminação de conteúdos flagrantemente ilícitos, prejudiciais ou danosos por plataformas de redes sociais, e dá outras providências. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-mjsp-n-351-de-12-de-abril-de-2023-476702096 Acesso em: 17.jun.2023.

BRASIL, Câmara dos Deputados, ***Parecer Proferido em Plenário ao Projeto de Lei nº 2.630, de 2020,*** *e apensados*, Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet, 2023ª. Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=2265334&filename=PRLP%201%20=%3E%20PL%202630/2020 Acesso em: 17.jun.2023.

BERTONCINI, Thays, MIACHON, Caio Tenório, (In)constitucionalidade do artigo 19 da Lei 12.965/14, Justiça? In: CAMARGO, Solano de, SOARES, Paulo Vinicius de Carvalho (Orgs), *Direito Digital*, **Novas Fronteiras e desafios em um mundo em transformação,** 1 ed., São Paulo, LBCA, 2023, p. 290-309.

CAVALCANTI, Ana Elisabeth Lapa Wanderley, LEITE, Beatriz Salles Ferreira, BARRETO JUNIOR, Irineu Francisco, Sistema de Responsabilidade Civil dos Provedores de Aplicações da Internet por Ato de Terceiros, Brasil, União Europeia e Estados Unidos da América, ***Revista Eletrônica do Curso de Direito da UFSM*****,** v. 13, n. 2, p. 506-531, 2008.

DOUEK, Evelyn, Content Moderation as Systems Thinking*,* ***Harvard Law Review***, v. 136 n. 2 p. 528-607, dez. 2022.

DOEUK, Evelyn, Facebook's Oversight Board, move fast with stable infrastructure

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

and humility, *North Carolina Journal of Law & Technology*, v. 21, p. 2-78, out. 2019.
ESTARQUE, Marina, ARCHEGAS, João Victor, ***Redes sociais e moderação de conteúdo*, criando regras para o debate público a partir da esfera privada**, **Instituto de Tecnologia & Sociedade do Rio**, 2021. Disponível em https://itsrio.org/pt/publicacoes/redes-sociais-e-moderacao-de-conteudo/ Acesso em 10.abril.2021.

EUROPEAN UNION, The European **Parliament and the Council of the European Union, Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC** *(Digital Services Act) (Text with EEA relevance)*. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065&qid=1666857835014 Acesso em: 17.jun.2023.

FAUSTINO, André, FUJITA, Jorge Shiguemitsu, O Princípio da Inimputabilidade da Rede e a Remoção de Conteúdos dos Provedores de Aplicações da Internet, ***Revista Jurídica Cesumar***, v. 17, n. 3, p. 809-829, set./dez. 2017.

MACEDO JUNIOR, Ronaldo Porto, Freedom of Expression, what lessons should we learn from US experience, ***Revista Direito*** *FGV*, v. 13, n. 1, p. 274-302, jan/abr 2017.

OVERSIGHT BOARD, *2021-001-FB-FBR*, **Former President Trump's suspension, 2021**. Disponível em: https://www.oversightboard.com/decision/PAO-NR730OFI/ Acesso em 13. maio.2023.

OVERSIGHT BOARD, *2023-001-FB-UA*, **Brazilian General Speech, 2023**. Disponível em: https://www.oversightboard.com/decision/FB-659EAWI8/ Acesso em: 08.jul.2023.

OVERSIGHT BOARD, **Public Comments Appendix,** 2023a. Disponível em: https://www.oversightboard.com/decision/FB-659EAWI8/ Acesso em: 08.jul.2023.

SCHAUER, Frederick, "Must Speech be Special?", **Northwestern University Law Review**, v. 78, p. 1284-1306, 1983.

SCORDATO, Marin Roger, Free Speech Rationalities After Septembre 11th, The First Amendment in Post-World Trade Center America, **Stanford Law & Policy Review**, v. 13, n.1, p. 185-203, 2002.

SRINIVASAN, Srinkhant, **Incidental Restrictions of Speech and the First Amendment, A Motive-Based Rationalization of the Supreme Court's Jurisprudence**, *Constitutional Commentary*, v. 12, p. 401-420, 1995.

STONE, Geoffrey R., **Free Speech in the Twenty-First Century, Ten Lessons from the Twentieth Century Lead Article.** *Papperdine Law Review*, v. 36, n. 273, p. 273-299, 2008.

Submetido em: 25/12/2022
Aprovado em: 08/01/2023
Avaliação: *Double Blind Reviewe*-
ISSN: **2316-7521**

TEFFÉ, Chiara Spadaccini, MORAES, Maria Celina Bodin de, Redes sociais virtuais, privacidade e responsabilidade civil*,* ***Pensar Revista de Ciências Jurídicas***, v. 22, n. 1, p.108-146, jan/abr 2017.

UNITED NATIONS, ***Annual report of the United Nations High Commissioner for Human Rights***, 2012. Disponível em: https://www.ohchr.org/sites/default/files/Rabat_draft_outcome.pdf Acesso em: 08.jul.2023.

UNITED NATIONS, ***General comment N.º 34.*** Disponível em: https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf Acesso em: 08.jul.2023.

UNITED NATIONS, **International Covenant on Civil and Political Rights**, Disponível em: https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights Acesso em: 08.jul.2023.

YOUNG, Greyson K, **How much is too much, the difficulties of social media content moderation.** Disponível em: https://ssrn.com/abstract=3792647 Acesso em 13.maio.2023.