

input_

new challenges in the liability of internet application providers

**implications for the digital economy and
e-commerce following the Federal
Supreme Court's decision on the Marco
Civil da Internet**

september, 2026

credits

project team

francisco brito cruz (coordinator)

fernando gallo

helená maluf

vinicius harada

support

mercado livre

how to cite this document?

BRITO CRUZ, Francisco (coord.); MALUF, Helena; and HARADA, Vinicius. *New challenges in the liability of internet application providers*. Input, São Paulo: 2026.

executive summary

This document summarizes the discussions of two workshops on the liability of internet application providers following the Federal Supreme Court (STF) decision on Article 19 of the Marco Civil da Internet (Brazil's Internet Civil Rights Framework), held in São Paulo on April 27, 2026 and August 10, 2026 under the *Chatham House* Rule, each with around 35 participants from the private sector, the public sector, academia and civil society.

The cross-cutting diagnosis is that **the new regime has significantly expanded the duties assigned to providers, but has left open a relevant share of the parameters needed to guide compliance with them.** Concepts such as duty of care, diligent conduct, systemic failure, systemic risks and reasonable doubt now occupy a central position in platform liability, without it yet being sufficiently defined what each of them requires in practice. The result is an environment in which providers, administrative authorities and the Judiciary may adopt different metrics and expectations regarding the same obligation, increasing legal uncertainty and creating incentives for the preventive removal of legitimate content and listings, which may in turn lead to an increase in litigation.

A first axis of discussions therefore examined the duties of care and diligence. The discussions point to the need to treat them predominantly as duties to adopt adequate processes, mechanisms and safeguards, rather than as a guarantee of the complete elimination of unlawful content. In this context, the document explores alternatives for bringing greater objectivity to oversight, such as anchoring the duty of care in the principle of objective good faith (*boa-fé objetiva*), using prevalence indicators and qualitative parameters on moderation architecture, as well as taking into account technical limitations, the scale of services and different business models. It also discusses problems specific to e-commerce, especially the need to distinguish consumer harms, subject to the specific regime of the Consumer Defense Code (CDC), from harms related to content and the circulation of speech, governed by the Marco Civil.

The second axis addresses the new role of extrajudicial notices. The main issue is preventing the regime from turning, in practice, into an automatic removal obligation. “Reasonable doubt” emerges as an essential mechanism for preserving a notice and action model, in which a notice triggers a duty to analyze and respond, but not necessarily to take content down. The workshops identified, however, significant risks of notices being weaponized by competitors, associations or other actors, as well as difficulties caused by mass requests that are poorly individualized or based solely on keywords, hashtags and hashes. In this scenario, the workshops emphasized the importance of precisely identifying content, preserving its contextual analysis and appropriately allocating duties among those who formulate the request, those who order the measure and those who must carry it out. The goal is to prevent the platform from being turned simultaneously into investigator, technical authority and adjudicator of disputes whose resolution depends on

information or jurisdiction held by the State itself, effectively transferring to it the exercise of police powers assigned to state authorities.

A third axis discusses the institutional governance of the new regime.

Assigning new competences to the National Data Protection Agency (ANPD) may help form a regulatory hub capable of producing more stable technical parameters, but its work takes place in an environment already marked by a multiplicity of administrative authorities. The challenge, therefore, is not merely to allocate powers, but to build coordination and some form of deference to technically grounded regulatory parameters. This discussion also shows that compliance with the new duties does not depend on platforms alone: when automated verification of public licenses, registrations or authorizations is required, for example, the State itself must make available structured, up-to-date and technically consumable databases.

Based on these assessments, the document proposes an agenda aimed at giving the new regime greater density and predictability.

Taken together, these proposals share the same premise: the effectiveness of the new liability regime depends less on the abstract expansion of obligations and more on building institutional, technical and procedural parameters capable of making them enforceable. The challenge opened by the STF's decision is thus to turn broad legal categories into a predictable system of incentives and responsibilities, capable of tackling unlawful content and activities without producing, as a side effect, over-removal, uncertainty for legitimate economic activities or fragmentation of Brazilian digital governance.

table of contents

introduction	6
project overview and background	6
work and method premises	6
I. challenges in due diligence and duty of care	8
scope of duties of care and diligence	8
between obligations “of result” and “of means”	9
parameters for “systemic” assessment and their challenges	10
operational limitations in compliance with obligations and their consequences	11
systemic risks vs. systemic failure	11
protection of vulnerable groups	12
what does the civil liability regime look like in e-commerce?	13
II. extrajudicial notices	15
weaponization of extrajudicial notices for abusive purposes	15
challenges in notices sent by public authorities	16
sharing data with public authorities	17
“reasonable doubt”	17
requirements for unequivocal identification of content	17
III. governance, enforcement and the role of the public sector	20
ANPD as the authority for the digital environment: tensions and risks	20
views on a regulatory deference agenda	21
public data as an input for providers to meet their obligations	22
IV. recommendations and convergences	23
self-regulation for technical standardization	23
regulatory deference agenda for legal certainty	23
compliance agenda and actions by public authorities	24
agenda for regulatory densification of the duty of care and diligent conduct	25
agenda for detailing the extrajudicial notice regime	26

introduction

project overview and background

This summary document presents and organizes the main ideas discussed at two workshops on the liability of internet application providers following the decision of the Federal Supreme Court (STF) on Article 19 of the Marco Civil da Internet (Brazil's Internet Civil Rights Framework, MCI), focusing on its effects on the digital economy and e-commerce. **The organization and funding of the workshops were supported by Mercado Livre.**

The first workshop focused on the scenario that emerged in response to the Federal Supreme Court (STF) ruling on leading cases no. 987 and 533, which concerned the constitutionality of Article 19 of the Marco Civil da Internet (Law No. 12,965/2014). In June 2025, when examining two extraordinary appeals,¹ the Court ruled that Article 19 of the MCI was partially and progressively unconstitutional, finding that the article did not afford sufficient protection to highly significant constitutional legal interests. The thesis established in the ruling thus significantly altered the overall framework of liability for internet application providers.

Subsequently, in May 2026, the Executive Branch issued Decrees No. 12,975 and No. 12,976/2026. The first addressed the general duties of internet application providers and amended the MCI's implementing decree, published in 2016²; the second establishes specific measures to combat violence against women in the digital environment.³ In June 2026, also contributing to the subsequent change in the new regulatory landscape, the STF ruled on the motions for clarification (embargos de declaração) filed against the extraordinary appeals. Indeed, a number of issues highlighted by participants at the first meeting were addressed in the texts of the decrees and in the STF's decision. For this reason, it was decided to hold a second workshop.

This new legal regime is still being absorbed by actors such as courts, companies, civil society and academia. Its practical implications for the digital economy (its incentives, business models and competitive dynamics) and for freedom of expression remain uncertain, and the workshops were organized precisely to map these implications. The meetings made it possible to (i) gather practical and theoretical perceptions from different actors about the scenario created by the STF's ruling, and (ii) map recommendations and forward-looking propositions on how to promote greater predictability in this framework.

work and method premises

Both workshops were held in person in São Paulo, on April 27, 2026 and August 10, 2026, respectively, running from 10:00 a.m. to 5:00 p.m. Each meeting brought together around 35 participants from different sectors – private-sector lawyers, academics, public-sector representatives and members of civil society organizations – all working on issues at the intersection of law, technology, digital markets and platform regulation. The workshops were structured in thematic sessions. Each session opened with guiding remarks on the topic to be addressed, followed by a broader debate open to all participants. **The workshops were held under the Chatham House Rule, so this summary document identifies neither the participants nor the authorship of any idea.**

As stated above, the **organization and funding of the workshops were supported by Mercado Livre**. However, this document does not necessarily reflect the company's views, but rather an account of the discussions held at the workshops.

The workshops did not aim to endorse the decision on leading cases no. 987 and 533, nor the decrees issued by the Executive Branch, but rather to discuss their practical implications across sectors. The first working premise, therefore, is that discussing these implications **does not exclude participants who hold critical views of decisions by the Judiciary or Executive Branch.**

¹ RE 1,037,396 and RE 1,057,258.

² Decree No. 8,771/2016.

³ Decree No. 12,976/2026.

The second premise is that the ruling, the issuance of the decrees and the decision on the motions for clarification have consolidated a new starting point for the debate on provider liability **that requires answers on how to operationalize these concepts and rules in institutional practice.** Accordingly, the insights and recommendations presented throughout this document result from the accumulated debates held over the course of this entire process, offering a contribution to a regulatory ecosystem that is still very much under construction.

I. challenges in due diligence and duty of care

The concepts of “duty of care” (associated with the obligation to immediately make unavailable certain content constituting crimes listed in the thesis issued by the STF) and “diligent conduct” (associated with the rebuttable presumption of fault in the case of paid ads and boosted content established in the STF's thesis) were among the most debated concepts at the workshops. The centrality of the topic stems largely from its indeterminacy, since these are categories still under construction, applied in varied ways in foreign jurisdictions.

scope of duties of care and diligence

Across both workshops, participants converged around a shared concern about the difficulty of interpreting the category of duty of care and other terms, such as diligent conduct. The semantic indeterminacy of the concept of duty of care was identified as a potential source of legal uncertainty, insofar as it hinders a clear delimitation of providers' responsibilities and the predictability of judicial and administrative decisions. In this regard, participants stressed the importance of progressively consolidating the duty of care on the basis of more precise parameters, guiding conduct and serving as a consistent reference both for the private sector and for administrative regulation and adjudication.

An example of the concern about the scope of the duty of care can be seen in a debate held at the second workshop regarding two layers of the provider liability regime established by the STF and detailed in the aforementioned decrees. On that occasion, some participants pointed out significant differences between the duty set out in Article 16-B of Decree 12,975/2026, concerning care for the “immediate removal” of content constituting serious crimes (indicating “systemic failure” as the case of non-compliance), and the obligations set out in Articles 16-K and 16-L of the same instrument, which establish a presumption of liability (or, according to the ruling on the motions for clarification, a *presumption of fault*) where there is no proof that providers “acted diligently and within a reasonable time to make the content unavailable” (without any reference to the idea of “systemic failure”, for example). These two provisions led workshop participants to ask whether the new regime in fact requires platforms to perform *two different duties of diligence*.

Still in the debate on the limits of the duty of care and diligent conduct, participants suggested that these concepts be legally interpreted in light of the classic private law notion of objective good faith (*boa-fé objetiva*).⁴ In this sense, participants questioned the uncritical use of the imported notion of *duty of care*, to the detriment of a reading grounded in the ancillary duties that Brazilian private law scholarship has consolidated over decades: loyalty, transparency, care and confidentiality. These four duties, anchored in legal doctrine and case law as corollaries of objective good faith in contractual relations, would offer an adequate doctrinal foundation for understanding the duty of care in the digital context. Under this possible reading, the duty of care would not be an autonomous category or one inaugurated by the STF's decision, but rather an “inheritance” of objective good faith. Anchoring the debate in this civil law framework was seen by some participants as a legally sound strategy for building, on the basis of the duty of care, a dam against a generalized incentive for providers to remove legitimate content, or to engage in indiscriminate removals. A clear reference to objective good faith would counter this incentive to remove content whenever there is even a minimal degree of uncertainty about its lawfulness, with all the risks this would entail for the fundamental rights involved in this decision.

Some participants also identified attempts to expand the interpretation of the concept duty of care beyond its original contours. It was reported that administrative authorities and industry associations are already carrying out, in practice, a genuine “transplant” of this notion into the sphere of personal data protection, taking the STF's decision as an enabling basis. This

⁴ A central aspect of objective good faith is that it gives rise to ancillary duties in the contractual relationship: obligations that are not expressly provided for in the contract, but stem from the principle itself and are widely recognized by Brazilian legal scholarship and case law. The main established ancillary duties include: the duty of loyalty (acting honestly and without intent to harm the other party), the duty of transparency (providing relevant information and not omitting essential facts), the duty of care (acting diligently to avoid harm to the counterparty) and the duty of confidentiality (preserving confidential information obtained in the course of the relationship). Breach of these duties, even in the absence of an express contractual clause, may give rise to civil liability.

expanded reading of the duty of care, however, has been advancing in a diffuse manner and without precise normative contours, moving into subject areas, such as privacy and data processing, that have their own well-established regulatory framework. Participants from academia and the private sector thus expressed skepticism about the need to bring the concept of duty of care into the scope of the Brazilian General Data Protection Law (LGPD), given the duties already consolidated by that statute.

In terms of scope, another topic discussed was the relationship between platforms' monetization model and the extent of their duty of diligence. The premise that whoever reaps the benefits must bear the burdens arising from its activity was discussed by some participants, who urged that this inference be drawn with caution, insofar as its application requires greater conceptual precision as to what, concretely, are the benefits and burdens at stake in the debate on digital platforms, and how relevant they are to determining the allocation and appropriate measure of liability arising from the duty of diligence.

In this aspect, a striking example from the debate was the distinction between organic content and boosted content. More specifically, when a platform profits directly from the artificial distribution of certain content through ad networks or paid posts, it was argued that the link between the platform and that content is qualitatively different from one in which, for example, users spontaneously post ads as “organic content” in the environments maintained by the provider. In the latter case, it would still be an ad in the broad sense, but not in the strict sense of content boosted by the platform, nor content from which it profits. On this point, workshop participants suggested that it is fair and proportionate for different liability regimes to apply to each scenario, with a stricter regime for cases in which the platform has effective control over the paid activity and derives direct economic benefit from it.

between obligations “of result” and “of means”

Following the STF's decision, the duty of care gained strength as a new lens for cases in which platforms should be held liable for “systemic failure” according to the obligation to “immediately make unavailable” content constituting certain serious crimes listed by the Court. “Diligent conduct”, in turn, also stood out as a parameter for rebutting the relative presumption of fault in cases of harm arising from boosted content and paid ads.

At the workshops, participants drew attention to different interpretations of these concepts, related to the notions of an “obligation of result” or an “obligation of means”. The former, focused on consequences, stems from the idea that compliance with the duty of care is assessed by analyzing the “results” the duty of care seeks, that is, the degree to which unlawful content is accessible and displayed to platform users. The latter, focused on conduct, is based on an understanding of the duty of care as a duty to adopt certain mechanisms and procedures expected of an internet application provider, capable of showing a sufficient degree of security in preventing or removing illegal content. Under this latter model, the parameter for assessing breaches of the duty of care would rest on the soundness of the moderation architecture developed by platforms to prevent unlawful activity, and not on the mere detection of such content.

When discussing the issuance of Decrees No. 12,975 and 12,976, participants at the second workshop indicated their provisions as evidence that the duty of care would predominantly constitute an “obligation of means”, noting that compliance would be assessed through the use of methodologies and efforts approved by the regulator to prevent the widespread circulation of illegal content. This is also reflected in the case of boosted content and paid ads, pointing to the demonstration of “diligent conduct”.

However, some workshop participants expressed concern that understanding the duty of care as an obligation of means might clash with the notion of “systemic failure” (connected specifically to the duty of care in the case of serious crimes, and not to diligent conduct in the context of advertising paid to the provider).

To illustrate this uncertainty, participants recalled statements made during the Federal Supreme Court's ruling which, although not included in the final text of the thesis, would suggest an outcome-oriented reading of liability. Under this stricter perspective, finding a significant volume of unlawful content in circulation would suffice to characterize the system

as ineffective and trigger the platform's liability, turning the assessment of means into a retrospective verification of results.

It was thus noted that there is a zone of indeterminacy that strains the application of the duty of care, since its qualification as a procedural standard coexists with a pressure toward quantitative control of posts through the concept of systemic failure. This interpretive indeterminacy creates the possibility that compliance verification will move away from examining the soundness of internal moderation processes and begin to demand a guarantee that irregularities will be eradicated, thereby bringing the debate back, once again, to the scope of an obligation of result.

As will be seen below, a question arose among participants as to which metrics would be most appropriate for establishing a platform's "systemic" lack of success, which generated perceptions of legal uncertainty and room for discretion in oversight. In particular, this question was connected, in the workshop discussions, to the understanding that the obligation imposed on platforms by the duty of care does not concern the individual review of illegal content analyzed in isolation, as explicitly stated in the thesis published by the STF and in the new Article 16-B, §3, now part of Decree 8,771/2016.

parameters for "systemic" assessment and their challenges

A number of participant contributions arose in the debate on the best metrics for assessing systemic failure (or success) under the new liability regime established by the STF and detailed in the decrees published by the Executive Branch. On this point, the same Article 16-B of Decree 8,771/2016, in its §1, provides that a provider is deemed to have committed a "systemic failure" if it cannot demonstrate adequate prevention and removal measures that "provide, according to the state of the art, the highest levels of security" (item I) and "inhibit mass circulation" of the specified unlawful content (item II). The debate then turned to the possible regulatory meaning of the concept of "mass circulation".

This approach would start from the premise established in the STF's thesis and in the aforementioned Article 16-B, §3, of Decree 8,771/2016 that, given the scale of the digital environment, responsible conduct cannot be measured by the absolute absence of failures, but rather by demonstrating that such failures are isolated episodes, a residual and non-representative share of the total universe of content. As will be seen below, standardization in this case was seen as key to consolidating an environment of legal certainty.

One approach raised in the workshop debate was the "prevalence" metric for measuring the success of platforms' diligence in content moderation. Implemented in a specific case in the private sector, the metric would estimate the percentage of times users view unlawful/violating content as opposed to total views of any type of content,⁵ representing a first attempt, within a given service, to produce a fraction with a numerator and a denominator for scrutinizing moderation activity. However, participants reported facing challenges in having this type of metric accepted in the judicial discussion of concrete cases, pointing once again to uncertain terrain regarding the technical standards applicable to assessing the level of diligence and effectiveness of safety systems implemented by different providers.

Participants also discussed structural gaps in building *ex ante* parameters for assessing "systemic failures", especially when new behavioral trends among users, coded terms and atypical mobilizations of content by platform users emerge. These innovative phenomena were identified as potentially capable of circumventing established moderation architectures, whether due to the speed at which they spread or the ease with which they can be camouflaged within seemingly lawful speech.

As a concrete illustration, participants cited the use of the expression "Selma's party" ("festa da Selma") in organizing the events of January 8, 2023, highlighting the difficulty of anticipating, in advance and exhaustively, coded content that emerges unexpectedly in public debate and may amount to offenses included in the duty-of-care list. The diversity and

⁵ Douek, Evelyn, Content Moderation as Systems Thinking (January 2022). Harvard Law Review Vol. 13. Available at: <https://harvardlawreview.org/print/vol-136/content-moderation-as-systems-thinking/>

unpredictability of trends in content production were thus identified as features of digital dynamics that regulation must also take into account.

operational limitations in compliance with obligations and their consequences

Given the regulatory assessment of the means and, possibly, the results of content moderation, a discussion arose among workshop participants about the expectations regarding the operational capacities of different providers in this area. On one hand, participants reported a perception among public authorities and civil society that hiring new workers at large scale seemed to be the appropriate response for handling the prior review of ads that escape detection, or content moderation in general. On the other hand, other voices pointed to the practical obstacles that make this expectation hard to reconcile with the reality of the digital economy, the complexity and volume involved in content moderation, and the different sizes of internet application providers.

At the same time, the role of artificial intelligence (AI) in this review and moderation process also sparked a substantial debate. Participants identified an ambiguity regarding the role of AI: while skepticism about the use of AI in regulatory contexts has traditionally prevailed, there is at the same time a growing tendency among numerous social actors to champion it as a solution capable of overcoming all the operational limitations of moderation, as if resorting to AI would, in theory, allow every single ad and piece of content published to be fully reviewed. Views expressed especially by private-sector participants indicated that both large-scale hiring of workers and widespread use of AI, when presented as "definitive solutions", prove to be unfeasible proposals disconnected from the concrete dynamics of the digital market. They also stressed that both initiatives heighten the complexity of processes for meeting regulatory requirements, and that both raise additional issues that may negatively affect activities aimed at giving practical effect to the duty of care.

Prompted by these claims of a "definitive solution", participants from different sectors discussed the evolution of case law on a possible duty of prior monitoring of content displayed to users (citing, in this regard, the prevailing understanding of the Superior Court of Justice (STJ) over the past decade, which regarded such a duty of prior monitoring as foreign to the essence of the services offered by digital platforms). Critically, it was assessed that this interpretation now appears to have been superseded by the Judiciary's inclination to recognize prior monitoring as a constitutive element of digital services, at least when reserved for certain more sensitive categories of content.

In this regard, multiple participants converged on the view that it is necessary to delimit this duty of prior monitoring (now connected to the idea of duty of care or diligent conduct), setting objective criteria on what can be required of platforms, with attention to the limits of their technical capabilities and the viability of legitimate business models. It was emphasized that such a delimitation of prior monitoring would be a way to strengthen and optimize efforts to implement the STF's decision, as opposed to the current scenario of erratic demands from different authorities (which currently adopt their own, mutually discrepant interpretations of what the duty of care entails) and a level of unpredictability that undermines platforms' ability to respond adequately. It was also stressed that this delimitation is crucial to prevent fundamental rights from being disproportionately weakened, especially given the risks that prior monitoring poses to privacy and freedom of expression.

From the standpoint of *marketplace* platforms, such moderation would directly affect the income of economic actors of different sizes, and could even expose the platform to legal retaliation over removal decisions, encouraged by the duty of prior monitoring as currently designed, which lacks more objective parameters.

systemic risks vs. systemic failure

Some participants at the second workshop also voiced concern about the interpretation of concepts introduced by the STF's decision and developed further by the decrees regulating the MCI: the notion of systemic risks in Article 16-C added to Decree 8,771/2016. Whereas in statutes such as the ECA Digital (Brazil's Digital Statute of Children and Adolescents), given the sectoral nature of the rule itself, the risks to be mitigated and addressed through

moderation include content capable of harming children and adolescents, participants pointed out that it is not explicit which systemic risks Article 16-C refers to. The comparative landscape also included a reference to the European *Digital Services Act* and its reference to legal interests to be protected, such as electoral processes and the protection of children, leading to the conclusion that there is an interpretive need to anchor this risk detection and mitigation activity, lest the provision be used discretionarily in different situations.

Another point of tension was the fact that Article 16-C refers broadly to the detection and mitigation of risks, without mentioning any kind of differentiation by sector. Participants at the second workshop therefore questioned the rationale and benefits of a universal and homogeneous risk mitigation obligation, given the structural differences between the business models underpinning platforms. In practical terms, it was pointed out that the risks seen in e-commerce are vastly different from those seen on social media platforms, for example.

protection of vulnerable groups

There was consensus among participants that operationalizing the duty of care and developing risk metrics must also incorporate the dimension of the differentiated vulnerability of socially minoritized groups. Content that systematically victimizes such groups was cited as a category requiring closer scrutiny by platforms, given the recurrence and severity of the harms it produces.

In this regard, a layered moderation approach was suggested, one that recognizes the specific risks faced by activists, journalists and historically vulnerable groups, without this implying the adoption of arbitrary criteria or the suppression of legitimate speech. The complexity of the topic was illustrated by the example of the non-consensual dissemination of intimate images, in which unlawfulness lies not in the content itself but in the absence of consent, thus requiring tools for contextual analysis.

It is worth mentioning, in this regard, the approach of Decree No. 12,976/2026 and its focus on combating violence against women in the digital environment. The decree establishes specific layered moderation duties for this category of content, including shorter removal deadlines, the obligation to act on its own initiative in cases of coordinated attacks, and specific safeguards for intimate content generated or manipulated by artificial intelligence, reinforcing the importance and timeliness of the approach proposed by participants and showing that building differentiated parameters for groups in vulnerable situations remains a matter of concern in the regulatory sphere.

With regard to this decree, workshop participants highlighted the need to give greater density and delimitation to the normative concepts aimed at combating violence on the internet. Participants' attention focused on the wording of Article 3, item II, which sets out a broad definition of violence associated with a merely illustrative list of conduct and with strict deadlines for intermediaries to act. In this context, two distinct views emerged on the practical and regulatory impacts of the provision.

On one hand, it was argued that the regulation had the legitimate purpose of consolidating essential safeguards, ensuring that the protection provided by Brazilian law effectively extends to the digital space. From this perspective, providing a catalog of scenarios seeks to respond to challenges such as the use of artificial intelligence tools to manipulate intimate images and the occurrence of harassment on social media. The systematization of these forms of conduct was understood as a reference point to guide the rapid identification of serious attacks, incorporating offenses already provided for in the legal system.

On the other hand, concern was expressed about the lack of definition for categories such as coordinated attacks and conduct targeting women with a public profile. It was argued that the breadth of such concepts may lead platforms to adopt a preventive removal posture to avoid the risk of penalties, with negative repercussions for freedom of expression. In scenarios involving news coverage, political debate and scrutiny of elected officials, the absence of strict guidelines for distinguishing unlawful offense from forceful criticism was identified as a source of uncertainty capable of leading to the suppression of legitimate expression.

what does the civil liability regime look like in e-commerce?

Within the duties of diligence involving proactive actions by internet application providers, one of the main topics discussed among participants was the confusion generated by the legal regime applicable to e-commerce and *marketplaces* established by the STF's decision.

As a source of confusion, it was noted that the thesis set by the Supreme Court may contain an antinomy, in that it states that strict liability of internet application providers will not be admitted following the declaration of partial and progressive unconstitutionality of Article 19 (item 12 of the STF thesis), while at the same time establishing that internet application providers operating as *marketplaces* will be civilly liable under the Consumer Defense Code (CDC, Law No. 8,078/1990) (item 7 of the thesis). Under this interpretation, attributing strict liability to *marketplaces* would also entail implementing a specific regime of prior monitoring or filtering of content, which is incompatible with the content of the STF thesis and is not even provided for in the body of law indicated by the court, namely the CDC.

Workshop participants pointed out that referring all aspects of the liability of *marketplaces* to the CDC would, in practice, subject them to a strict liability regime, which directly contradicts the very caveat in item 12 of the same thesis. These participants argued for ruling out strict liability for providers when the elements specific to consumer relations do not apply to the case.

Thus, as a possible way forward to promote clarity and legal certainty in interpreting the STF thesis, the workshop discussion indicated the need for judicial decisions that can distinguish cases typically covered by consumer legislation from cases specifically governed by the Marco Civil da Internet rules addressed by the STF's decision, particularly regarding the control of content published by third parties. One criterion proposed by participants was to distinguish strictly consumer harms, linked to logistics, delivery or paid advertising problems, on one hand, from content harms, on the other, that is, those linked to the sphere of speech.

The main conceptual and practical gain of this division would be a clear separation between liability regimes, making it possible to organize the regimes proposed by the STF's decision and detailed by the decrees regulating the MCI. Along these lines, workshop participants highlighted that formally consolidating this type of distinction would bring predictability to its application by the different levels of the Judiciary and greater clarity for the private sector and civil society as to when, how and under what circumstances the liability regimes operate, a notable need given the plurality of liability regimes established by the STF's decision.

In the same vein of discussion about the main risks to e-commerce and the digital economy, participants reported numerous cases in which providers find themselves required to identify unlawfulness or technical irregularities in products whose assessment is challenging even for the specialized state bodies themselves.

In these situations, authorities would be delegating to private actors a State oversight activity (which the State, in turn, is better equipped to carry out through such specialized bodies and through access to information reported under numerous laws, embodied, for example, in ancillary tax obligations). It was reported that the situation is aggravated by the spreading understanding that e-commerce platforms would be subject to a duty of prior monitoring of each and every ad or piece of advertising content – thus including more than the prior monitoring of ads offered or boosted for payment, as prescribed by the text of the STF thesis and by Article 16-K of Decree 8,771/2016.

As addressed in the following sections, the result identified was a scenario of significant legal uncertainty, in which platforms are driven to anticipate the positions of multiple authorities under threat of later challenge, running into gray areas and matters of high technical complexity.

What was observed, therefore, was an incentive to remove content or prevent ad sales in case of uncertainty, which favors over-removal and threatens the plurality of speech (and, in the commercial context, of sales) in the digital environment. The criticism of this phenomenon in e-commerce was underscored by the fact that preventive or wrongful removal, beyond infringing the user's freedom of expression (as in the wrongful removal of content posted on social media), can economically harm the advertiser, affecting the user in material and

concrete terms. The example of a micro-entrepreneur or small seller harmed by unjustified removal was cited by some participants to illustrate a scenario in which legal interests as significant as freedom of expression, the protection of competition and free enterprise are affected by unforeseen impacts of uncertain regulatory action. It is important to note, in this regard, that while large companies and sellers have channels to reverse or challenge wrongful suspensions, small merchants tend to be more affected by "false positives" generated by preventive filters, directly affecting their livelihood.

II. extrajudicial notices

The change in the liability regime of digital platforms promoted by the STF's decision had direct implications for the functioning of the *notice and takedown* regime. Whereas Article 19 of the Marco Civil previously established a relatively clear model for triggering platform action, the new framework raised questions about who may send notices, under what conditions and with what consequences, both for platforms and for the users affected.

Workshop participants debated the new nature of the MCI's Article 21 regime and the extent to which it would amount to a new “general” platform liability regime. A central issue was the uncertainty of this transposition, given that Article 21 was designed specifically to govern the unauthorized disclosure of intimate content, and its application to other offenses raises still unresolved questions, such as who would be entitled to send extrajudicial notices in contexts different from the one for which the regime was originally designed. A case identified as especially controversial is that of content linked to hate speech or of an anti-democratic nature, whose unlawfulness produces diffuse and collective harm, without a clearly identifiable individual victim able to play the role of notifier envisaged by the logic of Article 21.

After the Executive Branch published the implementing decrees, the discussion delved into the nature of the notice regime proposed by Articles 16-D and 16-E of Decree 8,771/2016.

On one hand, participants noted the understanding that the paradigm presented by the new legal text is one of *notice and action*, that is, a system in which sending an extrajudicial notice by a victim or other entitled party does not necessarily result in the content being taken down; rather, the platform must take steps to investigate its lawfulness and take a position on the matter, and may keep the notified content without becoming liable if there is reasonable doubt. In other words, platforms are not obliged to remove each and every piece of content that constitutes the subject of a notice, and the concept of reasonable doubt plays a fundamental role in ensuring that platforms can show they did not remain passive in the face of potentially illegal content.

Another group of participants, however, acknowledged that, without concrete parameters for establishing what constitutes reasonable doubt, this provision may be rendered meaningless and, in concrete and practical terms, establish a burden to remove content, shifting the scenario to a *notice and takedown* regime along the lines discussed at the first workshop.

Thus, given the technical difficulty of assessing whether content is lawful, the uncertainty itself regarding the standards of conduct for providers, and the pressure to comply with extrajudicial notices, even contested ones, several workshop participants emphasized their concern about a scenario in which content removal tends to become providers' default response. More specifically, the concern expressed is that the removal effort may go beyond improper content and produce *over-removal*, leading to the removal of legitimate content as well. Participants highlighted the potential harm of this dynamic to freedom of expression, which should be protected and pursued even in borderline or hard-to-classify cases (for example, journalistic criticism of content involving hate speech, since the indiscriminate removal of all posts carrying a given piece of content could suppress precisely the speech that denounces it).

In this regard, concern was raised that providers guided primarily by legal risk mitigation might be led to preventively remove minimally questionable content, without an adequate assessment of its purpose and context, effectively inverting the logic that should govern the relationship between content moderation and the preservation of public debate. Civil society representatives also stressed that the effects of over-removal are strongly felt by activists and journalists, whose content, due to its controversial or whistleblowing nature, is particularly vulnerable to inaccurate classification by automated moderation systems that fail to capture the purpose or context of speech.

weaponization of extrajudicial notices for abusive purposes

Given that the new regime expands the range of situations to which the notice regime applies, the discussion included questions about the absence of barriers preventing economic actors

from using extrajudicial notices to have their competitors' lawful ads removed, on the pretext of combating piracy or technical irregularities, for example. In light of this, concern was expressed about how to prevent natural and legal persons from weaponizing the notice mechanism for purposes unrelated to prohibiting improper content. It was stressed that any solution in this regard must also take into account that requests with anti-competitive purposes may originate not only from a competing company as a legal entity, but even from individuals, trade associations and other organizations used for that end.

In addition to anti-competitive instrumentalization, participants also voiced concern about other types of weaponization to which the expansion of the notice regime makes providers susceptible. In particular, they highlighted the risk of natural or legal persons reporting ads purely on grounds of preference or ideology, as a way of expressing their opposition to the sale of that product. As raised at the workshops, cases like these should not constitute a sufficient basis for removing content and consequently preventing the sale, but, in practice, given the broadening of the situations in which the regime applies, over-removal may occur. Thus, taking the literal interpretation of the new regulatory framework to its limit, offers of lawful products could be suppressed merely because a single consumer dislikes them.

In this regard, participants also highlighted the risk that content removal requests may conceal other conflicts of interest that are not always immediately visible. One example cited was content involving minors, in which the narrative of "child protection" is sometimes invoked by legal guardians requesting takedown, but whose real interest may be to preserve the family's own reputation in the face of criticism of how they expose their children on platforms. In such cases, the requested removal might not correspond to the best interests of the minor, but to the convenience of those who originally authorized and benefited from their exposure. This dimension of the debate reinforced questions about the limits of providers' role as executors of removal requests in contexts where identifying the legitimate interest to be protected is, in itself, a complex task.

challenges in notices sent by public authorities

Linked to this scenario, participants also identified a trend toward expanding the scope of extrajudicial notices issued by government bodies. The shared perception was that the STF thesis, combined with the guidelines of the new regime that are still being consolidated, has created an environment of interpretive ambiguity that, in practice, progressively broadens the reach of removal requests by public officials. In this context, participants reported situations in which notices extend to content whose unlawfulness is neither evident nor uncontroversial.

Added to this perception was an operational concern regarding the form in which some of these notices have been presented. Participants reported that government actors have sent lists containing hundreds of items to be removed, often generated from keyword searches. This method, although understandable given the volume of content circulating on platforms, prompted participants to question the context in which the identified keywords actually occur and, consequently, the appropriateness and legality of all the requested removals.

An additional point of tension reported by private-sector participants concerns challenges in dialogue with notifying authorities, whether technical in nature or not. When faced with non-specific or even possibly mistaken removal requests regarding content deemed improper, providers must contest such requests on the grounds that the content in question is lawful and, moreover, that its removal would be harmful and unjustified to their users (including those who are advertisers, in the case of *marketplaces*).

However, according to participants, such challenges rarely receive reasoned responses from the requesting bodies (in some cases, there is no response at all), thus placing providers before a practical dilemma. On one hand, they could keep the content and run the risk of being held liable. On the other hand, they could remove the content preventively, at the risk of curtailing the rights of the user who published it (their expression or their ability to offer lawful products on a *marketplace*, for example) and, ultimately, of also being held liable in a lawsuit brought by the user whose post was removed.

Faced with this dilemma, these participants reported that platforms often feel pressured to opt for preventive content removal, which threatens the preservation of legitimate content and incurs the already mentioned risk of over-removal, another central theme of the debate.

sharing data with public authorities

In the discussions on notice flows and cooperation with public authorities, participants debated how to interpret the new Article 16-H of Decree No. 8,771/2016, which entails forwarding information on criminal conduct to authorities. Indeed, the article is an innovation compared to the STF thesis on leading case no. 987 which received attention at the second workshop.

Some participants suggested that the mechanism for extrajudicial data sharing should draw inspiration from established experiences in comparative law, such as the U.S. model of the *National Center for Missing and Exploited Children*. The purpose of this approach would be to authorize the direct and swift forwarding of data solely in emergency situations, characterized by an imminent risk to life and physical integrity, with threats of school shootings cited as a paradigmatic example of such emergencies.

Workshop members expressed reservations about the broad and indeterminate wording of the provision, warning of the risks arising from a flexible interpretation of the article. It was argued that the absence of strict procedural guidelines could turn a tool meant for emergency and extraordinary response into a routine of active monitoring and preventive surveillance by platforms. From this critical perspective, it was reiterated that the transfer of authorship information and connection records without a court order must be subject to objective regulatory parameters, preserving the constitutional rights to privacy and due process of law.

“reasonable doubt”

The concept of reasonable doubt was widely valued by participants at the second workshop as an indispensable instrument for balancing and operationalizing the new notice regime. In the *notice and action* model, reasonable doubt was described as a proportionate safeguard that prevents an extrajudicial notice from turning into an obligation of immediate removal. In the context of e-commerce, it was argued that delimiting this concept is essential to curb incentives for the wrongful removal of lawful offers, preventing private intermediaries from being forced to assume adjudicative functions on controversial matters or on matters that require technical-regulatory expertise, and ensuring that platforms can assess the merits of complaints with due legal caution.

With regards to the practical application of reasonable doubt, the workshop debates outlined two main areas in which the concept applies. The first covers situations of technical uncertainty, in which verifying the regularity of an offer in a post depends on information held by public authorities. Examples cited, as in the discussion above on the duties of authorities, included real estate rental listings whose compliance requires checking municipal zoning and housing policy data, as well as the offer of products subject to oversight by regulatory agencies that do not maintain structured repositories for consultation. In these situations, the lack of information prevents the provider from attesting to (un)lawfulness by its own means, legitimizing keeping the content up under the protection of reasonable doubt.

The second area comprises scenarios of legal and contextual uncertainty, with an emphasis on posts associated with freedom of expression, news stories and expressions of public criticism. It would thus cover disputes involving possible factual errors or conflicting interpretations of unlawfulness that are giving rise to divergent case law, for example. Indeed, participants argued that the existence of disagreement within the Judiciary itself regarding the unlawfulness of certain speech is evidence of reasonable doubt, serving as a guideline to prevent platforms from preventively suppressing legitimate expression in response to extrajudicial notices.

requirements for unequivocal identification of content

In the discussions on how extrajudicial notices are materialized, the workshop explored a growing tension over the role of the URL (*Uniform Resource Locator*) as an instrument for identifying unlawful content. Traditionally conceived as the basic unit of reference / gold standard in removal orders, the URL has been progressively challenged and replaced by other methods of identifying content, such as hashtags, keywords and hash matching. Participants

described this transition as a structural change in the way public authorities have been materializing their requests.

Although it is acknowledged that the change has the valid aim of responding to the speed and scale at which unlawful content can be replicated in today's digital environment, it is notable that the new methods of identifying content would operate with a considerably greater degree of imprecision and uncertainty for the user. In the case of *hashtags*, for example, this is a technique that can be freely used by any user, on any type of content, to increase the reach of their post and connect it to a certain audience, even if there is no relationship between the hashtagged words and the actual content of the post.

For some workshop participants, this discussion takes on even more concrete contours in e-commerce, since the absence of URLs in orders or notices makes it significantly more complex to single out an irregular product among thousands of similar ads, turning compliance with such removal orders into a difficult task.

Furthermore, a central consequence identified by certain participants regarding the progressive abandonment of the URL-based model is the hollowing out of contextual analysis for content removal. In this regard, they point out that the URL makes it possible to understand the exact contextual surroundings of the post, which is essential for an analysis proportionate to the conduct and content in question. These participants noted that the topic gains special relevance in light of item 3.2 of the STF thesis, according to which "identical content" to that already recognized as offensive by a court decision subjects platforms to removal upon judicial or extrajudicial notice.

What criteria would determine whether a piece of content is identical to another? In the view of certain participants, this open-ended formulation risks neglecting nuances inherent to speech. The same content may take on radically different meanings depending on the framing and perspective in which it is presented, whether in a journalistic, critical, ironic or *counter-speech* register. In the case of *marketplaces*, the mere coincidence of keywords between two products does not guarantee that they are identical and may even indicate opposite intentions. It is the contextual and detailed analysis of each ad that makes it possible to distinguish, for example, between a product that glorifies a crime and another that denounces it – differences that are often subtle. Discarding context makes this distinction impossible, which is a central concern for participants given its potential to suppress legitimate discussions on platforms, raise the rate of false positives and produce significant side effects on freedom of expression.

Beyond the diagnosis of the erosion of the URL-based model, participants turned to a forward-looking question about which criteria should guide the identification of unlawful content in a rapidly changing digital environment. If the shift to *hashtags*, keywords and *hash* matching already represents a considerable challenge in terms of precision and contextualization, the advance of generative AI environments makes this problem even more complex.

In these new services, the very concept of content identifiable by a discrete unit of reference progressively loses meaning, insofar as content may be generated dynamically, in a personalized and non-repeatable way, making the task of identification even harder. Although ideas were sketched out as new approaches, participants signaled that any "overcoming" of the URL paradigm requires building new identification criteria that are, to some degree, precise, technically feasible and legally sound.

The workshops also reflected tensions between the expectation of a swift response to removal orders and the need for a structured due process in verifying whether content is unlawful, with safeguards for users' individual rights and guarantees. Participants underscored the legal dimension of this tension, insofar as the way removal requests are formulated and transmitted has a direct bearing on the quality and proportionality of the removals actually carried out.

In this context, participants discussed how the burden of accurately identifying the content deemed unlawful and the scope of the measures requested should be distributed among the different actors involved in the notice process. The concern expressed was to prevent this task from falling exclusively on a single link in the chain, especially on the provider responsible

for processing notices or complying with court orders. Accordingly, models that more clearly distribute the duties of identification, justification and execution among the actors involved in the process, from the formulation of the request to its fulfillment, were considered relevant.

III. governance, enforcement and the role of the public sector

Throughout the workshop, a set of perceptions emerged about the growing prominence of different public bodies in digital matters and about the increasingly major role that various regulatory bodies have been called upon to play, a picture that highlights the opportunity, and even the need, to strengthen forums for coordination capable of aligning and guiding this currently diffuse activity.

Within the Federal Executive Branch, it was noted that this activity has been distributed among bodies such as the National Data Protection Agency (ANPD), an institution whose competences are expanding with the entry into force of the ECA Digital and of the decrees regulating the MCI themselves, the Administrative Council for Economic Defense (CADE), the National Telecommunications Agency (Anatel) and the Office of the Union's Attorney General (AGU), further extending to consumer protection bodies and the various ministries and secretariats of the Executive Branch, reflecting the cross-cutting nature the agenda has acquired in recent years and its institutional plurality.

Added to this scenario is the involvement of the Judiciary, especially the STF, as exemplified by the ruling on Article 19 of the MCI itself, and the Superior Electoral Court (TSE), which has taken on a prominent role in regulating digital content, at times reaching issues that go beyond the strictly electoral domain. Different levels of the Judiciary were also mentioned as already frequently called upon (a trend possibly amplified by the new regulatory regime) to rule on digital issues of notable technical complexity. The elimination, by the STF's decision, of the requirement of a breached court order to establish individual harm, combined with the ease of filing lawsuits at scale with the help of AI tools, were identified as new factors likely to have significant effects in the coming years.

ANPD as the authority for the digital environment: tensions and risks

One of the major points of debate among participants at the second workshop was the institutional governance of the new regulatory regime and the *enforcement* of legal rules. Broadly speaking, two views on the new moment following the STF's decision and the decrees issued by the Executive can be described.

On one hand, some participants interpreted the designation of ANPD as a possible driver of a more centralized, less fragmented regime with greater legal certainty (the fact that ANPD's designation followed the trend already established in the regulation of the ECA Digital, and allowed the new Marco Civil da Internet rules to be harmonized with existing legal frameworks, was cited as a positive element). From this angle, therefore, the centralization of regulatory and oversight competences in ANPD operates as a key move in ensuring legal certainty and the effectiveness of the rules, albeit within a gradual process of building regulation and institutional legitimacy.

Other participants, however, were skeptical of the idea that *enforcement* by ANPD could put an end, in the short or medium term, to a trend of institutional fragmentation and overlapping action by public authorities. The period immediately following the approval of the LGPD was cited as a moment of plurality of public actors in action that shows how lack of coordination can indeed occur. The case of the Anti-Corruption Law (Law No. 12,846/2013) was cited as an example of a legal regime that, allowing for a plurality of actions and institutional avenues for its implementation, went through an initial period of lack of coordination, which was only addressed through technical cooperation among different public institutions.

Another example raised by participants who held a more optimistic view of the new institutional architecture was that of CADE and the development of the Brazilian competition system. The conclusion presented was that, despite the challenges in building the Council, CADE became, over the years and through the efforts of its technical staff and political leadership, an important administrative authority whose legitimacy is respected by other public actors, especially the Judiciary, offering a possible path for ANPD to follow. Other participants, however, objected to this interpretation, pointing out that CADE's institutional trajectory was marked by various challenges, including jurisdictional conflicts with the

Judiciary and the review of the agency's decisions by other bodies, showing that fragmentation was also a notable trend in the institution's historical development. This critical view was combined with the observation that, even today, ANPD appears as a co-defendant, alongside digital platforms, in public civil actions, indicating a possible challenge to the agency's regulatory and decision-making legitimacy.

Thus, institutional fragmentation was seen by some participants as a real and serious possibility, suggesting that multiple authorities, especially decentralized ones present at different levels of the federation, such as Public Prosecutor's Offices and Procons (consumer protection agencies), may overlap in applying the Marco Civil, with dissonant and varied interpretations of the various indeterminate concepts established in Brazilian law following the Supreme Court's decision on leading case no. 987.

In the case of the National Consumer Protection System – whose decentralized activity through hundreds of municipal and state Procons and civil society entities stems from normative competences consolidated by the Consumer Defense Code and recognized by the amendments from 2026 to Decree No. 8,771 of 2016 –, participants highlighted a complex overlap between the protection of rights provided for in the Marco Civil da Internet and consumer legislation. Without mechanisms ensuring some form of deference to the technical parameters set by ANPD, the digital ecosystem may be exposed to simultaneous and conflicting demands regarding content moderation.

Another point of institutional tension highlighted by participants concerns the asymmetry between the dynamics and timeframes of action of the public bodies involved in digital governance. On one hand, ANPD operates on a medium- and long-term regulatory horizon, grounded in a regulatory agenda, collegiate deliberative procedures, calls for input and public consultations aimed at the gradual consolidation of technical understandings. On the other hand, top-level bodies of the Administration, such as the Ministry of Justice and the Office of the Attorney General of the Union, operate on a vertical executive horizon focused on immediate responses to concrete security crises or matters of public repercussion.

Participants thus emphasized that the absence of formal coordination channels between the more emergency-driven management carried out by public actors and ANPD's structural regulatory planning can also generate legal uncertainty and antinomies, subjecting providers to divergent operational demands issued simultaneously by the State itself.

views on a regulatory deference agenda

The tension over the degree of regulatory centralization or fragmentation led to debates on the relevance of “regulatory deference”, embodied in the construction of the ANPD's legitimacy across different roles. This debate raised significant questions about the degree of acceptance its technical guidance will receive from other state actors, for example.

It was noted that the agency's choice to issue guidance documents rather than strictly normative resolutions, as occurred in the initial implementation of the ECA Digital, raises questions about the binding force of these instruments before the Judiciary. The central concern lies in the possibility that first-instance judges and small claims courts, for example, may disregard the technical guidelines issued by the authority, maintaining an autonomous and fragmented interpretation of what constitutes due diligence in fulfilling the duty of care. Should this scenario materialize, the question arises as to which due diligence the law actually prescribes: the one required by judicial institutions or the one recommended by ANPD in an administrative capacity.

Given the need to stabilize the regulatory environment without waiting for the long maturation periods historically observed in other agencies, participants discussed tools capable of accelerating ANPD's regulatory and decision-making legitimacy. Participants converged on the view that merely issuing advisory recommendations may not be sufficient to anchor the agency's authority over technical standards, risk detection and mitigation methodologies and diligence measures, making it necessary to mobilize regulatory instruments that connect these requirements and technical standards to pressing public needs.

Various solutions to this problem were raised at the workshops, with divergence among sectors. Among the proposals, the implementation of experimental regulatory environments,

known as regulatory *sandboxes*, stood out, for example. Others advocated for effectively regulating independent researchers' access to data, as provided for in Article 31 of the ECA Digital. It was argued that the creation of a scientific community capable of auditing regulators' activities and transparency reports and examining moderation systems would lend empirical support to the agency's decisions, strengthening its authority before the Judiciary and allowing civil society and academic entities to also act as important actors in the process of building ANPD, rather than being alienated from its development.

public data as an input for providers to meet their duties

Participants identified certain aspects of public-sector action as an indispensable condition for the viability of the liability regime and for platforms' effective exercise of due diligence. Participants reported that compliance with oversight duties regarding ads, offers and commercial registrations sometimes runs up against the opacity of the information provided by public authorities themselves. Regulatory agencies and bodies, such as Anvisa, Inmetro, Susep and ANTT, routinely publish administrative and enforcement documents without standardization or technical attributes that would enable an understanding of regularity (or irregularity) and allow the data to be processed at scale by providers' moderation tools, for example.

Thus, in e-commerce, the unavailability of open and structured databases was seen as capable of inducing practices of preventive and indiscriminate removal of ads. Given this bottleneck, it was argued that the State has an obligation to make available official, qualified public repositories⁶ or equivalent consumption formats that allow those subject to legal obligations to analyze the lawfulness of content. It was emphasized that, in the absence of functional public databases that can be consumed by automated systems, the technical uncertainty resulting from the State's omission may be recognized as a case of reasonable doubt, creating an exception to the duty to take content down. Along the same lines, participants stressed the importance of channels, systems or infrastructures that enable dialogue between platforms and authorities and the exchange of information and data needed to take down unlawful content.

On this topic, one example mentioned by some participants was the technical partnership entered into by the Central Bank of Brazil, the Ministry of Justice and digital platforms for screening financial ads, examined as a relevant case study of public-private dialogue and discussed at various points during the second workshop. Under this arrangement, the authority provides a structured list of institutions authorized to offer financial products, allowing the platform to use this official database as a prior validation criterion and entry filter for the publication of certain types of advertising on the internet.

Although the model was recognized as a practical advance in combating fraud and an important example of public-private cooperation, some interventions urged caution about expanding this model indiscriminately. Concern was raised about the feasibility of requiring similar prior verification for all categories of digital advertising, given the scale and diversity of products and services offered online. Moreover, it was stressed that smaller intermediaries lack both the technical infrastructure and the institutional channels needed to enter into specific agreements with state bodies.

Another case identified as problematic was the lack of infrastructures and systems for dialogue between platforms and public authorities to analyze the lawfulness of content and ads made available online. One example cited in this regard was the lack of technologies for communicating and sharing information in cases involving the rental of properties classified as Social Interest Housing (HIS) in the city of São Paulo: in the view of some participants, the absence of contact systems between platforms and public authorities contributed significantly to the circulation of irregular listings.

⁶ Several suggestions were raised on this topic, with particular concern for integrating government data into providers' moderation systems at scale.

IV. recommendations and convergences

Although some of these ideas have already been outlined throughout the document, this section develops the main recommendations on which the debate converged, as a guide for next steps in the regulatory and judicial discussion on the duties and liability of internet application providers and their impact on the digital economy and e-commerce.

self-regulation for technical standardization

The possible contributions that self-regulation initiatives could make to building legitimate technical standards for assessing the diligence of internet application providers emerged as one of the proposals with the greatest structural reach in the debate at the first workshop, and was discussed both in its institutional dimensions and in its practical implications for content moderation and platform liability. Some participants referred to the experience of the National Council for Advertising Self-Regulation (CONAR)⁷, whose ability to produce technical decisions recognized by the Judiciary was identified as a relevant reference.

The rationale of the proposal is that collective initiatives, endowed with sectoral legitimacy and specialized technical knowledge, would help fill certain gaps that remain after the STF's decision, offering a reference body for the interpretation and application of the duty of care, for example. Compared with each platform's individual terms of use, a body of this kind would give moderation guidelines greater social legitimacy and accountability.

Central to the proposal was the view that such initiatives should not adopt a monolithic structure, but rather be organized around sectoral annexes addressing the specificities of distinct segments, such as e-commerce, social media and search engines. This architecture would allow each sector to be assessed according to criteria suited to the nature of its services and the risks specific to it, avoiding the uniform application of parameters to very different realities. For the e-commerce sector in particular, this differentiation was seen as especially relevant, given that the challenges of moderating ads and products present their own technical complexities, distinct from those faced by platforms of a predominantly communicative nature.

In addition to its normative function, the proposed body would play a major role in producing administrative case law capable of guiding the Judiciary, especially in lower court decisions. In this regard, building a body of decisions and guidelines on complex cases, the so-called "gray areas", such as defining irregular products or classifying borderline speech, was presented as one of the main expected gains of the model, insofar as they would establish more objective parameters for assessing compliance with legal duties and possible systemic failures of the systems developed by platforms.

regulatory deference agenda for legal certainty

Given the institutional architecture chosen by Decree 12,975/2026 of centralizing regulatory oversight competences, transferring the issuance of guidance on the *enforcement* of the Marco Civil da Internet to a regulatory agency (ANPD), workshop participants considered it essential to propose measures to ensure the agency's legitimacy in its new functions and to avoid overlapping, contradictory orders from a plurality of public authorities.

1. **Sandboxes.** One measure suggested for this purpose is the implementation of experimental testing environments, known as regulatory sandboxes. These spaces for controlled experimentation allow for the empirical assessment of how new business models, artificial intelligence tools and moderation solutions work, enabling the production of practice-based rules and the calibration of safeguards before definitive

⁷ CONAR is a non-profit non-governmental organization founded in 1980 following the initiative of entities representing the Brazilian advertising market, in response to the threat of federal legislation being approved imposing prior censorship on advertising; it emerged as a voluntary product of the advertising market's self-organization. A noteworthy normative instrument of CONAR is the Brazilian Code of Advertising Self-Regulation, which sets out the ethical rules applicable to the creation and publication of advertisements in any medium. It is important to note that CONAR is not a state body and does not have coercive power in the strict sense.

resolutions are issued. In the view of some participants, tests and experiments of this kind could serve as direct safeguards of legitimacy for any general and binding rules issued by ANPD.

2. Research and expertise. Another measure highlighted in the discussions, with some convergence, was bringing ANPD closer to researchers and experts. Beyond expanding social oversight, the formation of this autonomous research network strengthens the technical legitimacy and authority of ANPD itself before other public bodies. Grounding regulatory guidelines and opinions in empirical assessments produced by academics lends support to the administrative authority's decisions, rebutting claims of arbitrariness and shielding its acts against annulment in lawsuits. The continuous production of scientific evidence turns regulation into a dynamic and participatory process, in which institutional deference to the agency's positions comes to be sustained by the technical consistency of the assessments produced by the research ecosystem.

3. Institutional coordination and articulation. In addition, workshop members suggested broader multi-stakeholder debates on building an agenda of this kind, with the aim of pursuing institutional coordination and instruments capable of reducing legal uncertainty and conflicting demands between different actors within the Executive Branch, between different levels of the federation, or disputes arising from interaction with the Judiciary.

compliance agenda and actions by public authorities

The workshops pointed to the importance of public authorities building an adaptation agenda, with proactive actions aimed at leading the construction of a new regulatory environment, with new roles and responsibilities. This agenda comprises the following discussions:

1. Transparency in requests issued by public authorities. Participants noted the asymmetry between the transparency obligations imposed on platforms and the absence of equivalent requirements for public actors themselves. While the STF's decision requires platforms to publish transparency reports, there is no analogous provision requiring public bodies or other sectoral entities to disclose data on the volume, criteria and outcomes of their notices and removal requests.

Participants suggested that incorporating this requirement into the regulatory model would help balance the system, making the notice process more traceable and creating the conditions for a well-informed assessment of its effectiveness and legitimacy, in particular by allowing scrutiny of the grounds for notices sent by public bodies. Beyond a formal parallelism, institutional transparency on behalf of public actors would help align expectations among the different stakeholders, fostering a more qualified institutional dialogue and allowing platforms to develop moderation architectures that are, from the outset, better aligned with the trends and priorities guiding public action. This alignment could even result in greater effectiveness in addressing unlawful content disseminated at scale.

2. Production of public databases and standardized protocols for engagement with the public sector. Participants recommended that the enforceability of automated due diligence routines for offers, ads and content be expressly conditioned on public authorities making available structured official repositories, accessible through machine-readable application programming interfaces. Imposing preventive large-scale checking duties relies on the existence of fully comprehensible – and preferably machine-readable – state data on health registrations, technical certifications, license revocations, sectoral authorizations, etc. Accordingly, in the absence of such databases, the technical impossibility of automated validation should be recognized as a case of reasonable doubt.

As a complement to this data infrastructure, it is suggested that permanent institutional channels for operational dialogue between public authorities and digital platforms be established. This integrated cooperation environment should enable regulatory bodies to share qualified information and to respond to queries from

providers regarding offers or content of questionable lawfulness. Establishing this two-way communication flow reduces the risk of indiscriminate preventive blocking, which unduly harms users and advertisers, ensuring that content restrictions or removal measures are backed by clear state guidance consistent with the limits of private action.

3. Proportionality in voluntary partnerships. Participants also recommended that bilateral cooperation agreements entered into between providers and sectoral bodies be understood as specific and voluntary practices, not necessarily to be converted into a mandatory standard for all segments of the digital economy. Transposing this arrangement to all sectors of commerce and advertising proves, arguably, unfeasible, given the extreme heterogeneity of economic activities and the lack of centralized public registries for all categories of products and services.

From a competition perspective, and in light of the principle of proportionality set out in Article 16-P of Decree 12,975/2026, generalizing this model also has the potential to deepen asymmetries among market players. Entering into customized agreements and maintaining complex technological integrations require structure and contacts that may not reflect the operational capacity of smaller intermediaries, regional services and new entrants. Accordingly, institutional partnerships should be encouraged as cooperative experiences necessary to ensure the safety of certain segments, but not as a common regulatory yardstick for the entire digital ecosystem.

agenda for regulatory densification of the duty of care and diligent conduct

Common ground was also found on the need for greater density in key concepts of the new liability regime for internet application providers, especially regarding their duties of diligence. This agenda comprises the following points:

1. Duty of care. Participants stressed the importance of issuing more concrete and precise elements in compliance with the duty of care. To this end, it was suggested that its interpretation be anchored in the classic private law notion of objective good faith, drawing on the ancillary duties of loyalty, transparency, care and confidentiality, in order to favor the acceptance of the concept by the courts and to discourage the preventive takedown of content that raises any degree of uncertainty.

Alongside this guideline, there was a recognized need to discuss benchmark metrics for assessing platforms' diligence. There was perceived convergence on seeking to replace the expectation of "zero error" with a parameter compatible with the scale and complexity of digital operations, making it possible to distinguish an inert platform from one that, in good faith, seeks to improve its moderation mechanisms.

2. Qualitative aspects in the supervision of the duty of care. One point raised was that supervision of the duty of care should be sensitive to qualitative arguments, moving beyond a strictly quantitative logic based on the ratio between the number of notices received and the volume of content removed. As different participants argued, assessing compliance through purely numerical metrics creates inappropriate incentives for preventive over-removal, penalizing prudent screening practices and encouraging the automatic blocking of legitimate offers and user expression in order to achieve high response rates. Accordingly, accountability may also focus on the detailed disclosure of the methodologies used in moderation tools, internal governance guidelines and the procedural flows adopted for analyzing complex complaints.

By documenting the reasoning used in cases of reasonable doubt and the criteria followed when faced with controversial notices or notices weaponized in bad faith, providers also become accountable for fulfilling their duties and for their impact on users' expression. Accordingly, the assessment of the intermediary by the regulator and the Judiciary is then based on the consistency of decision-making processes and the soundness of the safeguards implemented, ruling out the characterization of systemic failure as a breach of an obligation of result.

3. Distinguishing consumer issues from content issues. In the discussion on the place of e-commerce and consumption, a typology was proposed: distinguishing, on one hand, strictly consumer harms, associated with logistics, delivery or paid advertising problems, to which the applicable CDC regime would apply, and, on the other hand, content harms, associated with the sphere of speech, to which the Marco Civil da Internet would apply, under the terms of the STF's decision. The division would thus function as an organizing criterion for courts, administrative bodies and different sectors of society, a particularly relevant step given the plurality of liability regimes established by the STF's decision.

agenda for detailing the extrajudicial notice regime

The new role of the extrajudicial notice regime following the issuance of the STF thesis on leading case no. 987 also gave rise to the recommendation of a more extensive regulatory agenda on the subject, along the following takeaways.

1. Standards of reasonable diligence in response to notices. Participants converged on the importance of issuing standards on the diligence expected of providers upon receiving notices. Situations were reported in which removals carried out within very short timeframes, on the order of twenty-four to forty-eight hours, or compliance with significant portions of extensive lists of removal requests had no mitigating effect whatsoever on the sanction, which was maintained as if the platform in question had taken no action at all on the requests. For example, it would be illogical to consider it unimportant whether a platform removed 0 or 90 pieces of content in response to a removal request for 100 items; the incentive system must therefore reflect this difference and distinguish compliance efforts.

In light of this, the recommendation articulated by participants was that good faith, speed of response and the technical effort demonstrated by platforms be recognized as relevant factors in grading administrative sanctions. In addition to bringing greater predictability to the regulatory environment, this would help encourage continuous investment in robust moderation architectures and consolidate a culture of effective compliance, benefiting platforms, oversight bodies and users alike, and explicitly giving sanctions the parameter of proportionality.

2. Protection mechanisms against abusive use of notices. Workshop participants converged on the idea of detailed regulation of Article 16-F of Decree 12,975/2026, establishing procedural filters aimed at curbing the instrumental use of extrajudicial complaints. The practice of mass notices sent with unfair intent by competitors or other actors acting in bad faith is an irregular strategy that requires legal checks in building the new Marco Civil da Internet regime. To mitigate this risk, it is suggested that admissibility criteria be established for notifications reporting unlawful conduct, preventing generic complaints from leading to the summary removal of content or products. Accordingly, workshop members proposed providing for legal consequences and sanctions for economic agents who repeatedly use reporting channels to distort the purpose of notices.

3. Trusted flaggers. Another proposal discussed during the workshop was the creation of a structured accreditation mechanism for qualified notifiers (commonly referred to in the international literature as *trusted flaggers*) also for the new regime established for the MCI (noting that this figure is already covered under ECA Digital). The proposal consists of granting a differentiated status to certain entities with expertise and legitimacy, whose notices would be processed by platforms as a priority and through dedicated channels, including commitments to faster response.

The underlying logic is that formally identifying these groups serves a dual function: (i) ensuring a basic standard of quality for complaints, by linking their origin to entities with recognized technical capacity, and (ii) deterring abusive behavior, by associating the differentiated status with more demanding accountability standards. In this regard, it was argued that greater capacity to influence the notice system must be matched by a proportional commitment to objective technical criteria that justify and legitimize this status.

Finally, the debate also included the idea of creating institutional tools that allow platforms to build a structured memory of the notices sent by these qualified actors. The idea is that the notice history of trusted flaggers on certain topics or categories of content could serve as a benchmark for analysis when the same topic is raised by ordinary notifiers, providing greater security and certainty to all involved.

4. Return to the URL and preservation of contextual analysis. As a counterpoint to the growing trend of identifying unlawful content through hashtags, keywords and hash matching, participants recommended returning to the URL as the preferred instrument for singling out content in removal orders. Conceived as the gold standard for such identification, the URL preserves the direct link between reference and content, giving back to the platform and the user the ability to clearly delimit the object of removal, as opposed to methods that operate by approximation and tend to unduly broaden the scope of orders, with particularly sensitive consequences for the e-commerce sector, where singling out an irregular ad among thousands of similar ones is unfeasible without this precise reference.

Related to this discussion, participants stressed the need to preserve contextual analysis in complying with removal orders, particularly in light of the "identical content" criterion introduced by item 3.2 of the STF thesis. Participants recommend that assessing whether content is identical should not be reduced to the formal coincidence of textual or visual elements taken in isolation, but should consider the framing, purpose and context in which each piece is published.

It was also emphasized that responsibility for accurately identifying unlawful content should be understood as shared among the links in the notice chain, from the formulation of the request to its fulfillment, and not as an exclusive burden on the executing platform, especially given the new challenges that generative AI poses to content identification.

This document sought to portray the current stage of the debate and to inform next steps on aspects of Brazilian digital governance closely connected to recent legal frameworks, from the perspective of multiple actors attentive to the effects of this decision. Any developments relating to the operationalization of the new rules, far from settling the issues raised here, reaffirm their relevance, since the regulatory framework is still under construction, and the topics debated by participants will continue to require attention, further study and dialogue among the many actors involved. The topics debated at the workshop and set out here have certainly not been exhausted, and their developments will need to be monitored and discussed further over time.