

UNIVERSIDADE ESTADUAL DO MARANHÃO
CENTRO DE ESTUDOS SUPERIORES DE BACABAL
CURSO DE DIREITO BACHARELADO

RONILSON SILVA DE OLIVEIRA

A RESPONSABILIDADE CIVIL DAS PLATAFORMAS DIGITAIS NO COMBATE AO DISCURSO DE ÓDIO: uma análise da constitucionalidade do Art. 19 do Marco Civil após o entendimento do STF (2025-2026).

Bacabal

2026

RONILSON SILVA DE OLIVEIRA

A RESPONSABILIDADE CIVIL DAS PLATAFORMAS DIGITAIS NO COMBATE AO DISCURSO DE ÓDIO: uma análise da constitucionalidade do Art. 19 do Marco Civil após o entendimento do STF (2025-2026).

Monografia apresentada ao Curso de Direito Bacharelado da Universidade Estadual do Maranhão para obtenção do grau de Bacharel em Direito.

Orientador: Prof. Esp. Raimundo Miranda Teixeira Mendes Neto.

Oliveira, Ronilson Silva de.

A responsabilidade civil das plataformas digitais no combate ao discurso de ódio: uma análise da constitucionalidade do Art. 19 do Marco Civil após o entendimento do STF (2025-2026) / Ronilson Silva de Oliveira. - Bacabal - MA, 2026.

64 f.

Monografia (Graduação em Direito Bacharelado) - Universidade Estadual do Maranhão, Campus Bacabal, 2026.

Orientador: Prof. Esp. Raimundo Miranda Teixeira Mendes Neto.

1. liberdade de expressão. 2. discurso de ódio. 3. dignidade da pessoa humana. 4. plataformas digitais. 5. Supremo Tribunal Federal. I. Título.

CDU: 342.727:004.738.5

RONILSON SILVA DE OLIVEIRA

A RESPONSABILIDADE CIVIL DAS PLATAFORMAS DIGITAIS NO COMBATE AO DISCURSO DE ÓDIO: uma análise da constitucionalidade do Art. 19 do Marco Civil após o entendimento do STF (2025-2026).

Trabalho de Conclusão de Curso
apresentado ao Curso de Direito
Bacharelado da Universidade Estadual do
Maranhão como requisito para obtenção
do grau de bacharel em Direito.

Aprovado em: 20/07/2026

Nota: 10

BANCA EXAMINADORA

Documento assinado digitalmente
 **RAIMUNDO MIRANDA TEIXEIRA MENDES NETO**
Data: 23/07/2026 09:00:30-0300
Verifique em <https://validar.iti.gov.br>

Prof. Esp. Raimundo Miranda Teixeira Mendes Neto (Orientador)
Universidade Estadual do Maranhão

Documento assinado digitalmente
 **MARIA DO SOCORRO PEREIRA ALVES**
Data: 24/07/2026 11:18:16-0300
Verifique em <https://validar.iti.gov.br>

Professor(a)
Universidade Estadual do Maranhão

Documento assinado digitalmente
 **PEDRO HENRIQUE GONCALVES CLEMENTINO**
Data: 23/07/2026 08:16:46-0300
Verifique em <https://validar.iti.gov.br>

Professor(a)
Universidade Estadual do Maranhão

AGRADECIMENTOS

Chegar ao fim desta etapa é olhar para trás e ver o quanto o caminho foi longo e desafiador, mas também cheio de aprendizados. Escrever este Trabalho de Conclusão de Curso representa o encerramento de um ciclo muito importante, e eu não poderia deixar de expressar a minha profunda gratidão a todas as pessoas que caminharam ao meu lado até aqui.

Primeiramente, agradeço a Deus por me dar saúde, força e tranquilidade nos momentos em que o cansaço e a ansiedade pareciam tomar conta. Sem essa luz, a caminhada teria sido muito mais difícil. Aos meus pais, dedico a maior parte desta conquista. Obrigado por cada sacrifício que vocês fizeram para me dar a oportunidade de estudar, pelos conselhos nas horas certas e por nunca deixarem de acreditar nos meus sonhos. O amor, a comida de domingo e o apoio incondicional de vocês foram o meu maior combustível durante todos esses anos.

Agradeço também aos professores do curso de Direito da UEMA, que compartilharam seu conhecimento com tanta generosidade e me ajudaram a entender o verdadeiro papel da justiça na sociedade. Deixo um agradecimento muito especial ao meu orientador, Professor Neto, que teve toda a paciência do mundo para ler meus textos, corrigir meus rumos e me guiar na construção desta pesquisa sobre um tema tão importante. Sem a sua dedicação e palavras de incentivo, este trabalho não teria chegado a um bom resultado.

Aos meus amigos, especialmente aqueles que dividiram comigo as angústias das provas da faculdade e as preparações para a OAB, levarei a nossa parceria para sempre na memória. Por fim, agradeço a todas as pessoas que torceram por mim de alguma forma. A realização deste sonho é de todos nós!

"A tolerância ilimitada leva ao desaparecimento da tolerância. Se estendermos a tolerância ilimitada mesmo aos intolerantes, e se não estivermos preparados para defender a sociedade tolerante do assalto da intolerância, então os tolerantes serão destruídos e a tolerância com eles."

— Karl Popper

RESUMO

O presente trabalho analisa como o crescimento da internet e a hiperconectividade mudaram profundamente a forma como a sociedade se comunica, trazendo à tona a necessidade urgente de discutir quais são os limites reais da liberdade de expressão dentro das redes sociais. O objetivo central desta pesquisa é repensar a extensão do exercício desse direito quando ele passa a ser utilizado para propagar o discurso de ódio, violando diretamente a dignidade da pessoa humana e os direitos das minorias. Para compreender essa questão, o estudo parte dos fundamentos da Constituição Federal de 1988, que, ao mesmo tempo que garante a liberdade de manifestação do pensamento e veda a censura prévia, consagra a inviolabilidade da honra, da imagem e da dignidade humana. É fundamental diferenciar a livre opinião do discurso de ódio, ressaltando que ataques motivados por preconceito não encontram amparo constitucional como manifestações legítimas. No Brasil, o principal diploma regulatório da rede é o Marco Civil da Internet (Lei nº 12.965/2014). Durante anos, a regra do artigo 19 dessa norma determinava que as plataformas digitais apenas poderiam ser responsabilizadas civilmente por conteúdos de terceiros caso descumprissem uma ordem judicial de remoção. Entretanto, esse modelo passou a ser alvo de severos questionamentos judiciais, sob o argumento de que favorecia a perpetuação de publicações nocivas e danos irreparáveis. O ponto de maior destaque desta pesquisa é a substancial alteração jurisprudencial ocorrida em junho de 2025, quando o Supremo Tribunal Federal redefiniu o alcance do referido dispositivo. Através da análise dos votos dos ministros, constata-se que a Suprema Corte estabeleceu um novo panorama normativo, impondo às redes sociais um dever de cuidado mais rigoroso e um papel ativo no combate sistêmico ao discurso de ódio. Conclui-se que a hermenêutica do Marco Civil necessitava evoluir para acompanhar a realidade digital, fixando novos parâmetros que preservem a internet como um espaço democrático, mas que assegurem um ambiente seguro onde os direitos fundamentais de todos sejam respeitados.

Palavras-chave: liberdade de expressão; discurso de ódio; dignidade da pessoa humana; plataformas digitais; Supremo Tribunal Federal.

ABSTRACT

This research analyzes how the expansion of the internet and hyperconnectivity have profoundly transformed social communication, highlighting the urgent need to debate the actual limits of freedom of expression within social media environments. The primary objective is to re-evaluate the scope of this right when it is instrumentalized to propagate hate speech, thereby directly violating human dignity and the rights of minorities. To understand this issue, the study examines the foundations of the 1988 Brazilian Federal Constitution, which guarantees the freedom to manifest thought and prohibits prior censorship, while simultaneously establishing the inviolability of honor, image, and human dignity. It is essential to differentiate legitimate opinion from hate speech, as legal doctrine maintains that attacks motivated by prejudice do not find constitutional protection. In Brazil, the primary regulatory framework for the network is the Civil Framework of the Internet (Law 12,965/2014). For years, the original mechanism of Article 19 determined that digital platforms could only be held civilly liable for third-party content if they non-complied with a specific judicial order for removal. However, this model faced severe judicial questioning, under the argument that it facilitated the perpetuation of harmful publications and irreparable damages. The central focus of this research is the substantial jurisprudential shift that occurred in June 2025, when the Brazilian Supreme Federal Court (STF) redefined the interpretation of the aforementioned provision. By reviewing the ministers' votes, it is observed that the Supreme Court established a new normative landscape, imposing a more rigorous duty of care on social media providers and a proactive role in the systemic fight against hate speech. In conclusion, the research synthesizes that the interpretation of the legal framework needed to evolve to align with digital reality, establishing new parameters that preserve the internet as a democratic space while ensuring a secure environment where fundamental rights are respected.

Keywords: freedom of expression; hate speech; human dignity; digital platforms; Brazilian Supreme Federal Court.

SUMÁRIO

1	INTRODUÇÃO	10
2	FUNDAMENTOS CONSTITUCIONAIS E A LIBERDADE DE EXPRESSÃO NO CIBERESPAÇO	13
2.1	Evolução da Liberdade de Expressão	13
2.2	A liberdade de expressão no ciberespaço: entre a democratização e o limite do discurso de ódio	15
2.3	O desafio da regulação jurídica	19
3	O MARCO CIVIL DA INTERNET E A RESPONSABILIDADE CIVIL DAS PLATAFORMAS	23
3.1	O legado do Marco Civil da internet e os novos desafios regulatórios no Brasil	23
3.2	Da mecânica originária do artigo 19 e seus desafios	26
3.3	O papel dos provedores e redes sociais	28
4	O DIREITO COMPARADO E OS MODELOS INTERNACIONAIS DE REGULAÇÃO DIGITAL	31
4.1	O paradigma norte-americano: a <i>Section 230</i> e a imunidade ampla	31
4.2	A Ruptura Alemã: o NetzDG e a privatização do controle	32
4.3	A abordagem sistêmica da união europeia: o <i>Digital Services Act (DSA)</i>	33
5	A VIRADA JURISPRUDENCIAL DO STF (2025-2026)	35
5.1	A decisão da Suprema Corte	35
5.2	Fundamentos da decisão e os votos dos ministros	40
5.3	A análise da inconstitucionalidade	42
5.4	O contexto de instabilidade e a chegada dos casos ao Supremo	46
5.5	Estudo de caso I: dissecando o tema 987 (RE 1.037.396)	47
5.6	Estudo de caso II: dissecando o tema 533 (RE 1.057.258)	48
5.7	O dever fiduciário de cuidado e a inaplicabilidade da responsabilidade objetiva	49
5.8	A consolidação final em 2026: prazos e modulação de efeitos	49
6	A TUTELA COLETIVA E A EFETIVAÇÃO DA RESPONSABILIDADE CIVIL DIGITAL	51
6.1	A Insuficiência da Tutela Individual face às <i>Big Techs</i>	51
6.2	O papel do ministério público e a ação civil pública	52

6.3 O dano moral coletivo no ciberespaço	53
6.4 A Legitimidade das Associações, a Defensoria Pública e a Litigância Estratégica	53
6.5 O processo estrutural como resposta à complexidade algorítmica	54
6.6 A resolução extrajudicial e os Termos de Ajustamento de Conduta (TAC) ..	55
6.7 O desafio da efetividade: astreintes e medidas atípicas de coerção	56
7 CONSIDERAÇÕES FINAIS	58
REFERÊNCIAS	61

1 INTRODUÇÃO

A ascensão da internet e o advento das redes sociais transformaram profundamente as dinâmicas sociais, comunicativas e econômicas da sociedade contemporânea. Em um cenário de hiperconectividade, a livre manifestação do pensamento consolidou-se como um dos pilares da democracia, sendo um direito fundamental historicamente alicerçado em normativas internacionais como o art. 19 da Declaração Universal dos Direitos Humanos, o Art. 13 da Convenção Americana sobre Direitos Humanos e os Arts. 19 e 20 do Pacto Internacional sobre Direitos Civis e Políticos e plenamente consagrado pela Constituição Federal de 1988. Contudo, essa mesma facilidade e velocidade na disseminação de informações propiciaram o surgimento e a proliferação de condutas abusivas, com destaque para a propagação do discurso de ódio (*hate speech*). Esse fenômeno atenta diretamente contra a dignidade da pessoa humana e os direitos fundamentais de grupos minoritários, exigindo uma releitura sobre os limites constitucionais da liberdade de expressão.

Nesse contexto, a legislação brasileira passou a ser alvo de debates acerca de sua eficácia no combate aos abusos cometidos no ambiente virtual. O Marco Civil da Internet (Lei nº 12.965/2014), em seu art. 19, estabeleceu como regra geral que a responsabilização civil dos provedores de aplicação por danos decorrentes de conteúdo gerado por terceiros estaria condicionada à existência de ordem judicial prévia e específica. Tal dispositivo gerou intensa discussão doutrinária e jurisprudencial, visto que, para muitos, a exigência de notificação judicial prévia acabava por criar um “escudo” para as plataformas, dificultando a remoção célere de conteúdos nocivos e a proteção das vítimas. O cenário jurídico sofreu uma alteração substancial em junho de 2025, quando o Supremo Tribunal Federal (STF) julgou a constitucionalidade do referido artigo. O entendimento firmado pela Suprema Corte mudou o paradigma da moderação de conteúdo, redefinindo o papel das redes sociais e abrindo caminho para que sejam responsabilizadas, mesmo sem ordem judicial específica, no que tange ao combate a discursos que violem direitos fundamentais. Essa virada jurisprudencial reacendeu o debate entre a preservação do espaço democrático da rede e a necessidade de controle de abusos.

Diante dessa nova realidade, a presente pesquisa tem como objetivo geral reavaliar os limites do exercício da liberdade de expressão em um cenário de

hiperconectividade, analisando a constitucionalidade e a mutação hermenêutica do art. 19 do Marco Civil da Internet à luz dos novos paradigmas vinculantes fixados pelo Supremo Tribunal Federal, a fim de compreender se a referida diretriz jurisprudencial é adequada para harmonizar a livre manifestação do pensamento com a preservação irrestrita da dignidade da pessoa humana frente à disseminação algorítmica do discurso de ódio.

Para alcançar o objetivo proposto, o trabalho foi estruturado a partir de objetivos específicos sequenciais. Inicialmente, busca-se delinear os fundamentos constitucionais da liberdade de expressão em contraponto intransponível à proteção da dignidade humana e dos direitos da personalidade no combate ao discurso de ódio no ciberespaço. Em seguida, avalia-se a estrutura do modelo originário de imunidade condicionada (*safe harbor*) previsto no ordenamento pátrio, cotejando os desafios regulatórios brasileiros com matrizes internacionais, notadamente a *Section 230* dos Estados Unidos, o *NetzDG* da Alemanha e o *Digital Services Act (DAS)* da União Europeia. Sequencialmente, a pesquisa dedica-se a dissecar tecnicamente os fundamentos jurídicos da virada jurisprudencial promovida pelo STF nos julgamentos dos Temas 987 e 533 da Repercussão Geral, identificando o exato momento de ruptura, a instauração do dever fiduciário de cuidado sistêmico, o duplo regime de responsabilidade e as modulações temporais aplicadas ao ecossistema digital. Por fim, analisam-se os mecanismos de tutela jurisdicional coletiva, destacando a atuação do Ministério Público, a viabilidade do processo estrutural e a aplicação dos Termos de Ajustamento de Conduta (TAC) enquanto instrumentos hábeis a efetivar, na prática judiciária, a responsabilidade civil digital.

Para a consecução destes objetivos, a pesquisa estruturou-se sob a égide do método de abordagem dedutivo. O raciocínio encetado partiu de premissas teóricas gerais e basilares referentes à liberdade de expressão, aos direitos da personalidade e à própria arquitetura econômica e algorítmica das redes sociais, a fim de analisar, em um segundo momento e de forma verticalizada, as consequências jurídicas oriundas da aplicação da redação originária do art. 19 do Marco Civil da Internet e a posterior transição para um novo dever de cuidado no Direito pátrio. Como método de procedimento, optou-se pela pesquisa monográfica, permitindo a investigação aprofundada e delimitada do tema. No que concerne à técnica, o trabalho classifica-se estritamente como pesquisa bibliográfica e documental de natureza qualitativa, amparando-se no levantamento de doutrinas

especializadas em Direito Constitucional e Direito Digital, artigos científicos contemporâneos e legislação estrangeira comparada. Adicionalmente, empregou-se uma exaustiva análise da jurisprudência do Supremo Tribunal Federal, com ênfase técnica no estudo de caso e acompanhamento dogmático dos julgamentos de mérito e dos posteriores embargos de declaração referentes aos Temas 987 e 533 da Repercussão Geral, cujo escrutínio e consolidação final ocorreram em 17 de junho de 2026.

2 FUNDAMENTOS CONSTITUCIONAIS E A LIBERDADE DE EXPRESSÃO NO CIBERESPAÇO

2.1 Evolução da Liberdade de Expressão

A liberdade de expressão é um direito fundamental reconhecido em diversos documentos e constituições ao redor do mundo. A Organização das Nações Unidas, por meio da Declaração Universal dos Direitos Humanos, estabelece no artigo 19 que toda pessoa tem o direito de opinar e se expressar livremente. Isso inclui não ser perseguida por suas opiniões e poder buscar, receber e compartilhar informações e ideias, independentemente de fronteiras ou dos meios utilizados.

Esse direito abrange tanto a liberdade de formar opiniões quanto a possibilidade de divulgá-las sem interferência indevida.

Nos Estados Unidos, a Primeira Emenda à Constituição dos Estados Unidos assegura a liberdade de expressão. Já no Brasil, a Constituição Federal de 1988 garante esse direito no Artigo 5º, inciso IX, ao permitir a livre manifestação do pensamento sem censura.

Além disso, o inciso IV do mesmo artigo determina que a manifestação do pensamento é livre, mas proíbe o anonimato. O artigo 220 também reforça que a expressão, a criação e a informação não devem sofrer restrições. Contudo, a Constituição estabelece que, em caso de abuso, pode haver responsabilização, garantindo o equilíbrio entre a liberdade de expressão e a proteção de outros direitos fundamentais.

Para a sociedade brasileira, essa garantia representa uma mudança importante na história recente. Durante o período da Ditadura Militar no Brasil, a liberdade de expressão foi fortemente limitada. Nesse contexto, havia censura prévia rigorosa: jornais, revistas, livros, programas de televisão e filmes precisavam ser analisados e aprovados por órgãos do governo antes de serem divulgados ao público.

Durante esse período, qualquer conteúdo considerado contrário ao governo ou visto como ameaça à ordem era proibido. O Ato Institucional Número 5 marcou o auge da repressão, ao suspender garantias constitucionais, fechar o Congresso Nacional, formalizar a censura prévia e concentrar amplos poderes nas mãos do presidente. Com isso, a repressão se intensificou ainda mais, tornando a

censura extremamente rigorosa. Nesse contexto, a retomada da liberdade de expressão tornou-se um elemento essencial no processo de redemocratização do país.

O advento massificado da rede mundial de computadores reconfigurou substancialmente a morfologia e o alcance da liberdade de expressão, gerando externalidades disruptivas que atuam de forma paradoxal, ora como catalisadores democráticos, ora como vetores deletérios às relações informacionais. Antes, a circulação de ideias dependia de meios tradicionais, como jornais, rádio e televisão, que eram controlados por poucos grupos. Com a internet, qualquer pessoa passou a produzir e divulgar conteúdo em escala global, quase sem custo e de forma instantânea. Isso ampliou drasticamente o alcance da liberdade de expressão, democratizando o debate público e dando voz a grupos que antes eram invisibilizados.

Plataformas como Facebook, Twitter (X) e YouTube consolidaram-se como espaços centrais para a manifestação de ideias, opiniões políticas e mobilização social. Todavia, cumpre salientar que a expansão da liberdade de manifestação no ambiente virtual ensejou, concomitantemente, o surgimento de novos e complexos desafios jurídicos. Sob essa perspectiva, o mesmo arranjo tecnológico que descentraliza e amplia vozes também atua como um potente vetor para a disseminação célere de desinformação, discursos de ódio e ataques coordenados.

Além disso, surgiu um novo tipo de “controle”, não necessariamente estatal, mas privado. Grandes empresas de tecnologia, como Meta e Google, passaram a definir regras sobre o que pode ou não ser publicado em suas plataformas. Ou seja, mesmo sem censura estatal direta, existe moderação de conteúdo, o que levanta debates sobre limites, transparência e poder dessas empresas.

Outro ponto importante é que governos também passaram a atuar mais no ambiente digital. Alguns países adotam leis para combater fake News e crimes online, enquanto outros usam essas mesmas justificativas para restringir críticas e controlar a informação mostrando que o velho problema da censura não desapareceu, só mudou de forma.

O sistema jurídico brasileiro, estruturado a partir da Constituição Federal de 1988 e complementado por normas internacionais como a Convenção Americana

sobre Direitos Humanos, busca conciliar direitos fundamentais que possuem o mesmo nível de importância.

Ainda assim, essa convivência nem sempre é pacífica. O conflito surge justamente quando o exercício da liberdade de expressão ultrapassa os limites e atinge a esfera de proteção da dignidade de outra pessoa.

A Constituição reconhece a liberdade de expressão como elemento essencial do Estado Democrático. Os incisos IV e IX do art. 5º garantem a livre manifestação do pensamento e a liberdade de produção intelectual, artística e científica, sem censura prévia.

Além disso, o art. 220 reforça essa proteção ao vedar qualquer forma de restrição à manifestação do pensamento, à criação e à informação.

No âmbito internacional, o art. 13 da Convenção Americana amplia esse entendimento ao assegurar não apenas o direito de se expressar, mas também de buscar, receber e divulgar informações por qualquer meio, inclusive além das fronteiras nacionais.

Apesar de sua relevância, a liberdade de expressão não possui caráter absoluto. O próprio texto constitucional, no art. 5º, inciso X, estabelece a proteção à intimidade, à vida privada, à honra e à imagem das pessoas.

O problema aparece quando esses dois conjuntos de direitos entram em choque. Nesses casos, não há hierarquia pré-definida entre eles. A solução depende da análise do caso concreto, por meio da aplicação do princípio da proporcionalidade. Para isso, alguns critérios costumam ser considerados: a relevância pública da informação; a veracidade ou plausibilidade dos fatos divulgados; a intenção de informar, criticar ou ofender.

2.2 A liberdade de expressão no ciberespaço: entre a democratização e o limite do discurso de ódio

A ascensão da internet e a consolidação do ciberespaço reconfiguraram drasticamente a esfera pública, transformando a maneira como a sociedade interage, consome informação e exerce seus direitos fundamentais. Historicamente limitada aos meios de comunicação de massa tradicionais, onde a emissão de informações era vertical e centralizada, a comunicação humana encontrou no ambiente digital um modelo horizontal e descentralizado. Nesse cenário, a liberdade

de expressão encontrou um ambiente fértil para sua máxima expansão, permitindo que qualquer indivíduo com acesso à rede se tornasse não apenas um receptor, mas um emissor global de opiniões. Essa reconfiguração representou um avanço inegável para a democratização da informação e para o empoderamento cidadão. No entanto, o ambiente virtual não é uma zona alheia ao ordenamento jurídico. O grande desafio contemporâneo do direito constitucional e digital é estabelecer a complexa ponderação entre a garantia da livre manifestação do pensamento e a proteção de outros direitos da personalidade, como a honra, a intimidade e a dignidade da pessoa humana.

No ordenamento jurídico brasileiro, a liberdade de expressão é consagrada no artigo 5º, incisos IV e IX, da Constituição Federal de 1988, assumindo a posição de um dos pilares indispensáveis para a manutenção e o fortalecimento do Estado Democrático de Direito. Trata-se de uma garantia que assegura a livre manifestação do pensamento, a expressão da atividade intelectual, artística e científica, bem como a comunicação, independentemente de censura prévia ou licença governamental. Contudo, a arquitetura própria da internet e a sensação de distanciamento físico criaram, em parte considerável dos usuários, a falsa e perigosa percepção de que a liberdade de expressão no ciberespaço seria um direito de caráter absoluto. Essa ilusão de impunidade é frequentemente potencializada pelo uso de perfis falsos e do anonimato, ignorando a expressa vedação constitucional a essa prática. É imperativo compreender, sob a ótica da hermenêutica constitucional, que nenhum direito fundamental é absoluto. Eles encontram fronteiras intransponíveis no momento em que colidem com outros princípios de igual ou maior peso axiológico na ordem constitucional.

A linha tênue que separa o exercício regular e legítimo da liberdade de expressão da prática de ilícitos é diuturnamente testada e esgarçada no ciberespaço. A mesma rede que democratiza o acesso ao conhecimento atua como um potente vetor para a disseminação do chamado discurso de ódio (*hate speech*). Essas manifestações, longe de se configurarem como mero exercício crítico ou opiniões impopulares traduzem-se em mensagens que incitam a discriminação, a hostilidade, a segregação ou a violência contra indivíduos ou grupos, tendo como base marcadores identitários como raça, religião, gênero, orientação sexual ou origem nacional. O paradoxo da tolerância, formulado pelo filósofo Karl Popper, ilustra bem essa dinâmica no debate jurídico: a tolerância ilimitada fatalmente leva

ao desaparecimento da própria tolerância, uma vez que, ao se estender de forma irrestrita aos intolerantes, a sociedade corre o risco de ter suas próprias bases democráticas destruídas por eles.

Nesse contexto, o Supremo Tribunal Federal (STF) já consolidou o entendimento, em diversos precedentes históricos, de que a incitação ao ódio público não encontra guarida na cláusula constitucional da liberdade de expressão. Quando a fala tem o objetivo deliberado de inferiorizar, subjugar ou aniquilar a dignidade de terceiros, ela transborda as margens do debate salutar e adentra a esfera da ilicitude material. Tais condutas podem perfeitamente configurar crimes contra a honra, racismo, injúria racial, entre outros tipos penais, demonstrando que a Constituição não protege o discurso que visa corroer a igualdade inerente a todos os seres humanos. O sopesamento entre princípios exige que a dignidade da pessoa humana atue como um vetor de limitação quando a fala passa a ser uma arma de anulação do outro.

Para regular essas complexas interações no meio digital, a promulgação da Lei nº 12.965/2014, conhecida como Marco Civil da Internet, representou um avanço regulatório fundamental. A legislação estabeleceu a liberdade de expressão como um de seus fundamentos essenciais, mas também precisou enfrentar o dilema da responsabilidade das grandes plataformas de tecnologia e redes sociais. Para evitar o fenômeno da censura privada no qual as empresas poderiam excluir conteúdos indiscriminadamente por mero receio de litígios e proteger a livre circulação de ideias, o legislador optou, em seu artigo 19, pelo modelo da responsabilidade civil condicionada. Em regra, os provedores de aplicação não respondem de forma objetiva pelo conteúdo gerado por terceiros, sendo responsabilizados apenas se, após ordem judicial específica, não tomarem as providências para tornar o material ilícito indisponível. Essa escolha legislativa buscou um ponto de equilíbrio: resguarda a plataforma contra a obrigatoriedade de atuar como "censora prévia" da rede, mas garante à vítima um mecanismo jurídico eficaz para a remoção de conteúdos manifestamente lesivos quando o Poder Judiciário é acionado.

Em suma, o ciberespaço alterou de forma indelével a comunicação humana, democratizando a voz de bilhões de pessoas, mas, simultaneamente, amplificando o potencial de dano em escalas antes inimagináveis. A defesa intransigente da liberdade de expressão na era digital não pode, sob nenhuma

hipótese, significar a tolerância passiva com a intolerância e o aviltamento. A convivência harmônica nas redes exige uma filtragem constitucional constante e atenta, na qual o Estado-juiz, as plataformas de tecnologia e a própria sociedade civil devem atuar em conjunto. Somente por meio dessa corresponsabilidade será possível garantir que a internet permaneça um espaço de debate plural, crítico e aberto, mas implacavelmente resistente aos discursos que visam subverter o Estado de Direito e a dignidade fundamental.

A Delimitação do Discurso de Ódio (*Hate Speech*): Historicamente, o fenômeno do discurso de ódio não representa uma novidade na estrutura das relações sociais. Pelo contrário, ele tem se manifestado de forma cíclica sob diferentes roupagens. O exemplo mais emblemático e catastrófico da modernidade foi, sem dúvida, o regime nacional-socialista na Alemanha. Durante o nazismo, o ódio não era apenas uma manifestação individual ou periférica, mas sim uma política de Estado institucionalizada.

Através da máquina de propaganda liderada por figuras como Joseph Goebbels, o Estado nazista utilizou a retórica da exclusão para desumanizar o "outro" especificamente a população judaica, entre outros grupos minoritários. Ao rotular o diferente como uma ameaça à integridade nacional, o discurso de ódio serviu como o alicerce jurídico e moral que legitimou o extermínio e a perseguição sistemática. Nesse contexto, a palavra precedeu a ação; o discurso pavimentou o caminho para o campo de concentração.

A Ruptura do Século XXI: A Era da Hiperconectividade: Apesar desse lastro histórico, o debate contemporâneo sobre o discurso de ódio adquiriu matizes inéditos no século XXI. A transição da esfera pública analógica para o ecossistema digital alterou a natureza da comunicação humana. Hoje, o Direito enfrenta um desafio hercúleo: como regular manifestações que ignoram fronteiras geográficas e temporais?

O diferencial do problema hodierno reside na arquitetura das redes sociais. Se no passado o alcance de um discurso odioso dependia de palanques físicos ou da mídia de massa controlada, hoje qualquer indivíduo munido de um smartphone detém um poder de difusão massiva. A internet não apenas hospeda o ódio, ela o catalisa através de algoritmos que priorizam o engajamento e, lamentavelmente, o conteúdo que gera revolta e hostilidade tende a circular com maior velocidade do que o discurso racional e moderado.

A Mecânica da viralização e a Amplificação do Dano: O ponto central que exige uma revisão urgente das normas jurídicas é a velocidade cinestésica da disseminação digital. No ambiente virtual, a ofensa não é mais um evento isolado; ela se torna um fenômeno viral em questão de segundos, destacando-se: a) o alcance exponencial, pois uma postagem ofensiva atinge uma audiência global antes mesmo que a vítima possa tomar conhecimento de sua existência; b) a irreversibilidade do estigma, tornando o chamado “direito ao esquecimento” quase utópico em um cenário onde dados são replicados em servidores mundiais; e c) o constrangimento excessivo, gerando um impacto psicológico e social devastador para a vítima, uma vez que a exposição não se limita a um círculo imediato.

A Semiótica do Ódio Moderno: Memes, *Stickers* e Desinformação: A tecnologia não mudou apenas o "como" o ódio é distribuído, mas também o "quê" constitui essa agressão. Atualmente, presenciamos uma ressignificação estética do preconceito, onde o ódio se disfarça de humor através de: a) imagens e stickers, pequenas unidades visuais que ridicularizam grupos vulneráveis de forma rápida e lúdica; b) manipulação de vídeos e áudios, com o uso de trechos descontextualizados ou inteligência artificial para criar conteúdos falsos; e c) cultura do cancelamento e do linchamento virtual, na qual o discurso de ódio é camuflado sob uma pretensa "busca por justiça", resultando em perseguições coletivas.

2.3 O desafio da regulação jurídica

Diante desse cenário, torna-se imperativo que o ordenamento jurídico não apenas reaja aos danos causados, mas que se antecipe a eles. A discussão sobre a regulação das plataformas digitais não é um ataque à liberdade de expressão, mas sim uma medida de proteção à dignidade da pessoa humana.

É preciso encontrar um equilíbrio que responsabilize não apenas os autores diretos, mas também as plataformas que lucram com a circulação desses conteúdos. O ódio digital, em sua capacidade de fragmentar democracias e destruir vidas em tempo real, exige uma resposta que seja tão veloz e sofisticada quanto as tecnologias que o propagam. A proteção dos direitos fundamentais no século XXI depende, essencialmente, da nossa capacidade de domesticar a fúria das redes e devolver ao debate público a alteridade e o respeito mútuo.

O fenômeno da hostilidade direcionada a grupos minoritários ou indivíduos específicos, embora possua raízes profundas na história da civilização, assume na contemporaneidade uma roupagem técnica e estrutural que desafia as fronteiras do Direito e da Ética. Quando retrocedemos ao século XX, observamos que o discurso de ódio era frequentemente uma ferramenta de controle vertical, emanada de estruturas de poder centralizadas, como ocorreu de forma devastadora durante o regime nazista. Naquele período, a perseguição aos judeus e a outros grupos considerados dissidentes não era apenas uma manifestação de ódio interpessoal, mas um projeto de engenharia social legitimado por normas jurídicas e por uma máquina de propaganda estatal que desumanizava o "outro" para viabilizar a sua eliminação física. O Estado era o principal emissor da ofensa e o garantidor da sua eficácia destrutiva, criando um ambiente onde o preconceito era a lei e a violência era o método.

Contudo, ao cruzarmos o limiar do século XXI, a natureza dessa problemática sofreu uma mutação radical provocada pela revolução digital e pela democratização, ainda que caótica, dos meios de emissão de informação. O que antes dependia de grandes palanques ou da anuência de editores de jornais e emissoras de rádio, agora encontra terreno fértil na palma da mão de qualquer cidadão conectado. Essa transição do discurso de ódio da esfera institucional para a esfera capilarizada da internet transformou o que era uma questão de abuso de poder estatal em um problema de governança algorítmica e responsabilidade civil coletiva. A grande diferença do cenário atual reside não necessariamente na essência do ódio, que permanece pautada pela intolerância, mas na sua capacidade de disseminação, que atingiu níveis de velocidade cinestésica e alcance global até então inimagináveis.

A arquitetura das redes sociais, desenhada para maximizar o engajamento através da economia da atenção, acaba por atuar como um acelerador de partículas para o ódio. Algoritmos de recomendação, programados para manter o usuário conectado pelo maior tempo possível, identificam que conteúdos carregados de carga emocional negativa, como a indignação e a hostilidade, geram mais interações do que discursos moderados ou pautados pela racionalidade. Isso cria um ciclo vicioso onde o discurso de ódio não é apenas permitido, mas involuntariamente impulsionado pelas plataformas, alcançando uma audiência de milhões de pessoas em um intervalo de poucos minutos. Para a vítima, esse

processo resulta em um constrangimento excessivo e multidimensional, pois a ofensa se desprende do tempo e do espaço, permanecendo cristalizada na memória digital da rede e sendo replicada *ad infinitum*.

Nesse contexto, a gramática da agressão também se modernizou, abandonando frequentemente os textos longos em favor de unidades semióticas mais ágeis e de difícil moderação, como fotos manipuladas, vídeos curtos descontextualizados e os onipresentes *stickers* de aplicativos de mensagens. Essas ferramentas tornam a ridicularização de grupos vulneráveis algo aparentemente lúdico ou inofensivo, camuflando o preconceito sob o manto de um suposto humor. A facilidade com que se pode editar a imagem de um indivíduo para torná-lo alvo de escárnio público ou utilizar fragmentos de áudio para forjar confissões e opiniões absurdas criou um ambiente de insegurança jurídica onde a verdade é constantemente sacrificada no altar da viralização. O impacto na vida dos atingidos é devastador, gerando desde o isolamento social e prejuízos profissionais até danos psíquicos profundos, uma vez que a exposição digital funciona como um linchamento público que nunca cessa, já que o conteúdo, uma vez lançado na rede, torna-se quase impossível de ser totalmente erradicado.

Essa nova realidade exige que o Direito abandone posturas meramente reativas e passe a revisitar seus conceitos fundamentais sobre a liberdade de expressão. O debate clássico sobre até onde vai o direito de dizer o que se pensa ganha novas camadas de complexidade quando a fala não é apenas uma ideia, mas um vetor de destruição de reputações impulsionado por máquinas. A necessidade de regular as plataformas digitais torna-se, portanto, uma questão de proteção da dignidade humana, pois não se pode permitir que a inovação tecnológica se sobreponha aos direitos fundamentais básicos. É preciso entender que a velocidade da internet não pode servir de escudo para a impunidade, e que a responsabilidade pela manutenção de um ambiente público saudável deve ser compartilhada entre quem emite a ofensa, quem a replica e as corporações que provêm a infraestrutura técnica para que tal discurso circule. Somente através de uma nova hermenêutica jurídica, que compreenda a volatilidade e a perenidade do meio digital, será possível construir barreiras eficazes contra essa nova forma de barbárie que, sob o pretexto da conectividade, ameaça desintegrar os laços de alteridade e respeito que sustentam a convivência democrática.

Na perspectiva do direito constitucional contemporâneo, a liberdade de expressão não confere guarida ao discurso de ódio (*hate speech*). Conforme leciona Daniel Sarmento (2006), tais discursos consistem em manifestações de desprezo, ódio e intolerância direcionadas a grupos específicos, essencialmente motivadas por preconceitos ligados a marcadores como etnia, religião, gênero, orientação sexual ou deficiência. O autor defende que essa modalidade de manifestação não possui o condão de enriquecer o debate público livre; pelo contrário, seu objetivo central é atacar a dignidade da pessoa humana e corroer o princípio constitucional da igualdade. Sendo assim, a jurisprudência e a doutrina convergem no sentido de que o discurso de ódio mostra-se intrinsecamente incompatível com as aspirações de uma sociedade democrática, justificando-se a restrição à liberdade de expressão quando a fala é deliberadamente utilizada como mecanismo de opressão e exclusão de minorias estigmatizadas.

3 O MARCO CIVIL DA INTERNET E A RESPONSABILIDADE CIVIL DAS PLATAFORMAS

3.1 O legado do Marco Civil da internet e os novos desafios regulatórios no Brasil

Em abril de 2014, o ordenamento jurídico brasileiro inaugurava um de seus diplomas mais vanguardistas: a Lei nº 12.965, popularmente batizada como o Marco Civil da Internet. Sancionada logo após sua aprovação no Senado Federal, a normativa nasceu com o audacioso propósito de consolidar os direitos fundamentais, as garantias civis e os deveres dos usuários e provedores no ecossistema digital. À época de sua gênese, a legislação colheu severos elogios e reconhecimento global, posicionando o Brasil como uma referência internacional pioneira ao positivá-los sob os pilares da neutralidade de rede, da salvaguarda da privacidade e do absoluto respeito à liberdade de expressão e à pluralidade de ideias.

Contudo, o dinamismo tecnológico das últimas décadas impôs uma veloz metamorfose nas relações sociais e jurídicas travadas na rede. Como desdobramento natural dessa evolução e da necessidade de um microssistema de proteção mais robusto, sobreveio a Lei Geral de Proteção de Dados Pessoais (LGPD Lei nº 13.709/2018), que passou a disciplinar especificamente o tratamento de dados por entes públicos e privados, complementando a base principiológica outrora fixada em 2014.

Passada mais de uma década da promulgação do Marco Civil, o cenário contemporâneo descortina desafios complexos que transcendem a mera universalização do acesso à rede mundial de computadores. O debate atual deslocou-se para a urgência em se garantir um ambiente virtual seguro, ético e confiável. Entre as pautas mais prementes e debatidas no cenário jurídico e político nacional, destacam-se: a) o combate sistêmico à desinformação, uma vez que a proliferação deliberada de notícias falsas (fake news) tem tensionado os limites do debate público; b) a regulação da Inteligência Artificial (IA), que traz a necessidade de balizamentos éticos e de responsabilidade civil para o desenvolvimento e aplicação de novos algoritmos; e c) a transparência das Big Techs, com a exigência de maior clareza nos critérios de moderação de conteúdo e funcionamento dos algoritmos das grandes plataformas de redes sociais.

A ausência de respostas institucionais a esses fenômenos não atinge apenas a esfera individual, mas projeta riscos sistêmicos que ameaçam diretamente a integridade das instituições republicanas e a própria higidez do regime democrático.

O Projeto de Lei nº 2.630/2020, conhecido como “PL das Fake News”, está em tramitação no Congresso Nacional e tem como principal objetivo combater a disseminação de notícias falsas nas redes sociais e estabelecer regras para a atuação das chamadas *big techs*, empresas responsáveis pelas grandes plataformas digitais. A proposta foi apresentada pelo senador Alessandro Vieira, do MDB de Sergipe, e aprovada no Senado Federal em 2020 com relatório do senador Angelo Coronel, do PSD da Bahia.

O texto prevê medidas para aumentar a transparência das plataformas digitais e responsabilizar as empresas pela circulação de conteúdos impulsionados por seus algoritmos, especialmente quando contribuírem para a propagação de desinformação, discursos de ódio e conteúdos ilegais. A proposta também busca criar mecanismos de fiscalização e exigir maior controle das redes sociais sobre perfis falsos e campanhas de manipulação online.

Apesar da aprovação no Senado, o projeto encontrou forte resistência na Câmara dos Deputados, principalmente por parte de parlamentares e representantes das empresas de tecnologia, que alegam possíveis riscos à liberdade de expressão e ao funcionamento das plataformas digitais. Diante da falta de consenso para votação, o então presidente da Câmara, Arthur Lira, anunciou a criação de um grupo de trabalho para discutir o tema e buscar alterações no texto antes de uma possível votação. O debate em torno do projeto continua gerando grande repercussão no meio político e na sociedade, especialmente por envolver questões relacionadas à democracia, à liberdade de expressão e à responsabilidade das empresas de tecnologia no ambiente digital.

O Marco Civil da Internet, sancionado pela então presidente Dilma Rousseff, foi um marco pioneiro na legislação brasileira ao estabelecer direitos, deveres e garantias para o uso da internet no país. Entre seus principais pontos, a lei assegurou a inviolabilidade e o sigilo das comunicações realizadas no ambiente digital, garantindo maior proteção à privacidade dos usuários. No entanto, a norma também prevê que essas informações podem ser acessadas mediante ordem judicial, especialmente em investigações criminais ou processos judiciais.

Segundo o Marco Civil, os provedores e administradores de internet devem manter sob sigilo os registros de conexão dos usuários, como data e horário de acesso, duração da conexão e endereço de IP, pelo prazo mínimo de um ano. A legislação ainda permite que esse período seja ampliado mediante solicitação do Ministério Público ou de autoridade policial, mesmo sem autorização judicial prévia. Apesar disso, o acesso efetivo aos dados armazenados continua dependendo de decisão da Justiça.

No Congresso Nacional, o tema segue em debate por meio do Projeto de Lei nº 113/2020, de autoria do senador Angelo Coronel. A proposta amplia os poderes investigativos e autoriza membros do Ministério Público e delegados de polícia a requisitarem registros de conexão sem necessidade de autorização judicial antecipada. O texto original também previa a obrigatoriedade de apresentação de CPF ou CNPJ para criação de perfis nas redes sociais e plataformas digitais, mas essa exigência foi modificada pelo relator, o senador Marcos Pontes. Atualmente, o projeto aguarda votação na Comissão de Comunicação e Direito Digital do Senado.

Outro ponto importante do Marco Civil da Internet está previsto no art. 19, que determina que os provedores de internet somente podem ser responsabilizados por conteúdos publicados por terceiros caso descumpram uma ordem judicial de remoção. Esse dispositivo foi criado com a finalidade de proteger a liberdade de expressão e evitar práticas de censura nas plataformas digitais. Entretanto, após intenso debate sobre os abusos cometidos sob o escudo dessa norma, o Supremo Tribunal Federal consolidou o reconhecimento da inconstitucionalidade parcial e progressiva do art. 19 do Marco Civil da Internet, redefinindo, assim, os limites e os novos parâmetros da responsabilidade civil das plataformas digitais no país.

Além disso, tramita no Senado o Projeto de Lei nº 592/2023, apresentado pelo senador Jorge Seif, que estabelece novas regras para a moderação de conteúdos nas redes sociais e dificulta a remoção de publicações e a suspensão de contas pelas plataformas digitais. Segundo o autor, a proposta retoma parte das medidas previstas na Medida Provisória nº 1.068/2021, editada durante o governo do então presidente Jair Bolsonaro, a qual foi devolvida ao Poder Executivo pelo presidente do Senado, Rodrigo Pacheco, sob o argumento de inconstitucionalidade. O projeto recebeu parecer favorável do senador Hamilton Mourão na Comissão de Comunicação e Direito Digital e, dando prosseguimento ao seu rito legislativo, avançou para a Comissão de Direitos Humanos e Legislação Participativa, onde se

encontra com relatoria designada desde outubro de 2025, aguardando os trâmites subsequentes para deliberação e votação ao longo de 2026. Essa movimentação legislativa ganha contornos de extrema urgência e relevância diante do próprio item 13 da tese vinculante fixada pelo Supremo Tribunal Federal nos Temas 987 e 533. Em sua decisão, a Suprema Corte emitiu um apelo expresso e formal ao Poder Legislativo para que o Congresso Nacional não se omita e elabore, com brevidade, uma legislação regulatória robusta, capaz de sanar de forma definitiva as graves deficiências do atual regime jurídico e garantir a proteção perene dos direitos fundamentais da personalidade no ecossistema digital.”

3.2 Da mecânica originária do artigo 19 e seus desafios

Para a compreensão do regime de responsabilidade civil no ambiente virtual brasileiro, a análise da norma basilar contida no Artigo 19 do Marco Civil da Internet (Lei nº 12.965/2014) é imprescindível. O dispositivo consagrou, como regra geral, que os provedores de aplicação de internet não respondem objetivamente pelos conteúdos gerados por terceiros. A responsabilização civil por danos decorrentes de postagens ilícitas foi condicionada ao descumprimento de uma ordem judicial específica e prévia que determinasse a indisponibilização do material.

A opção legislativa pela responsabilidade civil subjetiva condicionada frequentemente denominada pela doutrina como modelo de *safe harbor* (porto seguro) teve como principal fundamento a proteção da liberdade de expressão e a vedação à censura prévia privada. Caso as plataformas fossem responsabilizadas sumariamente por qualquer denúncia, o receio do litígio e de prejuízos financeiros forçaria essas corporações a adotar uma postura excessivamente defensiva, excluindo de imediato qualquer conteúdo controverso. Tal dinâmica atribuiria às empresas privadas o papel de censores do debate público, comprometendo a pluralidade de ideias na rede. Portanto, a exigência de atuação do Poder Judiciário visou garantir que a verificação da ilicitude fosse realizada pela autoridade estatal competente, assegurando o contraditório e o devido processo legal antes da restrição de qualquer conteúdo. Conforme pondera Castro (2013), esse modelo buscou harmonizar a eficácia dos direitos da personalidade com a preservação da liberdade de expressão no ambiente virtual, evitando que o receio de litígios

massivos induzisse os intermediários a exercerem um monitoramento privado e excessivo sobre o fluxo informativo.

Apesar de sua abrangência estrutural, o diploma legal previu exceções pontuais a essa regra. A mais notória encontra-se no Artigo 21 do Marco Civil, que versa sobre a divulgação não autorizada de imagens ou vídeos contendo cenas de nudez ou atos sexuais de caráter privado. Nesses casos de grave violação à intimidade (frequentemente associados à “pornografia de vingança”), o legislador dispensou a ordem judicial, estabelecendo que a plataforma se torna civilmente responsável caso não remova o conteúdo de forma diligente após o recebimento de mera notificação extrajudicial direta da vítima. Violações a direitos autorais também possuem ressalvas legais próprias, não se submetendo à exigência estrita do Artigo 19.

Contudo, a evolução estrutural das redes sociais impôs desafios severos à manutenção desse modelo originário. A transição para um ecossistema digital pautado pela “economia da atenção”, no qual algoritmos de recomendação são programados para maximizar o engajamento, reconfigurou substancialmente a atuação dos provedores. As plataformas deixaram de ser meras hospedeiras passivas de dados para atuar como curadoras ativas e impulsionadoras de informação. Nesse contexto, evidenciou-se que a disseminação de discursos de ódio, teorias conspiratórias e conteúdos ilícitos gera alta lucratividade, uma vez que tais materiais costumam reter a atenção do usuário por meio de reações emocionais intensas.

Diante dessa nova realidade tecnológica, a aplicação irrestrita e literal da regra do Artigo 19 revelou-se insuficiente para a salvaguarda de outros direitos fundamentais. A exigência de ordem judicial prévia para a remoção de conteúdos manifestamente lesivos e deliberadamente impulsionados por algoritmos permitiu que danos irreversíveis à honra, à imagem e à integridade democrática se perpetuassem em escala global enquanto a vítima aguardava a tutela estatal. A constatação de que o modelo reativo deixava a coletividade desprotegida frente aos abusos massivos e à monetização da desinformação fomentou um intenso escrutínio jurídico sobre a defasagem da lei, culminando na reinterpretação de sua aplicabilidade e nos rigorosos questionamentos acerca de sua inconstitucionalidade parcial, temática que desaguaria na virada jurisprudencial do Supremo Tribunal Federal (STF) a partir de 2025.

3.3 O papel dos provedores e redes sociais

A compreensão contemporânea do direito constitucional evidencia que a liberdade de expressão não ostenta um caráter absoluto, encontrando barreiras intransponíveis sempre que colidir com outros direitos de igual estatura fundamental, cenário que se manifesta de forma nítida na problemática do discurso de ódio. Quando tais manifestações ilícitas e discriminatórias são propagadas no ambiente das redes sociais, a complexidade do fenômeno se acentua significativamente, exigindo que os mecanismos de enfrentamento e contenção sejam reformulados para se ajustarem às dinâmicas próprias da tecnologia digital.

A antiga premissa de que a internet se consolidaria como um território anárquico e imune à incidência jurídica demonstrou-se inteiramente equivocada frente à realidade atual, que é marcada por intensos esforços institucionais voltados ao aprimoramento dos sistemas de fiscalização e controle. Nesse contexto, é preciso reconhecer que o ciberespaço opera sob uma complexa engenharia moldada por códigos de programação e sistemas algorítmicos. Longe de representarem ferramentas tecnicamente neutras, essas estruturas são desenvolvidas por agentes humanos e, por conseguinte, respondem a interesses econômicos, políticos e sociais específicos de seus criadores.

Essa configuração digital impõe severos obstáculos à aplicação dos institutos clássicos do Direito Teórico e Positivo. A relação jurídica e comunicativa estabelecida entre o criador do conteúdo, a empresa que gerencia a plataforma e o destinatário final da informação não se amolda aos parâmetros tradicionais dos meios de comunicação analógicos. Nas redes sociais, inexistente uma linha editorial rígida ou uma seleção prévia do que será publicado, uma vez que o material digital é produzido, compartilhado e difundido pelos próprios usuários de maneira automatizada e praticamente instantânea, inviabilizando uma fiscalização antecedente por parte das empresas de tecnologia.

A fim de explicar a ruptura com os padrões tradicionais de regulação, Jack Balkin (2018) propõe a substituição do antigo modelo binário de proteção à liberdade de expressão por um arranjo pluralista de governança da internet, estruturado de maneira triangular. De acordo com essa formulação teórica, o primeiro vértice do triângulo é ocupado pelo poder político dos Estados-Nações. O segundo vértice

congrega as empresas privadas detentoras da infraestrutura tecnológica e as plataformas de redes sociais, enquanto o terceiro ponto da estrutura abriga a totalidade dos emissores de discursos, englobando desde indivíduos comuns até organizações da sociedade civil e ativistas digitais.

A interação entre os componentes desse triângulo revela uma dinâmica de poder multifacetada. Por um lado, o poder estatal tenta reprimir abusos valendo-se das ferramentas punitivas clássicas do ordenamento, como sanções civis e penais. Contudo, cientes das limitações práticas da jurisdição territorial frente à rede global, os governos buscam coagir ou cooptar as grandes empresas de tecnologia, valendo-se de novas legislações para transformá-las em agentes de fiscalização do debate público. Concomitantemente, as corporações detentoras da infraestrutura tecnológica passam a exercer um poder regulatório direto e privado sobre os cidadãos, ditando de forma autônoma as regras de remoção e moderação de conteúdos por meio de seus próprios termos de serviço e diretrizes comunitárias (Balkin, 2018, p. 2014-2015).

A atuação das mídias digitais no enfrentamento às manifestações ilícitas encontra seu fundamento primário nos termos de serviço e diretrizes comunitárias, instrumentos que delimitam as fronteiras do comportamento considerado aceitável no ambiente virtual. Conforme elucida Kate Klonick (2018), a elaboração dessas políticas de moderação de conteúdo é estruturada a partir de uma tríade de vetores essenciais: a profunda influência da cultura jurídica norte-americana de proteção à livre manifestação do pensamento, o imperativo da responsabilidade social corporativa e a necessidade pragmática de sustentabilidade econômica, que exige a manutenção de um ambiente online seguro e atraente para usuários e anunciantes.

Disso decorre a percepção de que os ecossistemas virtuais passam longe de qualquer neutralidade. Conforme elucida Nitrini (2021), a moderação de conteúdo nas plataformas digitais é operacionalizada, em grande medida, por meio de sistemas de filtragem algorítmica que gerenciam a dinâmica comunicativa a partir do binômio da visibilidade e da invisibilidade (banimento). Essas decisões arquitetônicas, intrínsecas aos códigos de programação, refletem inevitáveis intervenções humanas e propósitos comerciais delineados. Tais mecanismos sofrem pressões contínuas das dinâmicas de mercado, das convenções da sociedade civil, das limitações da própria tecnologia e das variadas regulamentações estatais em que as empresas operam.

No que tange à evolução e à operacionalização desse monitoramento, a escala massiva de postagens exigiu o desenvolvimento de métodos de fiscalização híbridos. Além do controle prévio, as plataformas confiam parcela expressiva da identificação de abusos à própria coletividade, utilizando o sistema de *flagging* (sinalização). Essa técnica consiste na disponibilização de ferramentas para que os próprios usuários denunciem postagens que violem as diretrizes. A sinalização atua, portanto, como um gatilho essencial para encaminhar o conteúdo suspeito ao escrutínio de moderadores humanos, consolidando um modelo de regulação privada e descentralizada (Nitrini, 2021).

A governança privada da internet conta com mecanismos complexos de fiscalização. Um exemplo prático dessa engrenagem operacional pode ser observado na estrutura hierárquica adotada por grandes plataformas, como o Facebook, minuciosamente detalhada por Klonick (2018). Na base dessa pirâmide de moderação, uma vasta força de trabalho frequentemente terceirizada atua na triagem inicial, resolvendo as demandas mais volumosas a partir de diretrizes objetivas. Em um patamar intermediário, moderadores com maior experiência técnica e treinamento específico coordenam os analistas de primeira linha, assumindo a análise de tópicos sensíveis ou urgentes. Por fim, o topo dessa estrutura é ocupado por especialistas seniores e equipes jurídicas internas, encarregados de proferir a decisão final sobre as controvérsias interpretativas mais ambíguas e complexas, ditando as normatizações internas e os termos de serviço que caracterizam a chamada 'jurisprudência privada' elaborada unilateralmente pelas plataformas de tecnologia.

4 O DIREITO COMPARADO E OS MODELOS INTERNACIONAIS DE REGULAÇÃO DIGITAL

A natureza transnacional e descentralizada da internet impõe ao operador do direito o desafio de compreender que a regulação do ciberespaço não pode ser analisada exclusivamente sob a ótica de um único ordenamento jurídico. O fluxo informacional ignora fronteiras geográficas, o que torna o Direito Comparado uma ferramenta metodológica indispensável para a compreensão dos diferentes paradigmas de responsabilidade civil adotados no enfrentamento ao discurso de ódio. Historicamente, a comunidade internacional dividiu-se em três grandes eixos regulatórios: o modelo de imunidade ampla e autorregulação (liderado pelos Estados Unidos), o modelo de responsabilização rigorosa e remoção sumária (inaugurado pela Alemanha) e, mais recentemente, o modelo de regulação sistêmica e assimétrica (consolidado pela União Europeia). A análise dessas matrizes fornece o substrato teórico necessário para compreender a evolução e os impasses da legislação brasileira.

4.1 O paradigma norte-americano: a *Section 230* e a imunidade ampla

O modelo regulatório que mais influenciou o desenvolvimento inicial da internet em escala global, inclusive inspirando a redação originária do Artigo 19 do Marco Civil da Internet no Brasil, teve sua gênese nos Estados Unidos por meio da *Communications Decency Act* (CDA) de 1996, especificamente em sua Seção 230. A *Section 230* estabeleceu um princípio basilar que moldou a economia digital: “Nenhum provedor ou usuário de um serviço de computador interativo deve ser tratado como o editor ou porta-voz de qualquer informação fornecida por outro provedor de conteúdo” (Estados Unidos, 1996, tradução nossa).

Conforme leciona o jurista Jeff Kosseff (2019, p. 15), em sua obra *The Twenty-Six Words That Created the Internet*, essa disposição legal criou um “porto seguro” (*safe harbor*) quase absoluto para as plataformas digitais. O objetivo do legislador estadunidense, fortemente influenciado pela Primeira Emenda e pela cultura de proteção extremada à liberdade de expressão, era duplo: proteger a incipiente indústria da internet contra uma enxurrada de litígios inviabilizadores e incentivar a moderação de “bom samaritano”, garantindo que, se a plataforma

tentasse remover conteúdos de boa-fé, não atrairia para si a responsabilidade editorial pelo que restasse publicado.

A jurisprudência norte-americana consolidou essa imunidade logo no ano seguinte à promulgação da lei, no emblemático caso *Zeran v. America Online, Inc.* (1997). Na ocasião, o tribunal federal decidiu que a *Section 230* impedia que a plataforma AOL fosse responsabilizada por postagens difamatórias de terceiros, mesmo após ter sido notificada pela vítima. Essa jurisprudência ergueu um escudo intransponível para as *Big Techs*.

Contudo, a evolução do modelo de negócios das redes sociais para o que Shoshana Zuboff (2020) denomina de "capitalismo de vigilância" onde o engajamento através da indignação é monetizado colocou o modelo da *Section 230* em profunda crise. A imunidade criada para proteger fóruns passivos na década de 1990 passou a ser utilizada como subterfúgio para eximir grandes corporações de responsabilidade sobre o impulsionamento algorítmico ativo de discursos de ódio, radicalização política e desinformação, gerando um forte movimento bipartidário nos EUA pela reforma ou revogação desse dispositivo.

4.2 A Ruptura Alemã: o *NetzDG* e a privatização do controle

Em diametral oposição à tolerância estadunidense, a Alemanha profundamente marcada pelos traumas históricos do regime nacional-socialista e pela consciência dos perigos da retórica do ódio inaugurou um modelo de enfrentamento rígido com a aprovação do *Netzwerkdurchsetzungsgesetz* (Lei de Melhoria da Aplicação da Lei nas Redes Sociais, ou simplesmente *NetzDG*), que entrou em vigor em 2017.

O legislador alemão partiu da premissa de que a autorregulação havia fracassado miseravelmente na contenção de crimes de ódio e propaganda extremista durante a crise migratória europeia. O *NetzDG* rompeu com a lógica da ordem judicial prévia ao estabelecer prazos drásticos e sanções pecuniárias severas. A lei obriga as redes sociais com mais de dois milhões de usuários registrados na Alemanha a remover conteúdos "manifestamente ilegais" o que inclui incitação ao ódio, racismo e apologia ao nazismo no prazo peremptório de 24 horas após o recebimento de uma denúncia de qualquer usuário, sob pena de multas que podem chegar a 50 milhões de euros (ALEMANHA, 2017).

Embora aplaudido por sua contundência pragmática, o modelo alemão atraiu severas críticas da doutrina e de organismos internacionais de direitos humanos. David Kaye (2019, p. 55), ex-Relator Especial das Nações Unidas para a Liberdade de Expressão, adverte que o *NetzDG* promove o *chilling effect* (efeito inibitório) em sua forma mais aguda. Ao transferir o ônus da análise jurídica (o que é ou não um discurso de ódio) para funcionários terceirizados das plataformas, sob a ameaça de multas milionárias, o Estado alemão acaba por incentivar a adoção do chamado *overblocking* (bloqueio profilático excessivo e temerário de conteúdos). Diante da dúvida, a decisão corporativa mais barata e segura é sempre a censura do conteúdo, culminando em uma perigosa privatização da jurisdição estatal, onde empresas ditam os limites do debate público sem o crivo do devido processo legal.

4.3 A abordagem sistêmica da união europeia: o *Digital Services Act* (DSA)

Buscando superar o *laissez-faire* norte-americano e os riscos de censura privada do modelo alemão, a União Europeia aprovou, em 2022, o Digital Services Act (Regulamento de Serviços Digitais - DSA), estabelecendo o marco regulatório mais sofisticado e abrangente do mundo contemporâneo em relação ao ambiente digital (União Europeia, 2022). O DSA não foca apenas na remoção pontual do conteúdo, mas na regulação arquitetônica e algorítmica das plataformas.

A grande inovação do DSA é o conceito de regulação assimétrica. A legislação compreende que nem todas as plataformas possuem o mesmo potencial lesivo. Dessa forma, impõe obrigações progressivas, criando a categoria das *Very Large Online Platforms* (VLOPs) plataformas de dimensão muito grande, com mais de 45 milhões de usuários ativos mensais na Europa. Para essas gigantes de tecnologia, a lei impõe deveres de cuidado proativos que vão muito além do tradicional sistema de *notice-and-takedown* (notificação e retirada).

O DSA exige que as VLOPs realizem auditorias independentes anuais e avaliações rigorosas de "riscos sistêmicos", o que inclui analisar de que maneira a configuração de seus algoritmos de recomendação contribui para a disseminação de discurso de ódio, violência de gênero e manipulação eleitoral (União Europeia, 2022). Em vez de focar apenas na responsabilização subjetiva baseada na remoção de uma postagem individual, a União Europeia foca na responsabilidade objetiva pelo design do serviço. Exige-se transparência algorítmica: o usuário tem o direito de

saber por que está visualizando determinado conteúdo e as plataformas devem oferecer alternativas de feed que não sejam baseadas em perfilamento massivo de dados.

Em síntese, a análise do Direito Comparado demonstra um movimento pendular. O mundo afasta-se da imunidade irrestrita da década de 1990 em direção a paradigmas que exigem das plataformas um dever de cuidado estrutural e proativo. É exatamente essa transição de uma moderação passiva para o controle do risco sistêmico que lança as bases interpretativas para a crise do Artigo 19 do Marco Civil e para a posterior intervenção do Supremo Tribunal Federal brasileiro.

5 A VIRADA JURISPRUDENCIAL DO STF (2025-2026)

5.1 A decisão da Suprema Corte

A decisão do Supremo Tribunal Federal não realizou apenas uma releitura da norma, mas consolidou o reconhecimento da inconstitucionalidade parcial e progressiva do artigo 19 do Marco Civil da Internet. A Suprema Corte atestou a existência de um estado de omissão parcial, entendendo que a regra geral do dispositivo original não confere proteção suficiente a bens jurídicos constitucionais de alta relevância, à defesa dos direitos fundamentais e à democracia. Diante da insuficiência do modelo puramente reativo, o Tribunal fragmentou a responsabilidade civil das plataformas digitais em um regime duplo. Por um lado, nas hipóteses envolvendo alegações de crimes contra a honra, a regra protetiva original permanece aplicável, de modo que os provedores só poderão ser responsabilizados e obrigados a pagar indenização mediante o descumprimento de uma ordem judicial expressa para a remoção do conteúdo. Em contrapartida, instituiu-se um dever de cuidado rigoroso aplicável à circulação massiva de conteúdos ilícitos graves, abrangendo condutas como racismo, atos antidemocráticos, incitação ao terrorismo e ao suicídio, além de crimes praticados contra a mulher. Nesses cenários específicos de risco majorado, as plataformas assumem a responsabilidade civil independentemente de notificação judicial caso falhem em atuar de forma diligente e imediata para retirar o material lesivo do ar.

Para compreender essa virada hermenêutica, é imperativo observar os fundamentos que balizaram a alteração de entendimento da Corte. O Supremo Tribunal Federal reconheceu que a exigência de notificação judicial prévia era um modelo adequado à fase inicial da internet na qual a rede funcionava de forma estática e a informação dependia da busca ativa do usuário. Contudo, as redes sociais contemporâneas operam com base em algoritmos de recomendação, sistemas desenhados para reter a atenção do usuário e que, com frequência, acabam impulsionando conteúdos extremistas, desinformação e discursos de ódio, em virtude do alto engajamento que geram e da consequente lucratividade publicitária. Diante dessa nova realidade arquitetônica, o STF firmou o entendimento de que, ao recomendar ativamente um conteúdo ilícito ou lucrar diretamente com sua disseminação, a plataforma abdica da imunidade conferida pelo Artigo 19.

Nessas circunstâncias específicas, a responsabilidade cessa de ser puramente subjetiva e dependente de ordem judicial, passando a incidir o que a doutrina e a jurisprudência convencionaram chamar de dever de cuidado (Mendes, 2023).

Esse dever de cuidado, central na argumentação da Suprema Corte em 2025, significa que as empresas de tecnologia precisam adotar medidas preventivas e proativas para manter o ambiente digital seguro, de forma muito parecida com as regras europeias de serviços digitais. Sob a estrita ótica dogmática, a Suprema Corte fixou o entendimento imperativo de que os provedores de aplicação detentores da infraestrutura não podem se eximir da responsabilidade omissiva delegando o ônus da ilicitude unicamente ao usuário emissor direto do conteúdo lesivo, sobretudo quando restarem demonstrados o impulsionamento algorítmico e a amplificação sistemática da ofensa. A decisão fixou que, em casos de conteúdos flagrantemente ilícitos, como imagens de abuso infantil, terrorismo, racismo ou ataques diretos ao Estado Democrático de Direito, as plataformas têm a obrigação de agir de forma rápida para remover o material assim que tomarem conhecimento, mesmo que esse conhecimento venha por meio de notificações de usuários e não de um juiz (Silva, 2024). Se não o fizerem, podem ser processadas e condenadas a pagar indenizações.

Essa modulação dos efeitos do artigo 19 foi cuidadosamente desenhada para tentar equilibrar a balança entre a inovação tecnológica e a proteção dos direitos fundamentais. A Suprema Corte reconheceu o receio de que, ao forçar as plataformas a removerem conteúdos por conta própria, haveria um risco de "efeito inibitório" (*chilling effect*), onde as empresas apagariam postagens legítimas por medo de serem processadas, prejudicando a liberdade de expressão. Para evitar que as redes sociais se tornassem censores privados absolutos, o STF estabeleceu limites claros. A remoção sem ordem judicial só gera responsabilidade se o conteúdo for inquestionavelmente ilegal e estiver causando danos graves (Leonardi, 2019). Para debates políticos cotidianos, críticas sociais ou situações em que a ilicitude do conteúdo seja controversa, a regra originária do Marco Civil permanece hígida: a plataforma só será obrigada a remover a publicação mediante a expedição formal de uma ordem judicial de remoção ou provimento jurisdicional específico emitido por autoridade judiciária competente. O impacto dessa decisão de 2025 é gigantesco para o ecossistema digital brasileiro. Na prática, a Suprema Corte consolidou a tese de que a tecnologia não vive em um vácuo jurídico e que o lucro proveniente da

economia da atenção atrai para si os riscos do negócio. Ao afastar a aplicação cega do artigo 19 em situações de amplificação algorítmica de ilícitos, o STF harmonizou o Marco Civil da Internet com os princípios da dignidade da pessoa humana e da defesa das instituições democráticas previstos na Constituição Federal (Brasil, 2025). Essa virada jurisprudencial obriga as empresas de tecnologia a investirem pesado em moderação de conteúdo, equipes de segurança e inteligência artificial para detectar abusos, encerrando a era em que a internet era vista como uma terra sem lei e inaugurando um período de co-regulação, onde Estado, plataformas e sociedade dividem a responsabilidade de manter o ambiente online saudável e seguro para todos.

Esse entendimento, inicialmente delineado pela Suprema Corte ao longo de 2025, encontrou seu desfecho e consolidação definitiva no ano seguinte. Em sessão realizada no dia 17 de junho de 2026, o Supremo Tribunal Federal concluiu o julgamento dos embargos de declaração referentes aos Temas 987 e 533 da Repercussão Geral. Na referida data, a Corte decretou o trânsito em julgado imediato das ações, extinguindo a possibilidade de interposição de novos recursos e tornando a nova interpretação vinculante para todas as instâncias do Poder Judiciário nacional. Esse julgamento final encerrou as incertezas jurídicas sobre a aplicação do Marco Civil da Internet, estabelecendo não apenas a nova responsabilidade das plataformas, mas também prazos estruturais peremptórios para que as corporações de tecnologia se adequassem à nova dinâmica de proteção no ciberespaço.

Avançando nessa compreensão de que a internet deixou de ser uma terra sem lei, a Suprema Corte, em sua histórica decisão de 2025, aprofundou a análise sobre a arquitetura das redes sociais para tratar de um dos pontos mais sensíveis da atualidade: a complexa fronteira entre a proteção à liberdade de expressão e a proliferação do discurso de ódio. Durante muito tempo, as empresas de tecnologia utilizaram o princípio da liberdade de expressão como um escudo quase absoluto para se eximirem de qualquer responsabilidade sobre conteúdos discriminatórios, violentos ou antidemocráticos que circulavam em seus servidores. Contudo, os ministros entenderam que o discurso de ódio, especialmente quando impulsionado por sistemas automatizados, não é uma mera opinião impopular, mas sim uma ferramenta que corrói as bases da própria sociedade. O STF deixou muito claro que não existe direito fundamental absoluto, destacando que a liberdade de

manifestação do pensamento não pode servir de esconderijo para o cometimento de crimes ou para a destruição da reputação e da dignidade de minorias e instituições (Mulholland, 2024).

Um fato novo e determinante na fundamentação dos ministros durante esse julgamento foi a constatação empírica, apoiada em diversos relatórios técnicos anexados ao processo, de que o modelo de negócios das plataformas muitas vezes lucra diretamente com a radicalização. A Corte percebeu que postagens contendo discurso de ódio ou falsidades chocantes geram reações emocionais intensas nos usuários, sejam elas de apoio ou de repúdio. Essa revolta faz com que as pessoas passem mais tempo conectadas, comentando e compartilhando o material, o que permite às plataformas exibir mais anúncios e, conseqüentemente, aumentar seu faturamento. Diante dessa realidade inegável, a virada jurisprudencial se consolidou ao determinar que, se a empresa lucra com a viralização de um conteúdo de ódio por meio de seu algoritmo, ela passa a integrar a cadeia de responsabilização pelos danos causados. Não se trata mais de um mero quadro de avisos passivo, mas de um agente ativo que lucra com o engajamento gerado por conteúdos ilícitos (Schreiber, 2025).

A fim de conferir concretude a esse novo paradigma hermenêutico, a jurisprudência tem se socorrido da aplicação adaptada da teoria do risco do empreendimento ao ambiente digital. Historicamente sedimentada no direito do consumidor e no direito civil pátrio, referida teoria preconiza que o agente econômico que explora atividade no mercado tem o dever legal de responder, independentemente de culpa, pelos vícios e danos decorrentes de sua atuação. Ao transpor essa premissa estrutural para o ecossistema tecnológico, compreende-se que as falhas operacionais nos algoritmos de recomendação e moderação os quais frequentemente falham em mitigar conteúdos perigosos e acabam impulsionando discursos de ódio para usuários vulneráveis consubstanciam um risco intrínseco ao próprio modelo de negócios. Conforme leciona Ana Frazão (2024), a manutenção de um regime jurídico que exige da vítima de violações virtuais a obtenção prévia de um provimento jurisdicional, enquanto as plataformas se apropriam financeiramente do engajamento gerado pelo ilícito, afigura-se materialmente incompatível com a proteção constitucional da honra e da imagem. Dessa forma, a aplicação do risco-proveito justifica de forma robusta a

responsabilização civil das corporações que auferem as vantagens financeiras resultantes dessa dinâmica.

Ademais, a reorientação jurisprudencial não se restringiu à reparação pretérita de danos, estabelecendo também obrigações estruturantes de índole preventiva e prospectiva, com ênfase no dever de transparência algorítmica. As plataformas digitais foram compelidas a fornecer relatórios periódicos e detalhados acerca da operacionalização prática de seus sistemas de recomendação, explicitando os parâmetros técnicos empregados para o impulsionamento ou a supressão de postagens. Conforme pontuam Bioni e Nunes (2025), a opacidade algorítmica e a assimetria informacional constituem óbices severos à regulação eficaz do ecossistema digital e à proteção dos usuários. Para sanar essa falha, a exigência de transparência abarcou a obrigatoriedade de demonstração da alocação de equipes de moderação humana sediadas no país e devidamente capacitadas para compreender as nuances semânticas e o contexto sociocultural nacional. Tal medida visou superar o modelo ineficaz de moderação automatizada transnacional, que reiteradamente demonstrava inaptidão técnica para interpretar regionalismos e as especificidades sociológicas do discurso de ódio no Brasil.

Por fim, é fundamental destacar que essa virada na jurisprudência exigiu a criação de canais de comunicação diretos e ininterruptos entre as plataformas e as autoridades brasileiras. O STF determinou que a inércia reiterada em remover conteúdos que configuram crimes graves contra o Estado Democrático de Direito ou violações severas a direitos humanos, após notificação de órgãos competentes, pode resultar em multas pesadíssimas, proporcionais ao faturamento global da empresa, e não apenas no valor simbólico de uma indenização individual. Com isso, o Tribunal não revogou o artigo 19 do Marco Civil da Internet, mas o reinterpretou à luz da Constituição de 1988 e da realidade de 2025. A regra da ordem judicial prévia foi mantida como o padrão para a grande maioria das discussões online, preservando a liberdade de expressão em seu sentido mais amplo. A exceção, contudo, foi desenhada de forma cirúrgica para as situações em que a própria infraestrutura tecnológica da plataforma se torna a arma que potencializa o dano, consolidando um novo paradigma de responsabilidade civil digital no Brasil.

5.2 Fundamentos da decisão e os votos dos ministros

Para compreender a magnitude dessa transformação no cenário jurídico brasileiro, é indispensável mergulhar nos fundamentos que embasaram a decisão e na forma como os votos dos ministros do Supremo Tribunal Federal foram construídos e articulados. A redefinição da responsabilidade das plataformas digitais não ocorreu de maneira abrupta ou sem um denso debate constitucional, mas foi o resultado da soma de diferentes visões dentro da Corte, que convergiram para um ponto comum: a inadequação do modelo antigo frente às novas tecnologias. Ao longo das sessões de julgamento, o argumento principal que unificou o plenário foi a constatação de que as redes sociais deixaram de ser meros "muros de recados" passivos para se tornarem os principais editores e curadores da informação no mundo moderno. Os relatores dos recursos extraordinários destacaram em seus votos que a redação original do artigo 19 do Marco Civil da Internet foi pensada para uma realidade tecnológica que não existe mais, conhecida como Web 2.0, onde a plataforma apenas hospedava o conteúdo. No cenário contemporâneo, por meio do uso de inteligência artificial, a plataforma digital dota-se de discricionariedade algorítmica autônoma para determinar a filtragem, o alcance, a temporalidade e a capilaridade da distribuição do conteúdo informacional, o que a coloca em uma posição de responsabilidade direta sobre o impacto das postagens amplificadas (Doneda, 2024).

Durante a leitura dos votos, um dos pilares mais fortes para a mudança de entendimento foi a necessidade urgente de combater o discurso de ódio e proteger a democracia, tema que toca diretamente no núcleo da tensão entre a liberdade de expressão e a preservação dos direitos humanos. Alguns ministros, em falas contundentes, argumentaram que a liberdade de manifestação do pensamento, embora seja um dos direitos mais sagrados da Constituição de 1988, não pode ser utilizada como um escudo intransponível para a prática de crimes, como racismo, homofobia, incitação à violência e ataques às instituições. Foi amplamente debatido que o discurso de ódio possui um efeito destrutivo não apenas para a vítima individual, mas para toda a coletividade, pois tenta calar e excluir grupos inteiros do debate público. Nesse sentido, os votos firmaram a tese de que exigir uma ordem judicial prévia para remover conteúdos de ódio explícito e massivamente impulsionados era o mesmo que exigir que o Estado enxugasse gelo, pois a

velocidade de disseminação da ofensa na internet é infinitamente superior ao tempo de resposta do Poder Judiciário (Sarlet, 2025). Assim, transferiu-se para as empresas de tecnologia o dever de estancar a sangria logo no início, sob pena de conivência.

Outro argumento central explorado nos votos dos ministros foi a análise econômica do modelo de negócios das chamadas "*Big Techs*", utilizando a teoria do risco do proveito. Ministros com forte atuação na área do direito privado e econômico trouxeram para o centro do debate a ideia de que a "economia da atenção" lucra deliberadamente com a polarização. Os votos detalharam como os algoritmos são programados para privilegiar postagens que causam indignação, medo ou raiva, simplesmente porque esses sentimentos mantêm os usuários conectados por mais tempo, gerando mais visualizações de anúncios publicitários. A argumentação jurídica foi simples e direta: se a empresa de tecnologia auferir lucros bilionários a partir de um sistema de recomendação que, por sua própria natureza, favorece a viralização de conteúdos ilícitos e discursos de ódio, ela deve assumir os riscos dessa operação. O Supremo Tribunal Federal consolidou o entendimento de que não é justo, nem legalmente aceitável, privatizar os lucros gigantescos da rede social e socializar os prejuízos e danos morais com a sociedade e com as vítimas dos ataques virtuais (Tepedino; Silva, 2024).

Apesar da forte inclinação para responsabilizar as plataformas, os votos também trouxeram ponderações essenciais para evitar que a internet se tornasse um ambiente de censura corporativa. Alguns ministros, em votos complementares e de ressalva, demonstraram uma preocupação legítima com o chamado "efeito inibitório". Eles argumentaram que, se a responsabilidade das redes sociais fosse ampla e irrestrita, o caminho mais fácil e barato para as empresas seria apagar qualquer conteúdo que recebesse uma denúncia, mesmo sendo uma crítica política legítima ou uma piada inofensiva, apenas para evitar processos judiciais. Para solucionar esse dilema, os votos construíram uma saída intermediária e inteligente: a modulação da regra. O STF decidiu que as plataformas só podem ser responsabilizadas sem ordem judicial quando o conteúdo for inquestionavelmente ilícito, ou seja, quando não houver qualquer dúvida sobre a violação da lei, como nos casos de pedofilia, apologia ao terrorismo ou racismo escancarado. Em casos onde há zona cinzenta, onde a postagem pode ser interpretada de várias formas e exige uma análise mais profunda do contexto, a plataforma não tem o dever de ser o juiz

da verdade, e a necessidade de ordem judicial prévia do artigo 19 permanece intacta (Abboud, 2025).

A soma de todos esses votos formou uma jurisprudência híbrida e moderna, que colocou o Brasil na vanguarda da regulação digital global. Os ministros conseguiram costurar um acordo interpretativo que reconhece a liberdade de expressão como regra, mas que não fecha os olhos para os abusos algorítmicos. O resultado do julgamento deixou como legado o entendimento de que a neutralidade da rede não significa cegueira das plataformas. As empresas passaram a ter o dever constitucional de cuidado, precisando agir com diligência, transparência e responsabilidade social. Os fundamentos da decisão de 2025 mostraram que o Direito não é estático; ele precisa caminhar na mesma velocidade das inovações tecnológicas para garantir que ferramentas criadas para conectar pessoas não sejam sequestradas por grupos que desejam espalhar o ódio e a desinformação, garantindo, assim, que a dignidade da pessoa humana prevaleça tanto no mundo físico quanto no ambiente virtual.

5.3 A análise da inconstitucionalidade

A consolidação do novo entendimento do Supremo Tribunal Federal abre espaço para uma profunda e necessária discussão crítica acerca da constitucionalidade e dos limites práticos desse novo panorama de responsabilização das plataformas digitais no combate ao discurso de ódio. O grande nó górdio dessa questão reside em determinar se a imposição de um dever de cuidado proativo às empresas de tecnologia viola, ou não, o núcleo essencial da liberdade de expressão garantido pela Constituição Federal de 1988. Por um lado, defensores da constitucionalidade da virada jurisprudencial argumentam que a interpretação dada pela Corte cumpre o mandamento constitucional de proteção à dignidade da pessoa humana e de repúdio ao racismo e a outras formas de discriminação. Sob essa ótica, o antigo modelo do artigo 19 do Marco Civil da Internet havia se tornado inconstitucional por omissão, uma vez que protegia de forma desproporcional o poder econômico das grandes empresas de tecnologia em detrimento dos direitos de personalidade das vítimas e da integridade do próprio debate democrático.

Por outro lado, críticos dessa guinada jurisprudencial apontam para riscos severos de inconstitucionalidade por via transversa, especialmente no que diz respeito à criação de um ambiente de censura privada. Ao transferir para entidades corporativas internacionais o poder e o dever de julgar o que constitui ou não um discurso de ódio inaceitável, o Estado corre o risco de terceirizar sua função jurisdicional. As plataformas digitais, movidas pelo receio legítimo de sofrerem sanções financeiras pesadíssimas e processos de responsabilidade civil em massa, tendem a adotar uma postura extremamente conservadora e defensiva. Na prática, isso significa que sistemas de inteligência artificial e moderadores humanos passarão a remover conteúdos de forma preventiva ao menor sinal de polêmica. Esse fenômeno atinge diretamente o debate político, a sátira, as manifestações artísticas e as críticas sociais legítimas, que muitas vezes utilizam termos fortes ou provocativos que podem ser facilmente confundidos por um algoritmo com discurso de ódio.

A análise crítica desse novo ecossistema revela que a linha que separa o combate legítimo à discriminação e a mordida ao pensamento divergente tornou-se perigosamente tênue. O discurso de ódio, diferentemente de ilícitos manifestos e objetivos como a pornografia infantil, é um conceito juridicamente fluido, altamente dependente de contexto, ironia, história e cultura local. Quando o Judiciário exige que as redes sociais realizem essa filtragem complexa de forma célere e sob pena de punição, ele acaba por incentivar a eliminação do dissenso. O cidadão comum, percebendo que suas postagens estão sendo constantemente derrubadas ou rebaixadas pelos mecanismos invisíveis de distribuição das plataformas, pode passar a se autocensurar, enfraquecendo o pluralismo de ideias que serve de base para qualquer sociedade democrática.

Portanto, o cenário pós-2025 desenha um desafio interpretativo sem precedentes para o direito digital brasileiro. A mudança de postura do Supremo Tribunal Federal buscou corrigir uma assimetria de poder onde as empresas lucravam com o caos informativo sem pagar o preço social por ele, mas a solução adotada traz consigo o efeito colateral de empoderar essas mesmas empresas como os novos censores do espaço público virtual. A constitucionalidade dessa nova engenharia jurídica dependerá, exclusivamente, de como os tribunais de instâncias inferiores aplicarão as balizas fixadas pela Suprema Corte, exigindo transparência absoluta nos critérios de moderação e garantindo que o usuário afetado tenha

sempre direito ao devido processo legal e à ampla defesa dentro da própria plataforma, sob o risco de transformarmos a internet em um espaço de conformismo e controle corporativo.

Nesse sentido, o aprofundamento dessa análise de inconstitucionalidade passa obrigatoriamente pela ausência de um devido processo legal no ambiente corporativo privado. Quando o Estado delega às plataformas o poder de julgar e punir a expressão do pensamento, cria-se um sistema punitivo paralelo onde as regras não são votadas pelo parlamento, mas redigidas em termos de uso frequentemente obscuros, extensos e impostos de forma unilateral. O cidadão que tem seu conteúdo removido ou sua conta suspensa sob a alegação de estar propagando discurso de ódio encontra-se, na imensa maioria das vezes, diante de um tribunal kafkiano e inatingível. As notificações de remoção costumam ser mensagens padronizadas e genéricas, emitidas por sistemas automatizados que não explicam qual palavra, imagem ou contexto específico gerou a punição. Essa dinâmica impossibilita o exercício efetivo do contraditório e da ampla defesa, princípios consagrados no ordenamento jurídico brasileiro. Esse cenário afronta diretamente garantias constitucionais básicas, pois a punição virtual, que na sociedade conectada atual muitas vezes significa a exclusão brutal do indivíduo do debate público contemporâneo ou a destruição de seu meio de expressão, de trabalho e de socialização, é aplicada de forma sumária sem que haja a figura de um julgador imparcial ou a chance real de recurso a uma instância humana devidamente qualificada para reverter eventuais equívocos.

Avançando nessa perspectiva, é preciso lançar um olhar técnico e crítico sobre a própria natureza das ferramentas tecnológicas empregadas para cumprir essa nova ordem jurisprudencial focada no dever de cuidado. A exigência de proatividade e velocidade no combate ao discurso de ódio empurra, inevitavelmente, as redes sociais para uma dependência quase total da inteligência artificial e dos filtros algorítmicos. No entanto, é de conhecimento consolidado na área de tecnologia e direitos humanos que esses algoritmos não são neutros; eles herdam, reproduzem e muitas vezes amplificam os vieses estruturais e preconceitos presentes nos imensos bancos de dados com os quais foram treinados. Consequentemente, surge um paradoxo cruel na moderação automatizada de conteúdo: ao tentar proteger grupos minoritários e vulneráveis do discurso de ódio, as plataformas frequentemente acabam silenciando esses mesmos grupos.

Movimentos sociais que denunciam o racismo ou a violência institucional, por exemplo, muitas vezes precisam reproduzir as ofensas que sofreram em suas denúncias públicas para expor a gravidade da situação. Contudo, os sistemas robóticos, incapazes de compreender nuances linguísticas, ironia, sarcasmo ou a denúncia sociológica por trás de uma palavra agressiva, tratam o material de forma literal como violação de diretrizes e banem a vítima em vez de identificar o agressor. Essa moderação cega e mecanizada, exigida pelo temor do risco do empreendimento e da responsabilização civil imposta pela Suprema Corte, pode violar a Constituição de forma transversal ao promover uma discriminação algorítmica disfarçada de assepsia, asfixiando as vozes de ativistas e minorias que historicamente já sofrem com a invisibilidade no espaço público tradicional.

Diante desses severos entraves estruturais e operacionais, a discussão sobre a constitucionalidade da virada jurisprudencial promovida pelo Supremo Tribunal Federal em 2025 não pode se encerrar de maneira simplista em um embate maniqueísta entre proteger cegamente a liberdade de expressão ou combater de forma absoluta o discurso de ódio. A reflexão jurídica precisa evoluir para a construção urgente de um modelo de correção inteligente, transparente e auditável. Para que a decisão da Corte seja considerada material e procedimentalmente constitucional no longo prazo, não basta exigir que as grandes corporações de tecnologia apaguem conteúdos ilícitos com rapidez; é absolutamente imperativo que o Estado exija total transparência sobre como essa exclusão em massa está sendo realizada. A sociedade civil, a academia e as instituições democráticas precisam ter acesso claro e desimpedido aos critérios lógicos que orientam o design e as decisões desses algoritmos moderadores. Torna-se indispensável a formulação de leis e regulamentações que exijam a realização de auditorias externas periódicas e independentes, permitindo que pesquisadores, órgãos do Ministério Público e entidades de defesa do consumidor analisem o comportamento interno das redes sociais. O objetivo dessa supervisão pública é verificar se a moderação está sendo aplicada de forma isonômica e equilibrada ou se os mecanismos estão servindo, na surdina, para silenciar adversários políticos, suprimir o debate democrático legítimo e impor uma hegemonia cultural estritamente alinhada aos interesses comerciais e acionários das "*Big Techs*". Sem a consolidação dessa camada essencial de transparência ativa, a delegação do dever de monitoramento e cuidado transmuta-se de uma medida protetiva em uma carta

branca manifestamente inconstitucional para a manipulação invisível da opinião pública e da vontade popular.

Por fim, ao encerrar esta análise crítica sobre o capítulo da responsabilização dos provedores de aplicação frente aos conteúdos gerados por terceiros, conclui-se que o novo panorama jurídico instaurado pelo Supremo Tribunal Federal representa um amadurecimento complexo, doloroso, porém historicamente inevitável da relação entre o Direito, as garantias fundamentais e a evolução exponencial da tecnologia da informação no Brasil. A superação do modelo irresponsável e puramente reativo originalmente previsto no texto do Marco Civil da Internet não configura, por si só, uma violação do texto constitucional, mas sim um esforço hermenêutico monumental para adequar a proteção da dignidade humana, da honra e da democracia à ferocidade destrutiva da economia da atenção impulsionada por algoritmos preditivos. O verdadeiro desafio jurídico e social que se apresenta agora não reside mais em fechar os olhos e negar o fato de que as plataformas digitais possuem controle e responsabilidade direta sobre a arquitetura do ecossistema informacional que criaram e monetizam. O desafio atual é calibrar milimetricamente essa responsabilidade delegada para que o remédio não se revele mais tóxico e letal que a própria doença que se propõe a curar. O sucesso dessa complexa empreitada regulatória dependerá, nos próximos anos, da resiliência e da capacidade analítica dos operadores do direito, da doutrina e da jurisprudência em vigiar incansavelmente as fronteiras da moderação privada de conteúdo. Apenas por meio dessa vigilância constante será possível garantir que o espaço digital não se consolide definitivamente como um imenso latifúndio corporativo onde a liberdade de expressão é uma concessão tolerada apenas enquanto for favorável aos negócios, assegurando que a internet permaneça preservada como uma verdadeira arena democrática, plural e vibrante, onde a divergência ideológica e o debate de ideias possam prosperar de forma livre da toxicidade coordenada, da disseminação lucrativa do ódio e da sombra sufocante da censura algorítmica invisível.

5.4 O contexto de instabilidade e a chegada dos casos ao Supremo

A redação do artigo 19 do Marco Civil da Internet (Lei nº 12.965/2014) instituiu no ordenamento jurídico brasileiro a regra geral de que os provedores de aplicação apenas poderiam ser responsabilizados civilmente por danos decorrentes

de conteúdo gerado por terceiros se, após ordem judicial específica, não tomassem as providências para tornar o conteúdo indisponível. Embora a intenção original do legislador fosse resguardar a liberdade de expressão e evitar a censura privada, a aplicação literal e irrestrita desse dispositivo passou a gerar profunda instabilidade jurídica e social.

Com a sofisticação dos algoritmos e a proliferação em massa de discursos de ódio, campanhas de desinformação e ataques a direitos da personalidade, a exigência inexorável de ordem judicial prévia revelou-se um escudo desproporcional para as *Big Techs*. A jurisprudência dos tribunais estaduais começou a oscilar vertiginosamente: enquanto algumas câmaras cíveis aplicavam o artigo 19 de forma estrita, outras o afastavam sob o argumento de violação direta à proteção constitucional da dignidade da pessoa humana e da honra (art. 5º, incisos V e X, da CRFB/88). Diante dessa fragmentação do entendimento nacional e da inércia do Poder Legislativo em atualizar a norma frente à nova realidade tecnológica, o Supremo Tribunal Federal (STF) foi instado a pacificar a controvérsia, reconhecendo a repercussão geral de dois casos paradigmáticos: os Temas 987 e 533.

5.5 Estudo de caso I: dissecando o tema 987 (RE 1.037.396)

O Recurso Extraordinário 1.037.396, que originou o Tema 987 da Repercussão Geral, sob a relatoria do Ministro Dias Toffoli, debruçou-se centralmente sobre a constitucionalidade do artigo 19 do Marco Civil da Internet. O fato gerador do litígio envolvia a manutenção de perfis falsos em redes sociais, situação na qual a vítima, mesmo após notificar extrajudicialmente a plataforma apontando a flagrante ilicitude do conteúdo e a usurpação de sua identidade, viu seu pedido de remoção negado sob a justificativa de ausência de ordem judicial.

O núcleo do debate no Tema 987 não é a censura de opiniões divergentes, mas a análise da proporcionalidade da imunidade conferida aos provedores. O voto e os debates na Suprema Corte evidenciaram que condicionar a remoção de conteúdos manifestamente ilícitos como a pedofilia, o racismo estrutural e a criação de perfis **fakes** para a prática de crimes a uma prévia e demorada decisão judicial significa, na prática, cancelar a perpetuação do dano.

A exigência do artigo 19, quando aplicada a infrações inequívocas, esvazia o núcleo essencial do direito à reparação e à proteção da imagem.

A tese em construção pela Corte reflete o entendimento de que o modelo de *notice and takedown* (notificação e retirada) não pode ser sumariamente descartado pelo legislador infraconstitucional em nome de uma suposta defesa absoluta da liberdade de expressão. Como bem pontua a doutrina especializada em direitos fundamentais, não existe direito absoluto no ordenamento jurídico brasileiro. A ponderação de interesses exige que, em casos de violação cristalina a direitos fundamentais, a plataforma assumo o ônus de agir mediante simples notificação da vítima, sob pena de responsabilidade solidária.

5.6 Estudo de caso II: dissecando o tema 533 (RE 1.057.258)

De forma complementar e simultânea, o STF debruçou-se sobre o Recurso Extraordinário 1.057.258 (Tema 533), relatado pelo Ministro Luiz Fux. Este caso, cujas raízes remontam a litígios envolvendo antigas redes sociais (como o Orkut) e a criação de comunidades difamatórias, carrega uma discussão dogmática profunda sobre o dever de fiscalização prévia e a natureza do serviço prestado pelos provedores de aplicação.

A análise do Tema 533 exige a dissecação do comportamento do provedor diante da ciência inequívoca do dano. Enquanto o Tema 987 foca estritamente na constitucionalidade da exigência de ordem judicial do Marco Civil, o Tema 533 explora a base principiológica da responsabilidade civil extracontratual no ambiente virtual. Nos debates conduzidos pelo Ministro Fux, consolidou-se a premissa de que a liberdade de expressão não pode ser utilizada como biombo para a prática de ilícitos cibernéticos.

A Suprema Corte assentou que, embora não se possa exigir das plataformas um monitoramento prévio genérico e universal de todas as postagens o que configuraria censura prévia e inviabilizaria tecnicamente a internet, essa dispensa de monitoramento não as exime de agir com diligência assim que tomam conhecimento da lesão. O provedor que se omite após ser formal e inequivocamente notificado sobre a existência de conteúdo difamatório ou de ódio passa a ser um agente causador do dano por omissão (teoria da causalidade adequada), atraindo para si a responsabilidade civil subsidiária ou solidária.

5.7 O dever fiduciário de cuidado e a inaplicabilidade da responsabilidade objetiva

Síntese da análise conjunta dos Temas 987 e 533 pelo STF culmina na consagração de um novo paradigma para o Direito Digital brasileiro: a aplicação direta da teoria do risco do empreendimento (ou risco-proveito) aos provedores de aplicação, conjugada com a exigência de um dever de cuidado objetivo.

Historicamente cunhada pela doutrina civilista para as relações de consumo, a teoria do risco-proveito, magistralmente defendida por Sérgio Cavalieri Filho, estabelece que todo aquele que se dispõe a exercer alguma atividade no mercado de consumo tem o dever de responder pelos eventuais vícios ou defeitos dos bens e serviços fornecidos, independentemente de culpa. O STF, ao julgar os referidos Temas, transpôs essa lógica para a arquitetura das redes sociais.

As *Big Techs* deixaram de ser meros repositórios passivos (murais de recados) para se tornarem curadoras ativas de conteúdo por meio de inteligência artificial. Se a plataforma estrutura seu modelo de negócios e auferir lucros bilionários através de algoritmos que impulsionam o engajamento frequentemente favorecendo discursos polarizados e de ódio pela sua alta capacidade de retenção de atenção, ela internaliza o risco dessa atividade. Portanto, a virada jurisprudencial do Supremo Tribunal Federal decreta o fim da irresponsabilidade sistêmica: o lucro obtido com a economia da atenção atrai, indissociavelmente, o dever de cuidado profilático e a responsabilidade civil em caso de inércia frente a violações de direitos fundamentais.

5.8 A consolidação final em 2026: prazos e modulação de efeitos

O desfecho definitivo da virada jurisprudencial operada pelo Supremo Tribunal Federal ocorreu na sessão do dia 17 de junho de 2026, com a conclusão do julgamento dos embargos de declaração atrelados aos Temas 987 e 533 da Repercussão Geral. Ao decretar o trânsito em julgado imediato dessas ações, a Suprema Corte encerrou as margens para questionamentos e pacificou o novo regime de responsabilização civil em todo o território nacional. O aspecto mais incisivo desta consolidação foi a determinação de um prazo peremptório de 60 dias, a contar da publicação da ata de julgamento, para que as plataformas e provedores

de aplicação promovam a implementação das medidas estruturais necessárias à adequação ao novo dever de cuidado.

No que tange à segurança jurídica e ao direito intertemporal, o STF promoveu uma modulação de efeitos de extrema relevância dogmática. A Corte definiu que a decisão tem eficácia prospectiva (*ex nunc*), valendo a partir da data de publicação da ata do julgamento de mérito (5 de agosto de 2025). Contudo, os ministros abriram uma exceção vital para coibir a perpetuação de danos nas redes: a modulação temporal não alcança os chamados atos continuados ou permanentes.

Na prática processual e tecnológica, isso significa que publicações antigas contendo discursos de ódio, violações à honra ou ameaças antidemocráticas, que continuem armazenadas e ativas nos servidores das redes sociais, configuram uma infração contínua. Ao manter esses conteúdos ilegais acessíveis ao público e retroalimentando os algoritmos de recomendação, as corporações de tecnologia atraem a aplicação imediata da nova tese de responsabilização, não podendo invocar a data original da postagem para se eximirem do dever de remoção ou da responsabilidade solidária.

6 A TUTELA COLETIVA E A EFETIVAÇÃO DA RESPONSABILIDADE CIVIL DIGITAL

A virada jurisprudencial promovida pelo Supremo Tribunal Federal em 2025, ao afastar a incidência cega do Artigo 19 do Marco Civil da Internet para casos de impulsionamento algorítmico de conteúdos manifestamente ilícitos, representou um avanço dogmático inquestionável. Contudo, no âmbito do Direito, o reconhecimento material de um direito não garante, por si só, a sua eficácia prática. A proteção da dignidade da pessoa humana e o combate ao discurso de ódio no ciberespaço exigem mecanismos processuais capazes de lidar com a escala massiva, a velocidade cinestésica e a pulverização dos danos gerados pelas redes sociais. É nesse cenário de hiperconectividade que a tutela jurisdicional coletiva, capitaneada pelo Ministério Público, assume um papel de protagonismo inafastável.

6.1 A Insuficiência da Tutela Individual face às *Big Techs*

O sistema clássico de responsabilidade civil foi estruturado a partir de uma lógica essencialmente bipolar e individualista: de um lado, o ofensor; do outro, a vítima perfeitamente individualizada. No entanto, o fenômeno do discurso de ódio (*hate speech*) digital subverte essa lógica. As manifestações de desprezo e intolerância raramente visam atingir apenas um indivíduo isolado; seu objetivo central é corroer a igualdade e estigmatizar grupos inteiros, fundamentando-se em marcadores de raça, religião, gênero ou orientação sexual. Consequentemente, o dano transborda a esfera individual e atinge a coletividade, configurando-se como uma grave violação a direitos difusos e coletivos *stricto sensu*.

Diante de ataques coordenados e impulsionados por algoritmos projetados especificamente para a ‘economia da atenção’, a via judicial individual revela-se materialmente insuficiente e processualmente ineficiente. A vítima individual que tenta processar corporações multinacionais de tecnologia depara-se com uma brutal assimetria de poder técnico, informacional e econômico. As chamadas *Big Techs* operam sob um manto de extrema opacidade algorítmica, o que inviabiliza a produção de provas pelo usuário comum sobre como e por que o conteúdo ofensivo foi impulsionado. Ademais, a propositura de milhares de ações individuais pulverizadas nos juizados especiais buscando a remoção de URLs específicas resulta no mero ‘enxugamento de gelo’. Enquanto o indivíduo litiga por

meses para derrubar uma única postagem, o material já foi replicado em escala exponencial por múltiplos servidores, perpetuando o dano psíquico e a exclusão social da vítima.

6.2 O papel do ministério público e a ação civil pública

A Constituição Federal de 1988 inovou ao alçar o Ministério Público à condição de instituição permanente e essencial à função jurisdicional do Estado, atribuindo-lhe, em seu artigo 129, inciso III, a função institucional de promover o inquérito civil e a Ação Civil Pública (ACP) para a proteção dos interesses difusos e coletivos (Brasil, 1988). No enfrentamento à toxicidade do ambiente digital, a ACP (Lei nº 7.347/1985) consagra-se como a via processual idônea para domesticar a fúria das redes e exigir o cumprimento do dever de cuidado fixado pelo STF.

A legitimação do Ministério Público permite uma alteração fundamental de foco: sai de cena a análise microscópica de uma postagem isolada e entra em pauta a análise macroscópica do modelo de governança da plataforma. O Parquet possui competência e envergadura institucional para requisitar perícias técnicas, firmar Termos de Ajustamento de Conduta (TAC) e ajuizar demandas estruturais.

Na prática processual, a atuação do Ministério Público por meio da Ação Civil Pública permite uma alteração fundamental de foco: afasta-se a análise microscópica de uma postagem isolada e adentra-se o escrutínio macroscópico do modelo de governança e do design algorítmico da plataforma. O Parquet possui competência e envergadura institucional para exigir a imposição de obrigações de fazer estruturais que sanem o ecossistema informacional em sua raiz tecnológica, antes que o ilícito viralize. No âmbito do combate ao discurso de ódio, isso se traduz em pleitos judiciais voltados a compelir as *Big Techs* a: a) alterarem a programação de seus sistemas de recomendação que sabidamente priorizam o engajamento através da hostilidade; b) estabelecerem equipes de moderação humana baseadas fisicamente no Brasil, dotadas de capacidade técnica para interpretar nuances semânticas e o contexto sociocultural do preconceito local; e c) fornecerem relatórios periódicos de transparência ativa sobre as providências automáticas adotadas contra a desinformação coordenada e os ataques às instituições democráticas. A intervenção ministerial, portanto, transmuta-se de uma mera fiscalização reativa

para uma verdadeira atividade de regulação de risco sistêmico, essencial para harmonizar a exploração econômica das redes com os ditames constitucionais.

6.3 O dano moral coletivo no ciberespaço

Outro reflexo direto da insuficiência do modelo individual e da aplicação da teoria do risco do empreendimento (risco-proveito) ao modelo de negócios das plataformas é o reconhecimento e a quantificação do dano moral coletivo. O Superior Tribunal de Justiça (STJ) já possui jurisprudência consolidada no sentido de que o dano moral coletivo ocorre quando há uma lesão injusta e intolerável a valores fundamentais da sociedade, provocando repulsa e indignação na consciência coletiva.

Quando uma plataforma digital se omite de forma reiterada diante da proliferação de campanhas de ódio neonazistas, ataques racistas estruturados ou injúrias homofóbicas, e, simultaneamente, auferir vultosos lucros decorrentes da retenção dos usuários estimulada por tais reações emocionais negativas, há uma evidente degradação do patrimônio moral da sociedade brasileira. A omissão corporativa, sob a égide da teoria do risco-proveito, agride o princípio da solidariedade e a própria dignidade da pessoa humana. A lesão, portanto, deixa de ser um mero aborrecimento individual e passa a configurar um dano socializável, uma vez que a infraestrutura técnica da empresa foi utilizada como vetor de amplificação de preconceitos estruturais que fragmentam o tecido social.

6.4 A Legitimidade das Associações, a Defensoria Pública e a Litigância Estratégica

Embora o Ministério Público detenha inegável protagonismo na tutela de direitos transindividuais, a efetivação da responsabilidade civil das plataformas digitais encontra um reforço indispensável na atuação de outros atores legitimados. O microsistema processual coletivo brasileiro, consolidado pelo Código de Defesa do Consumidor e pela Lei da Ação Civil Pública, confere legitimidade ativa às associações civis e à Defensoria Pública, cuja atuação no ciberespaço tem se mostrado cada vez mais decisiva.

A litigância estratégica promovida por organizações não governamentais (ONGs) e associações de defesa de direitos humanos possui uma vantagem

peculiar: a especialização temática. Conforme pontua Zaneti Jr. (2021), a chamada representatividade adequada (*adequacy of representation*) é um requisito vital para o sucesso da tutela coletiva. Quando uma associação voltada exclusivamente para o combate ao racismo ou para a defesa dos direitos LGBTQIAPN+ ingressa com uma demanda coletiva contra uma grande empresa de tecnologia, ela leva ao Poder Judiciário não apenas a legitimidade formal, mas um acervo probatório denso sobre como o discurso de ódio afeta concretamente o cotidiano e a saúde mental daquele grupo vulnerável.

Nesse mesmo viés, a intervenção de terceiros sob a modalidade de *amicus curiae* (amigo da corte) ganha contornos de extrema relevância em litígios digitais. Em demandas que exigem a compreensão de algoritmos complexos, viés de aprendizado de máquina e moderação de conteúdo, a contribuição técnica da sociedade civil e de centros de pesquisa em tecnologia e sociedade auxilia o magistrado a proferir decisões mais conectadas com a realidade cibernética. Segundo Didier Jr. (2022), o *amicus curiae* pluraliza o debate processual e democratiza a decisão judicial, evitando que sentenças proferidas contra gigantes da tecnologia sejam inexequíveis ou tecnicamente falhas.

6.5 O processo estrutural como resposta à complexidade algorítmica

Um dos maiores desafios enfrentados pela jurisdição moderna no combate ao discurso de ódio é a limitação do modelo tradicional de decisões judiciais. Sentenças clássicas operam de forma binária (pagar/não pagar, fazer/não fazer), o que se mostra flagrantemente inadequado para resolver falhas sistêmicas na moderação de bilhões de postagens diárias. Para solucionar essa defasagem, a doutrina tem invocado a aplicação do Processo Estrutural aos litígios envolvendo redes sociais.

O litígio estrutural caracteriza-se pela necessidade de intervenção contínua, dialógica e prospectiva do Poder Judiciário para reorganizar instituições sejam elas entidades públicas ou corporações privadas que violam direitos fundamentais de forma crônica e generalizada. No ambiente virtual, o processo estrutural não busca apenas punir um ilícito pretérito ou remover uma postagem isolada, mas sim reformular a estrutura causal e sistêmica que permitiu a proliferação daquele conteúdo deletério. Reconhece-se, assim, que a resposta para

a toxicidade das redes não reside em provimentos judiciais estáticos, mas na modificação das diretrizes operacionais que comandam o fluxo de informações.

Aplicada à responsabilidade civil digital pós-*virada* jurisprudencial do STF, a tutela estrutural permite que o magistrado, em sede de Ação Civil Pública, expeça decisões em cascata e determine a elaboração de planos de transição e metas de conformidade pelas plataformas digitais. Em vez de simplesmente aplicar condenações pecuniárias retroativas, o juiz instaura um procedimento flexível no qual a empresa de tecnologia é compelida a apresentar, em etapas cronológicas, um plano de aprimoramento de seus filtros de inteligência artificial, submeter-se a auditorias externas independentes e redesenhar seus sistemas algorítmicos de recomendação. Essa maleabilidade procedimental, típica das medidas estruturais, revela-se como a única via processual capaz de gerar modificações reais nas políticas internas de governança corporativa das *Big Techs*, equilibrando de forma *perene* a inovação tecnológica com o estrito cumprimento do dever de cuidado.

6.6 A resolução extrajudicial e os Termos de Ajustamento de Conduta (TAC)

A velocidade do fluxo de informações na internet é, por natureza, assimétrica em relação ao tempo exigido para o trâmite regular de um processo judicial. Diante do princípio da inafastabilidade da jurisdição, as tutelas de urgência são fundamentais, mas a judicialização extrema carrega altos custos e imprevisibilidade. Por essa razão, a busca por mecanismos extrajudiciais de resolução de litígios coletivos tornou-se uma diretriz preferencial na regulação da internet.

O Termo de Ajustamento de Conduta (TAC), previsto no artigo 5º, parágrafo 6º, da Lei nº 7.347/1985, desponta como um instrumento de imensa valia no cenário do Direito Digital. O TAC consiste em um negócio jurídico processual por meio do qual os legitimados coletivos comumente o Ministério Público ou a Defensoria Pública estabelecem um acordo direto com a plataforma digital ofensora. Através desse mecanismo, a corporação compromete-se a adequar sua conduta às exigências legais e constitucionais, aplicando soluções tecnológicas imediatas sob pena de multa cominatória preestabelecida, o que evita a instauração ou a continuidade de um longo e desgastante processo judicial.

A agilidade desse instituto permite alcançar resultados práticos e desenhados sob medida para a dinamicidade das inovações digitais. Por meio da celebração de TACs, as plataformas de tecnologia aceitam voluntariamente implementar obrigações que dificilmente seriam obtidas com rapidez na via judicial tradicional, tais como: a criação de centrais exclusivas e aceleradas para receber notificações de autoridades brasileiras, o financiamento de campanhas massivas contra a desinformação e o discurso de ódio, e o ajuste imediato de seus termos de serviço em conformidade com o dever de cuidado estabelecido pelo STF. Ao optar pelo consenso extrajudicial, a plataforma resguarda sua imagem pública perante o mercado e seus anunciantes, enquanto a sociedade civil colhe os frutos de uma moderação de conteúdo aprimorada em tempo recorde.

6.7 O desafio da efetividade: astreintes e medidas atípicas de coerção

A despeito da existência de um arcabouço processual sofisticado, o desfecho da tutela coletiva esbarra constantemente no desafio da efetividade da execução. As grandes corporações de tecnologia detêm orçamentos globais que superam o Produto Interno Bruto de diversos países, o que mitiga o efeito intimidatório das tradicionais multas diárias (astreintes).

Para garantir que a responsabilidade civil digital não se transmute em uma promessa inócua, o Poder Judiciário tem sido compelido a aplicar sanções pecuniárias em patamares inéditos e a valer-se de medidas executivas atípicas, autorizadas pelo artigo 139, inciso IV, do Código de Processo Civil. As tutelas inibitórias e de remoção do ilícito direcionadas ao combate ao discurso de ódio apenas ganham concretude prática quando a sanção pelo descumprimento afeta severamente o patrimônio do ofensor. No caso das *Big Techs*, a jurisprudência caminha para a aplicação de astreintes calculadas com base no faturamento bruto anual da empresa no território nacional, garantindo a proporcionalidade necessária entre a capacidade econômica da corporação e a gravidade de sua omissão fiduciária.

Em situações de descumprimento sistêmico e afronta direta à soberania nacional como a recusa reiterada em desarticular redes coordenadas de intolerância ou em cumprir o dever de cuidado, surge a necessidade de adoção de meios coercitivos de ultima *ratio*. A jurisprudência pátria tem sustentado a legitimidade de

medidas extremas, tais como a suspensão temporária das atividades da plataforma no país ou a desmonetização forçada de suas contas corporativas. Embora polêmico por impactar reflexamente a comunidade de usuários, o bloqueio técnico temporário revela-se um mecanismo indispensável para compelir conglomerados transnacionais a se submeterem aos ditames da Constituição Federal. Portanto, a efetivação da responsabilidade civil digital passa, inexoravelmente, pela superação das ferramentas executivas tradicionais e pela afirmação de uma jurisdição coletiva rigorosa e adaptada à magnitude da economia digital.

7 CONSIDERAÇÕES FINAIS

A presente pesquisa investigou as profundas transformações ocorridas nas dinâmicas sociais e jurídicas em decorrência da ascensão da internet e do cenário de hiperconectividade contemporâneo. O cerne deste estudo consistiu em reavaliar os limites reais do exercício da liberdade de expressão no ciberespaço, especialmente quando esse direito fundamental é desvirtuado e instrumentalizado para promover o discurso de ódio, atentando frontalmente contra a dignidade da pessoa humana e os direitos essenciais de grupos minoritários. Demonstrou-se que a internet revolucionou a forma como a sociedade se comunica, democratizando a emissão de informações de maneira inédita, mas, de forma simultânea e prejudicial, amplificou o alcance, a perpetuidade e a velocidade de discursos lesivos. Diante dessa dicotomia, a análise do embate constitucional entre a livre manifestação do pensamento e a proteção aos direitos da personalidade exigiu um olhar técnico, revelando que não há direitos absolutos no ordenamento jurídico e que o paradoxo da tolerância impõe balizas rígidas àqueles que utilizam o espaço público virtual para violar direitos de terceiros.

Nesse contexto, dedicou-se parte expressiva deste estudo à evolução legislativa brasileira, com destaque central para o Marco Civil da Internet (Lei nº 12.965/2014). Verificou-se que a regra original encartada no art. 19, que condicionava a responsabilização civil das provedoras de aplicação à prévia inércia diante de uma ordem judicial específica, foi estruturada com a intenção legislativa de evitar a censura prévia e resguardar o debate democrático. Contudo, a pesquisa evidenciou que esse modelo demonstrou insuficiência frente à complexa arquitetura das redes sociais modernas. O ciberespaço atual deixou de atuar como um mero ambiente passivo de hospedagem de conteúdo para se converter em um ecossistema moldado por algoritmos de recomendação, desenhados para reter a atenção do usuário. Considerando que o conteúdo extremado frequentemente gera mais engajamento e, conseqüentemente, retorno financeiro com publicidade, as plataformas passaram a impulsionar ativamente materiais ilícitos. Nesse cenário, a exigência irrestrita de notificação judicial prévia converteu-se em uma proteção desproporcional para as corporações tecnológicas, prolongando o dano das vítimas.

O marco central dessa problemática consistiu na decisão do Supremo Tribunal Federal proferida em junho de 2025 e consolidada em 2026. A análise dos

votos dos ministros revelou a instauração de um novo regime de responsabilização, fundamentado na aplicação da teoria do risco do empreendimento e na necessidade de instituir um dever de cuidado por parte das redes sociais. O STF consolidou o entendimento de que, ao lucrar com a viralização algorítmica de conteúdos ilícitos, a empresa passa a integrar a cadeia de responsabilização civil. A Suprema Corte adotou parâmetros de imputação precisos: não declarou a inconstitucionalidade total nem revogou o art. 19, mas promoveu uma filtragem constitucional, adequando seus efeitos práticos. Estabeleceu-se a obrigação de moderação autônoma e imediata, sob pena de responsabilização civil, para casos de violações manifestamente ilícitas, como racismo, abuso infantil e atos antidemocráticos. Concomitantemente, preservou-se a exigência de tutela jurisdicional para situações controvertidas, assegurando a liberdade de expressão e mitigando o risco do "efeito inibitório" (*chilling effect*).

Apesar de reconhecer os avanços trazidos por essa nova hermenêutica protetiva, a pesquisa também apontou as contradições e os riscos de inconstitucionalidade por via transversa que o novo paradigma apresenta. Demonstrou-se que, ao delegar o dever de moderação às empresas de tecnologia, o Estado arrisca terceirizar parte de sua função jurisdicional, transferindo para entidades privadas o poder de delimitar a manifestação na esfera pública. Em face do risco de imposição de sanções pecuniárias, essas empresas tendem a adotar posturas de bloqueio profilático excessivo (*overblocking*), valendo-se de sistemas automatizados que, frequentemente, podem suprimir o dissenso político legítimo, silenciar denúncias sociais e prejudicar as próprias minorias que a lei visa proteger. A consequência prática desse cenário pode ser a exclusão do usuário do debate público de forma sumária e sem a devida transparência, o que tensiona princípios processuais constitucionais como o contraditório, a ampla defesa e o devido processo legal.

Conclui-se, portanto, que a efetividade da responsabilidade civil digital no Brasil dependerá da capacidade do ordenamento jurídico de avançar além da mera imposição de deveres de remoção de publicações. O desafio atual impõe a construção de um modelo de correção entre Estado, iniciativa privada e sociedade civil. Para que a delegação desse monitoramento não resulte em um ambiente de censura corporativa, é imperativo consolidar mecanismos de transparência algorítmica ativa. Instituições democráticas, o Ministério Público, a

Defensoria Pública, a academia e pesquisadores independentes devem dispor de meios para auditar o funcionamento das plataformas, mitigando o risco de manipulação de filtros de conteúdo. A proteção do ciberespaço contra o discurso de ódio deve coexistir com a preservação de sua natureza plural e democrática. Somente por meio da corresponsabilidade institucional e processual coletiva será possível assegurar que o avanço tecnológico atue em conformidade com a dignidade da pessoa humana.

REFERÊNCIAS

ABBOUD, Georges. **Direito Constitucional e a era digital**: limites da jurisdição na internet. 4. ed. São Paulo: Thomson Reuters Brasil, 2025.

ARENHART, Sérgio Cruz. **Processos estruturais no direito brasileiro**: reflexões a partir da teoria da jurisdição. São Paulo: Thomson Reuters Brasil, 2020.

BALKIN, Jack M. *Free speech is a triangle*. **Columbia Law Review**, Nova York, v. 118, n. 7, p. 2011-2056, nov. 2018. Disponível em: <https://columbialawreview.org/wp-content/uploads/2018/11/Balkin-FREE-SPEECH-IS-A-TRIANGLE.pdf>. Acesso em: 15 jun. 2026.

BESSA, Leonardo Roscoe. **Interesses difusos e coletivos**. 8. ed. Rio de Janeiro: Forense, 2023.

BIONI, Bruno Ricardo; NUNES, Vinícius. Transparência algorítmica e a moderação de conteúdo: os novos deveres de cuidado das plataformas digitais no Brasil. **Revista Brasileira de Direitos Fundamentais e Justiça**, Porto Alegre, v. 19, n. 48, p. 115-140, jan./jun. 2025.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 22 jun. 2026.

BRASIL. **Lei nº 7.347, de 24 de julho de 1985**. Disciplina a ação civil pública de responsabilidade por danos causados ao meio-ambiente, ao consumidor, a bens e direitos de valor artístico, estético, histórico, turístico e paisagístico (VETADO) e dá outras providências. Brasília, DF: Presidência da República. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l7347orig.htm. Acesso em: 22 jun. 2026.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil (Marco Civil da Internet). Brasília, DF: Presidência da República. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 22 jun. 2026.

BRASIL. Supremo Tribunal Federal. **Recurso Extraordinário nº 1.037.396 / São Paulo (Tema 987 da Repercussão Geral)**. Discussão sobre a constitucionalidade do art. 19 da Lei 12.965/2014. Relator: Ministro Dias Toffoli. Brasília, DF: Supremo Tribunal Federal, 2025. Disponível em: <https://portal.stf.jus.br/jurisprudenciaRepercussao/tema.asp?num=987>

BRASIL. Supremo Tribunal Federal. **Recurso Extraordinário nº 1.057.258 / Minas Gerais (Tema 533 da Repercussão Geral)**. Discussão sobre a responsabilidade civil de provedores por conteúdo ofensivo anterior ao Marco Civil. Relator: Ministro Luiz Fux. Brasília, DF: Supremo Tribunal Federal, 2025. Disponível em: <https://portal.stf.jus.br/jurisprudenciaRepercussao/tema.asp?num=533>. Acesso em: 22 jun. 2026.

CASTRO, Clarissa de. **A responsabilidade civil dos provedores de serviços de internet por atos de terceiros: uma análise à luz dos direitos fundamentais**. 2013. 142 f. Dissertação (Mestrado em Direito). Departamento de Direito, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013. Disponível em: <https://www.maxwell.vrac.puc-rio.br/22481/22481.PDF>. Acesso em: 16 jun. 2026.

CAVALIERI FILHO, Sergio. **Programa de responsabilidade civil**. 15. ed. São Paulo: Atlas, 2021.

DIDIER JR., Fredie. **Curso de direito processual civil: introdução ao direito processual civil, parte geral e processo de conhecimento**. 24. ed. Salvador: Juspodivm, 2022.

DONEDA, Danilo. **A proteção de dados e a responsabilidade algorítmica: do Marco Civil à inteligência artificial**. Rio de Janeiro: Renovar, 2024.

ESTADOS UNIDOS. **Communications Decency Act**. 47 U.S.C. § 230. Promulgada em 8 de fevereiro de 1996. Washington, DC: Congresso dos Estados Unidos, 1996.

ESTADOS UNIDOS. Corte de Apelações do Quarto Circuito. **Zeran v. America Online, Inc.**, 129 F.3d 327 (4th Cir. 1997). Decisão proferida em 12 de novembro de 1997.

FRAZÃO, Ana. Responsabilidade civil, algoritmos e plataformas digitais: a aplicação da teoria do risco do empreendimento na economia da atenção. **Revista de Direito Civil Contemporâneo**, São Paulo, v. 38, n. 2, p. 45-70, jan./mar. 2024.

KAYE, David. Speech Police: **The Global Struggle to Govern the Internet**. Nova York: Columbia Global Reports, 2019.

KLONICK, Kate. The new governors: The people, rules, and processes governing online speech. **Harvard Law Review**, Cambridge, v. 131, n. 6, p. 1598-1670, abr. 2018. Disponível em: <https://harvardlawreview.org/print/vol-131/the-new-governors-the-people-rules-and-processes-governing-online-speech/>. Acesso em: 15 jun. 2026.

KOSSEFF, Jeff. **The Twenty-Six Words That Created the Internet**. Ithaca: Cornell University Press, 2019.

LEONARDI, Marcel. **Responsabilidade Civil dos Provedores de Serviços de Internet**. 2. ed. São Paulo: Thomson Reuters Brasil, 2019.

MARINONI, Luiz Guilherme. **Tutela inibitória e tutela de remoção do ilícito**. 8. ed. São Paulo: Thomson Reuters Brasil, 2021.

MAZZILLI, Hugo Nigro. **A defesa dos interesses difusos em juízo: meio ambiente, consumidor, patrimônio cultural, patrimônio público e outros interesses**. 36. ed. São Paulo: Saraiva, 2024.

MENDES, Laura Schertel. **Algoritmos, inteligência artificial e o dever de cuidado nas plataformas digitais**: novos contornos para a responsabilidade civil. Revista de Direito Civil Contemporâneo, São Paulo, v. 35, p. 89-112, out./dez. 2023.

MILAGRES, Marcelo de Oliveira. **Liberdade de expressão, responsabilidade civil e discurso de ódio**. Migalhas, 13 maio 2021. Coluna Migalhas de Responsabilidade Civil. Disponível em: <https://www.migalhas.com.br/coluna/migalhasderesponsabilidade-civil/345380/liberdade-de-expressao-responsabilidade-civil-e-discurso-de-odio>. Acesso em: 8 maio 2024.

MORAES, Alexandre de. **Direito Constitucional e a defesa da Democracia no ambiente digital**. 39. ed. São Paulo: Atlas, 2024.

MULHOLLAND, Caitlin. **Responsabilidade civil e processos decisórios autônomos em sistemas de inteligência artificial (IA)**: autonomia, imputabilidade e responsabilidade. In: GOSTINSKI, Aline (Ed.). Estado, Regulação e Transformação Digital. São Paulo: Tirant lo Blanch, 2024. p. 54-75.

NITRINI, Rodrigo Vidal. **Liberdade de expressão nas redes sociais: o problema jurídico da remoção de conteúdo pelas plataformas**. São Paulo: Editora Dialética, 2021.

POLITIZE!. **Fundamentos da liberdade de expressão: Politize!**, [s.d.]. Disponível em: <https://www.politize.com.br/liberdade-de-expressao/>. Acesso em: 8 maio 2024.

POPPER, Karl. **A sociedade aberta e seus inimigos**. Tradução de Milton Amado. Belo Horizonte: Itatiaia, 1998.

SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais e a moderação de conteúdo no ciberespaço**. 15. ed. Porto Alegre: Livraria do Advogado, 2025.

SARMENTO, Daniel. A liberdade de expressão e o problema do hate speech. Revista de Direito do Estado, Rio de Janeiro, n. 4, p. 53-105, out./dez. 2006.

SCHREIBER, Anderson. **STF e a nova responsabilidade civil das plataformas digitais**. JOTA, São Paulo, 29 jul. 2025. Coluna do Anderson Schreiber. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/coluna-do-andersonschreiber/stf-e-a-nova-responsabilidade-civil-das-plataformas-digitais>. Acesso em: 15 jun. 2026.

SENADO NOTÍCIAS. **Marco Civil da Internet completa dez anos ante desafios sobre redes sociais e IA**. Brasília, DF: Senado Federal, 26 abr. 2024. Disponível em: <https://www12.senado.leg.br/noticias/materias/2024/04/26/marco-civil-da-internet-completa-dez-anos-ante-desafios-sobre-redes-sociais-e-ia>. Acesso em: 8 maio 2024.

SILVA, João Paulo. **Democracia e desinformação: o papel das Big Techs no direito brasileiro**. Rio de Janeiro: Lumen Juris, 2024.

SOUZA, Carlos Affonso; BOTTINO, Celina. **O Marco Civil da Internet e a encruzilhada da moderação de conteúdo**: reflexões sobre a responsabilidade de

intermediários. *Revista Brasileira de Direitos Fundamentais*, Belo Horizonte, v. 12, n. 2, p. 45-67, 2024.

TEPEDINO, Gustavo; SILVA, Rodrigo da Cunha e. **Economia da atenção, algoritmos e responsabilidade civil**: a revisão do artigo 19 do Marco Civil da Internet pelo STF. *Revista de Direito Civil Contemporâneo*, São Paulo, v. 40, p. 112-145, 2024.

UNIÃO EUROPEIA. Regulamento (EU) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022, relativo a um mercado único de serviços digitais e que altera a Diretiva 2000/31/CE (Regulamento dos Serviços Digitais - DSA). **Jornal Oficial da União Europeia**, Bruxelas, L 277, p. 1-102, 27 out. 2022. Disponível em: <https://eur-lex.europa.eu/legalcontent/PT/TXT/?uri=CELEX%3A32022R2065>. Acesso em: 22 jun. 2026.

VITORELLI, Edilson. **O devido processo legal coletivo**: dos direitos aos litígios coletivos. 3. ed. São Paulo: Thomson Reuters Brasil, 2022.

WATANABE, Kazuo (Coord.). **Tutela coletiva**: Ação Civil Pública e Ação Popular. 2. ed. São Paulo: Thomson Reuters Brasil, 2023.

ZANETI JR., Hermes. **Direito processual coletivo**: tutela dos direitos transindividuais. 4. ed. São Paulo: Saraiva, 2021.

ZUBOFF, Shoshana. **A era do capitalismo de vigilância**: a luta por um futuro humano na nova fronteira do poder. Tradução de George Schlesinger. Rio de Janeiro: Intrínseca, 2020.