REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

A RESPONSABILIDADE CIVIL DAS PLATAFORMAS DIGITAIS PELA NÃO REMOÇÃO DE CONTEÚDO ILÍCITO

CIVIL LIABILITY OF DIGITAL PLATFORMS FOR FAILURE TO REMOVE UNLAWFUL CONTENT

Ciências Sociais Aplicadas • 11/11/2026

REGISTRO DOI: 10.70773/revistatopicos/794295184

Francisco Campos Costa[1]

Paulo Renan Oliveira Santos[2]

Alex José Batista Cardoso Junior[3]

RESUMO

O presente artigo analisa a responsabilidade civil das plataformas digitais pela manutenção de conteúdos ilícitos, com ênfase no artigo 19 da Lei nº 12.965/2014 e nas alterações decorrentes do julgamento dos Temas 987 e 533 pelo Supremo Tribunal Federal. Investiga-se em que medida os provedores de aplicações podem ser responsabilizados pela não remoção de conteúdos produzidos por terceiros diante da necessidade de conciliar proteção dos direitos fundamentais, liberdade de expressão e prevenção de danos no ambiente digital. A pesquisa adota abordagem qualitativa, de natureza exploratória, desenvolvida mediante procedimentos bibliográficos e documentais, com análise da legislação brasileira, de decisões do Supremo Tribunal Federal e do Superior Tribunal de Justiça, de literatura especializada e, complementarmente, de modelos regulatórios estrangeiros. O estudo demonstra que o regime originalmente centrado na exigência de ordem judicial específica passou a apresentar estrutura diferenciada, na qual fatores como natureza do conteúdo, ciência do provedor, diligência, impulsionamento remunerado, disseminação artificial e falhas sistêmicas podem assumir relevância para a responsabilização. Verificou-se também que inteligência artificial e mecanismos automatizados ampliam a capacidade de moderação, embora apresentem limitações relacionadas ao contexto, à transparência e à possibilidade de decisões equivocadas. Conclui-se que a jurisprudência brasileira evoluiu para um modelo mais proativo e graduado, que mantém o artigo 19 em hipóteses específicas e procura equilibrar proteção contra conteúdos ilícitos e preservação da liberdade de expressão.

Palavras-chave: Responsabilidade civil; Plataformas digitais; Marco Civil da Internet; Moderação de conteúdo; Liberdade de expressão.

ABSTRACT

This article analyzes the civil liability of digital platforms for maintaining unlawful content, with emphasis on Article 19 of Brazilian Law No. 12,965/2014 and on the changes resulting from the Brazilian Federal Supreme Court's rulings on General Repercussion Themes 987 and 533. It investigates the extent to which application providers may be held liable for failing to remove third-party content while considering the need to reconcile the protection of fundamental rights, freedom of expression, and prevention of harm in the digital environment. The research adopts a qualitative and exploratory approach based on bibliographic and documentary procedures, including the analysis of Brazilian legislation, decisions of the Federal Supreme Court and the Superior Court of Justice, specialized literature and, complementarily, selected foreign regulatory models. The study demonstrates that the regime originally centered on the requirement of a specific judicial order has evolved into a differentiated structure in which factors such as the nature of the content, the provider's knowledge, diligence, paid amplification, artificial dissemination, and systemic failures may become relevant to liability. It was also found that artificial intelligence and automated mechanisms increase content moderation capacity, although they present limitations concerning context, transparency, and the possibility of erroneous decisions. The study concludes that Brazilian case law has evolved toward a more proactive and graduated model that retains Article 19 in specific cases and seeks to balance protection against unlawful content and the preservation of freedom of expression.

Keywords: Civil liability; Digital platforms; Brazilian Civil Rights Framework for the Internet; Content moderation; Freedom of expression.

1. INTRODUÇÃO

As plataformas digitais assumiram posição central nas relações comunicacionais contemporâneas, não apenas porque permitem criação e compartilhamento de informações, mas também porque organizam sua circulação por meio de mecanismos de recomendação, classificação, impulsionamento e moderação.

Essa atuação amplia a influência de empresas privadas sobre os fluxos informacionais e torna mais complexa a discussão acerca da responsabilidade por conteúdos produzidos por terceiros, especialmente quando publicações ofensivas, desinformação, práticas discriminatórias ou outras violações alcançam grande número de usuários em curto intervalo de tempo (Klonick, 2018; Santos et al., 2023; Lemos; Estima, 2026).

A Constituição Federal protege a manifestação do pensamento e a liberdade de comunicação, ao mesmo tempo em que assegura a intimidade, a vida privada, a honra e a imagem, com direito à reparação quando esses bens forem violados. A retirada de conteúdo, portanto, não pode ser analisada apenas como instrumento de combate a ilícitos, pois também interfere nas condições de exercício da liberdade de expressão (Brasil, 1988).

Nesse contexto, o tema da pesquisa corresponde à responsabilidade civil das plataformas digitais pela não remoção de conteúdo ilícito, delimitando-se ao regime brasileiro aplicável aos provedores de aplicações de internet, com ênfase no artigo 19 da Lei nº 12.965/2014 e na interpretação construída pelo Supremo Tribunal Federal nos Temas 987 e 533 da repercussão geral. Tal recorte, portanto, considera ainda a influência dos mecanismos de moderação, dos

algoritmos e da inteligência artificial e compreende o período iniciado com o Marco Civil da Internet, em 2014, alcançando os desdobramentos jurisprudenciais verificados até 2026.

Em sua formulação original, o art. 19 condicionou a responsabilidade civil do provedor por conteúdo gerado por terceiro ao descumprimento de ordem judicial específica de indisponibilização, tendo em vista que o modelo buscava preservar a liberdade de expressão e reduzir incentivos à censura privada, diferenciando a responsabilidade do usuário daquela atribuível ao intermediário. Entretanto, a expansão das plataformas e de seus sistemas de recomendação tornou essa estrutura progressivamente controvertida, sobretudo diante da velocidade de circulação dos danos e da participação tecnológica das empresas na ampliação de determinados conteúdos (Brasil, 2014; Flumignan, 2018; Lemos; Estima, 2026).

A controvérsia culminou no julgamento dos Recursos Extraordinários nº 1.037.396 e nº 1.057.258, quando o Supremo Tribunal Federal reconheceu, em 2025, a inconstitucionalidade parcial e progressiva do artigo 19 e estabeleceu regimes diferenciados de responsabilização. A decisão não eliminou integralmente o dispositivo nem instituiu responsabilidade automática das plataformas, passando a considerar fatores como natureza do ilícito, forma de circulação, ciência do provedor, impulsionamento e existência de falhas sistêmicas, aspectos posteriormente aperfeiçoados nos desdobramentos do julgamento em 2026 (Supremo Tribunal Federal, 2025a; Supremo Tribunal Federal, 2026).

Diante desse cenário, formula-se o seguinte problema de pesquisa: em que medida as plataformas digitais podem ser responsabilizadas civilmente pela manutenção de conteúdos ilícitos, à luz do artigo 19 do Marco Civil da Internet e da interpretação conferida pelo Supremo Tribunal Federal? A hipótese nula, H0, considera que a responsabilidade permanece predominantemente condicionada à ordem judicial específica.

A hipótese alternativa, H1, sustenta que ocorreu evolução para um regime mais proativo, no qual determinadas situações podem gerar deveres de prevenção, moderação ou remoção independentemente de decisão judicial prévia.

A pesquisa justifica-se pela necessidade de compreender os efeitos dessa transformação sobre a proteção das vítimas e sobre a liberdade de expressão, dado que um regime excessivamente restritivo pode dificultar a interrupção de danos que se propagam rapidamente, enquanto uma responsabilização demasiadamente ampla pode incentivar remoções preventivas e ampliar o poder privado de controle sobre manifestações legítimas. A definição dos deveres das plataformas exige, assim, critérios capazes de conciliar proteção de direitos, segurança jurídica e proporcionalidade (Lefouili; Madio, 2022; Lemos; Estima, 2026).

O objetivo geral, então, consiste em analisar a responsabilidade civil das plataformas digitais pela manutenção de conteúdos ilícitos à luz da interpretação conferida pelo Supremo Tribunal Federal ao artigo 19 do Marco Civil da Internet.

Já como objetivos específicos, busca-se identificar as principais categorias de conteúdos ilícitos relevantes para a responsabilização,

examinar as implicações da decisão do STF sobre os deveres de moderação e remoção e avaliar de que maneira a inteligência artificial, os algoritmos de recomendação e os mecanismos automatizados interferem na retirada de conteúdos e na proteção de direitos fundamentais.

Para alcançar esses objetivos, utiliza-se abordagem qualitativa, de natureza exploratória, desenvolvida mediante pesquisa bibliográfica e documental. São examinados a Constituição Federal, o Marco Civil da Internet, decisões do STF e do STJ, literatura especializada e, de forma comparativa, modelos regulatórios estrangeiros selecionados.

O artigo organiza-se em cinco seções. Após esta introdução, a fundamentação teórica (item 2) aborda os fundamentos da responsabilidade civil das plataformas, o artigo 19 do Marco Civil, suas críticas e a evolução do regime. O item 3 apresenta a metodologia. O item 4 expõe e discute os resultados: a decisão do STF, a liberdade de expressão, a moderação de conteúdo, a inteligência artificial e a experiência comparada. O item 5 traz a conclusão, que retoma o problema de pesquisa e as hipóteses formuladas.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Fundamentos da Responsabilidade Civil das Plataformas Digitais

A responsabilidade civil no ambiente digital apresenta particularidades decorrentes da multiplicidade de agentes envolvidos na circulação das informações. Nas plataformas, o responsável pela publicação não é necessariamente o único cuja conduta interessa ao Direito, pois manutenção, organização e

difusão podem envolver decisões técnicas e empresariais posteriores.

Por essa razão, torna-se necessário distinguir a responsabilidade daquele que produz o conteúdo da eventual responsabilidade do intermediário decorrente de sua própria atuação ou omissão (Flumignan, 2024).

Essa distinção é relevante porque a expressão provedor de internet compreende atividades diferentes. O artigo 18 do Marco Civil determina que o provedor de conexão não responde civilmente pelos danos decorrentes de conteúdo produzido por terceiros, enquanto os provedores de aplicações se submetem ao regime específico dos artigos seguintes (Brasil, 2014).

A diferença relaciona-se à posição ocupada por cada agente e à possibilidade de interferência sobre o material disponibilizado, já que quem apenas proporciona acesso à rede não dispõe dos mesmos mecanismos existentes em serviços que hospedam, organizam ou distribuem conteúdos (Brasil, 2014; Flumignan, 2024).

A posição dos provedores de aplicações tornou-se mais complexa com a evolução das redes sociais. Muitas plataformas organizam *feeds*, recomendam publicações, comercializam impulsionamento e utilizam sistemas automatizados para aumentar ou reduzir visibilidade.

Isso não transforma a empresa na autora da manifestação inserida por terceiro, mas impede que sua atuação seja compreendida exclusivamente como armazenamento passivo. Lemos e Estima (2026) destacam que a materialidade tecnológica e a participação algorítmica influenciam o alcance das informações e, por

consequência, precisam integrar a discussão sobre responsabilidade.

Antes da Lei nº 12.965/2014, a ausência de disciplina específica fazia com que esses conflitos fossem solucionados mediante normas gerais e construções jurisprudenciais. Nesse período, consolidou-se entendimento de que a disponibilização do serviço não tornava o intermediário automaticamente responsável pelo conteúdo ofensivo de terceiro, nem lhe impunha fiscalização prévia de todas as manifestações (Lima, 2015; Flumignan, 2018).

A responsabilidade poderia surgir, contudo, quando o provedor recebesse comunicação suficientemente determinada acerca do material e permanecesse inerte apesar da possibilidade de retirada (Lima, 2015; Flumignan, 2018).

A ciência extrajudicial assumia, portanto, função importante na identificação da omissão do provedor. O fundamento da responsabilidade não estava na autoria da publicação, mas no comportamento posterior do intermediário depois de informado sobre sua existência, e tal distinção continuou relevante após o Marco Civil, embora a legislação tenha alterado o acontecimento necessário para caracterizar a omissão juridicamente relevante, substituindo, como regra, a notificação privada por uma ordem judicial específica (Brasil, 2014; Lima, 2015).

A evolução desse regime demonstra que a responsabilidade das plataformas nunca se resumiu à escolha entre responsabilização integral e completa imunidade, e o ponto central está na identificação dos deveres que podem ser razoavelmente exigidos

segundo a função desempenhada pelo provedor, sua relação com o conteúdo e as circunstâncias do dano.

Essa perspectiva tornou-se ainda mais relevante diante de serviços que participam ativamente da organização dos fluxos informacionais, sem que isso elimine a responsabilidade primária daquele que produz a manifestação ilícita (Flumignan, 2024; Lemos; Estima, 2026).

2.2. O Artigo 19 da Lei Nº 12.965/2014 e o Modelo de Responsabilização dos Provedores

O Marco Civil da Internet procurou conferir maior previsibilidade às relações jurídicas desenvolvidas na rede e estabeleceu disciplina específica para conteúdos gerados por terceiros, tendo em conta que o art. 19 determinou que o provedor de aplicações somente poderia ser responsabilizado civilmente quando, após ordem judicial específica, deixasse de tornar indisponível o conteúdo indicado dentro dos limites técnicos de seu serviço. A existência da publicação, por si só, não produzia responsabilidade automática da plataforma (Brasil, 2014).

A opção pela ordem judicial estava diretamente relacionada à liberdade de expressão e à preocupação com censura privada, pois o modelo procurava evitar que qualquer reclamação conduzisse à retirada preventiva de conteúdo cuja ilicitude fosse discutível, atribuindo inicialmente ao Poder Judiciário a função de estabelecer o dever de indisponibilização. Em relação à jurisprudência anterior, a mudança reduziu os efeitos da notificação extrajudicial e consolidou um sistema no qual a omissão juridicamente relevante dependia,

como regra, do descumprimento de determinação judicial (Lima, 2015; Flumignan, 2018).

A referida exigência buscava impedir determinações excessivamente genéricas e evitar que a plataforma recebesse a tarefa de escolher autonomamente quais manifestações deveriam ser retiradas. Esse cuidado era especialmente importante em controvérsias envolvendo crítica, opinião, reputação ou interesse público, nas quais publicações sobre o mesmo assunto poderiam possuir significados e enquadramentos jurídicos distintos (Brasil, 2014).

O próprio Marco Civil, entretanto, reconheceu que algumas violações justificavam tratamento diferente, por conta do art. 21, que estabeleceu disciplina específica para divulgação não autorizada de imagens ou materiais contendo nudez ou atos sexuais de caráter privado, atribuindo relevância à notificação da pessoa atingida independentemente de ordem judicial prévia. A exceção demonstrava que, desde sua origem, a legislação não tratava todas as formas de conteúdo ilícito de maneira absolutamente uniforme (Brasil, 2014; Flumignan, 2018).

A exigência de ordem judicial para caracterização da responsabilidade também não impedia a plataforma de moderar espontaneamente seu serviço. No REsp nº 2.139.749/SP, o Superior Tribunal de Justiça reconheceu que provedores podem remover conteúdos incompatíveis com a legislação ou com suas próprias regras, desde que não atuem abusivamente (Superior Tribunal de Justiça, 2024).

Dessa forma, a faculdade empresarial de moderar deve ser distinguida do momento em que a permanência de determinada publicação passa a gerar responsabilidade civil perante terceiro.

Essa diferença é relevante porque as plataformas desenvolveram políticas próprias de conteúdo, procedimentos de denúncia e formas privadas de sanção, ao passo que Santos et al. (2023) demonstram que serviços digitais utilizam regras internas para tratar, por exemplo, manifestações enquadradas como discurso de ódio. A governança do conteúdo ocorre, assim, pela interação entre legislação, decisões judiciais, termos de uso e mecanismos empresariais, o que torna inadequada a percepção de que toda decisão de moderação dependeria de prévia atuação estatal.

O modelo do artigo 19, como visto, procurava equilibrar riscos opostos. Proteção excessiva ao intermediário poderia dificultar a interrupção de danos, enquanto uma responsabilização facilitada poderia estimular remoção preventiva de manifestações legítimas (Lefouili; Madio, 2022).

Esse segundo risco não é apenas teórico, pois regimes de responsabilidade alteram incentivos econômicos, investimentos e estratégias de moderação das empresas. A calibragem jurídica, portanto, produz efeitos que ultrapassam o conflito individual e alcançam o próprio funcionamento do espaço comunicacional digital (Lefouili; Madio, 2022).

2.3. Críticas e Limitações da Aplicação do Artigo 19 do Marco Civil da Internet

Uma das principais críticas ao artigo 19 relacionou-se à necessidade de judicialização como pressuposto geral da responsabilidade do

provedor. Embora a intervenção judicial aumentasse a segurança diante de conteúdos controvertidos, ela também exigia que a pessoa atingida buscasse decisão específica mesmo quando a plataforma já tivesse sido informada da publicação.

Em um ambiente caracterizado por rápida reprodução, esse intervalo poderia prolongar a exposição ao dano e reduzir a eficácia prática da posterior retirada (Lima, 2015; Flumignan, 2018). A dificuldade decorria principalmente da diferença entre o tempo da tutela judicial e a velocidade da comunicação digital. Uma publicação pode ser compartilhada, reproduzida ou armazenada por diversos usuários antes da obtenção de uma decisão.

Isso evidencia a limitação de aplicar uma mesma exigência a situações muito distintas em natureza e gravidade (Flumignan, 2018; Supremo Tribunal Federal, 2025a). A transformação tecnológica também colocou em questão a ideia de intermediário essencialmente passivo, pois, em serviços organizados por recomendação, a visibilidade das informações não resulta apenas da escolha do autor ou de compartilhamentos espontâneos.

Algoritmos selecionam aquilo que será apresentado e podem interferir significativamente no alcance das publicações, fazendo com que a arquitetura da plataforma se torne elemento relevante para compreender sua participação na circulação do conteúdo (Lemos; Estima, 2026).

Essa diferença aparece de forma ainda mais clara no impulsionamento remunerado e na disseminação artificial. Quando a empresa recebe pagamento para ampliar uma publicação, sua

relação com a circulação deixa de ser equivalente à simples hospedagem.

Da mesma forma, mecanismos destinados a produzir alcance inorgânico acrescentam uma dimensão tecnológica à propagação do material. Esses fatores foram posteriormente considerados pelo STF ao construir tratamentos diferenciados para determinadas formas de amplificação (Lemos; Estima, 2026; Supremo Tribunal Federal, 2025a).

Outra crítica envolveu a tutela dos direitos da personalidade. Lima (2015) apontou que a exigência de intervenção judicial modificou o regime anterior, no qual a ciência extrajudicial possuía relevância maior, e suscitou debate sobre a adequação da proteção conferida aos intermediários diante de danos à honra, à imagem e à privacidade. O problema, contudo, não seria solucionado pela simples adoção de responsabilidade ampla, pois isso poderia transferir à empresa forte incentivo para remover manifestações sempre que existisse algum risco de condenação.

A literatura econômica reforça essa preocupação ao demonstrar que regimes mais rigorosos podem produzir sobre-remoção, dado que, quando o custo potencial da permanência é elevado, excluir o conteúdo pode se tornar a alternativa empresarial mais segura mesmo diante de dúvida sobre sua ilicitude. O desafio regulatório consiste, portanto, em incentivar resposta adequada a danos reais sem criar uma estrutura na qual a prevenção de responsabilidade privada reduza indevidamente o espaço de manifestações legítimas (Lefouili; Madio, 2022).

Por essa razão, a controvérsia não pode ser reduzida a uma escolha entre proteção da vítima e liberdade de expressão. A própria revisão constitucional do artigo 19 envolveu a busca por critérios capazes de diferenciar situações (Supremo Tribunal Federal, 2025a; Lemos; Estima, 2026).

A necessidade de respostas mais intensas diante de determinadas práticas não implica que toda notificação deva provocar retirada automática, especialmente quando a qualificação do conteúdo como ilícito exige interpretação jurídica mais cuidadosa. O debate também se deslocou para danos de natureza sistêmica. Redes de disseminação de práticas criminosas e conteúdos capazes de atingir coletivamente grupos ou instituições desafiam uma análise limitada à publicação individual.

Nessas situações, a qualidade das estruturas de prevenção e resposta da plataforma passa a ser juridicamente relevante, aspecto incorporado pelo STF ao desenvolver a noção de falha sistêmica e o dever de cuidado aplicável a categorias específicas de conteúdos graves (Supremo Tribunal Federal, 2025a; Supremo Tribunal Federal, 2026).

2.4. Evolução da Responsabilidade Civil no Ambiente Digital

A evolução brasileira pode ser compreendida a partir de três momentos principais. Antes do Marco Civil, a jurisprudência atribuiu relevância à notificação extrajudicial. A partir de 2014, a ordem judicial específica tornou-se, como regra, pressuposto da responsabilidade pela não retirada. Em 2025, o STF promoveu nova reformulação ao estabelecer tratamentos diferenciados conforme a natureza do ilícito e a atuação do provedor. O percurso revela

sucessivas tentativas de conciliar proteção das vítimas, liberdade de expressão e segurança jurídica (Superior Tribunal de Justiça, 2025; Supremo Tribunal Federal, 2025a).

No primeiro período, não se exigia que o provedor fiscalizasse previamente todas as manifestações, nem se considerava que o simples uso de sua estrutura por terceiro produzisse responsabilidade objetiva. Buscava-se identificar uma omissão própria depois que o intermediário adquirisse ciência suficientemente determinada do material. Com o Marco Civil, esse ponto de referência foi deslocado para a ordem judicial, fazendo com que o descumprimento da determinação específica assumisse papel central na caracterização da responsabilidade subjetiva por omissão (Flumignan, 2018; Superior Tribunal de Justiça, 2023).

Paralelamente, consolidou-se o reconhecimento de que o artigo 19 não retirava das plataformas a possibilidade de estabelecer regras próprias e moderar seu ambiente. Responsabilidade civil e moderação privada passaram, assim, a ser compreendidas como questões relacionadas, mas juridicamente distintas. Essa separação seria particularmente importante para a fase posterior, quando a atuação empresarial sobre a distribuição e a governança do conteúdo passou a receber maior atenção na definição dos deveres dos intermediários (Superior Tribunal de Justiça, 2024).

O REsp nº 1.783.309/SP, julgado em 2025, exemplifica essa evolução. Ao examinar ofensas à honra, o STJ reconstruiu a diferença entre os regimes anterior e posterior ao Marco Civil e já incorporou a orientação do STF. O Tribunal reconheceu a preservação do artigo 19 nessa categoria e admitiu que reproduções idênticas de conteúdo anteriormente declarado ilícito fossem retiradas após notificação

capaz de identificá-las, sem necessidade de uma nova decisão judicial para cada veiculação (Superior Tribunal de Justiça, 2025).

A solução procura adaptar a tutela jurídica à capacidade de replicação própria da internet, tendo-se que exigir processo autônomo para cada nova reprodução poderia tornar praticamente ineficaz decisão que já reconheceu determinado conteúdo como ilícito. Ao mesmo tempo, a exigência de identidade com o material anteriormente analisado busca impedir que a ordem inicial seja utilizada como autorização genérica para excluir manifestações apenas semelhantes ou relacionadas ao mesmo tema (Superior Tribunal de Justiça, 2025).

A evolução tecnológica também modificou as possibilidades de intervenção dos provedores, embora não tenha eliminado as limitações da moderação. Ferramentas automatizadas permitem analisar grandes quantidades de conteúdo, mas enfrentam dificuldades na compreensão de contexto, ambiguidade e significados sociais. A maior capacidade técnica, portanto, amplia possibilidades de prevenção e resposta sem justificar a expectativa de identificação perfeita de toda ilicitude antes de sua circulação (Gorwa; Binns; Katzenbach, 2020).

O percurso brasileiro demonstra, assim, que a responsabilidade não evoluiu em linha reta para um sistema necessariamente mais rigoroso. Cada regime redistribuiu riscos e competências entre usuários, plataformas e Poder Judiciário.

A alteração mais recente deslocou parte da análise de uma pergunta predominantemente formal, centrada na existência de ordem judicial, para uma avaliação mais ampla da natureza do conteúdo,

da ciência do provedor, da forma de circulação e da qualidade de sua atuação diante de determinados riscos (Flumignan, 2024; Supremo Tribunal Federal, 2025a).

3. METODOLOGIA

A pesquisa adota abordagem qualitativa, de natureza exploratória, desenvolvida mediante procedimentos bibliográficos e documentais. Essa opção decorre das características do objeto, que exige interpretação de normas, decisões judiciais e produção doutrinária relacionadas à responsabilidade civil das plataformas digitais, ao artigo 19 do Marco Civil da Internet, à liberdade de expressão, à moderação de conteúdos e aos efeitos das tecnologias utilizadas na circulação das informações.

A pesquisa documental compreende a Constituição Federal de 1988, a Lei nº 12.965/2014, documentos e decisões relacionados aos Temas 987 e 533 da repercussão geral e precedentes do Superior Tribunal de Justiça utilizados para reconstruir a evolução da responsabilidade dos provedores. Foram considerados ainda os desdobramentos do julgamento do STF ocorridos até 2026, permitindo confrontar o regime original do artigo 19 com a interpretação constitucional posteriormente consolidada.

A etapa bibliográfica utilizou livros, dissertações, artigos científicos e publicações especializadas diretamente relacionadas ao problema e aos objetivos da pesquisa. Foram considerados estudos sobre responsabilidade dos provedores, governança das plataformas, direitos fundamentais, moderação e sistemas automatizados, combinando trabalhos anteriores necessários à compreensão

histórica com produções recentes sobre as transformações jurídicas e tecnológicas do ambiente digital.

O recorte temporal principal inicia-se com a promulgação do Marco Civil da Internet, em 2014, e alcança os desdobramentos disponíveis em 2026. Fontes anteriores foram empregadas quando necessárias para reconstruir o regime jurisprudencial que antecedeu a lei. A análise foi organizada segundo os eixos correspondentes aos objetivos, passando pelos fundamentos da responsabilidade, revisão do artigo 19, conteúdos ilícitos, liberdade de expressão, moderação, inteligência artificial e experiência comparada.

A perspectiva comparativa utilizou, de forma auxiliar, o Digital Services Act da União Europeia, o Online Safety Act do Reino Unido, a Section 230 norte-americana e estudos sobre experiências estrangeiras. A finalidade não foi equiparar ordenamentos distintos, mas identificar diferentes formas de distribuição de deveres, avaliação de riscos e proteção da liberdade de expressão capazes de contribuir para a compreensão crítica do modelo brasileiro.

A interpretação confrontou legislação, jurisprudência e produção bibliográfica, buscando identificar convergências, limitações e transformações nos deveres atribuídos às plataformas. Por se tratar de pesquisa bibliográfica e documental, não foram empregados questionários, entrevistas ou outros instrumentos empíricos. Os resultados decorrem da análise crítica das fontes e foram discutidos no item 4 em articulação com o problema, os objetivos e as hipóteses formuladas.

4. RESULTADOS E DISCUSSÕES

Esta seção apresenta os resultados da análise documental e bibliográfica e os discute à luz da literatura. Inicia-se pela redefinição do regime de responsabilidade pelo Supremo Tribunal Federal (4.1) e prossegue com a liberdade de expressão (4.2), a moderação de conteúdo (4.3), a inteligência artificial (4.4) e a perspectiva comparada (4.5), em articulação com o problema, os objetivos e as hipóteses da pesquisa.

4.1. A Redefinição da Responsabilidade das Plataformas Digitais Após o Julgamento do Supremo Tribunal Federal

A mudança mais significativa ocorreu no julgamento conjunto dos Recursos Extraordinários nº 1.037.396 e nº 1.057.258, correspondentes aos Temas 987 e 533 da repercussão geral. Em junho de 2025, o Supremo Tribunal Federal reconheceu a inconstitucionalidade parcial e progressiva do artigo 19, entendendo que a utilização uniforme da ordem judicial como requisito da responsabilidade não oferecia proteção suficiente em determinadas hipóteses, sem, contudo, eliminar integralmente o regime previsto pelo Marco Civil (Supremo Tribunal Federal, 2025a, 2025b).

A expressão inconstitucionalidade parcial e progressiva é essencial à compreensão do julgamento. O STF preservou o artigo 19 para determinadas situações e instituiu outros critérios enquanto não houver nova disciplina legislativa, construindo um regime diferenciado conforme o conteúdo, as características do serviço e o comportamento adotado pelo provedor. A decisão, portanto, não autoriza concluir que qualquer notificação extrajudicial gera automaticamente responsabilidade civil (Supremo Tribunal Federal, 2025a).

Nas situações alcançadas pela nova disciplina, a ciência do provedor e a diligência de sua resposta adquiriram maior relevância, cuja formulação aperfeiçoada em 2026 admite responsabilização em hipóteses que não dependem necessariamente da ordem judicial anteriormente exigida, ao mesmo tempo em que permite considerar a existência de dúvida razoável acerca da ilicitude após análise qualificada e diligente. O modelo procura evitar tanto a completa inércia quanto a obrigação de retirar automaticamente toda manifestação denunciada (Supremo Tribunal Federal, 2026).

Para crimes e atos ilícitos civis contra a honra, entretanto, permaneceu relevante o regime do artigo 19, visto que a plataforma pode retirar espontaneamente uma publicação diante de suas regras internas ou de sua própria avaliação, mas a possibilidade de moderação não se confunde com os pressupostos da responsabilidade indenizatória. A preservação de tratamento diferenciado para essa categoria demonstra que a Corte reconheceu a dificuldade contextual frequentemente envolvida na avaliação de manifestações que atingem reputação e honra (Supremo Tribunal Federal, 2026).

O STF também disciplinou a disseminação de conteúdos previamente reconhecidos como ilícitos. Quando já existe decisão judicial sobre determinada manifestação, novas veiculações idênticas podem ser removidas mediante notificação capaz de identificá-las, sem repetição da mesma discussão para cada publicação. A solução busca impedir que a multiplicação de endereços eletrônicos esvazie a tutela anteriormente concedida e foi posteriormente incorporada pelo STJ na análise de casos concretos (Supremo Tribunal Federal, 2026; Superior Tribunal de Justiça, 2025).

Outro elemento importante é a diferenciação de serviços destinados predominantemente a comunicações privadas e daqueles que interferem no fluxo informacional. Ao preservar tratamento específico para certas comunicações interpessoais e provedores de menor ingerência sobre a circulação, o Tribunal incorporou características da arquitetura tecnológica à definição dos deveres jurídicos. O nível de participação da plataforma na distribuição das informações passa, desse modo, a integrar a análise da responsabilidade (Supremo Tribunal Federal, 2026).

Nos casos de publicidade ou impulsionamento remunerado, a Corte adotou disciplina mais rigorosa, prevendo presunção relativa de culpa e possibilidade de demonstração de atuação diligente pelo provedor.

Tratamento semelhante alcança mecanismos artificiais de disseminação inorgânica de conteúdo ilícito. O fundamento dessa diferenciação está na maior participação econômica ou tecnológica da plataforma na ampliação da publicação, situação distinta daquela em que o serviço apenas recebe conteúdo inserido espontaneamente pelo usuário (Supremo Tribunal Federal, 2026; Lemos; Estima, 2026).

O julgamento estabeleceu ainda dever de cuidado associado à circulação massiva de determinadas categorias de conteúdos ilícitos graves. Nesses casos, a responsabilidade foi relacionada à existência de falha sistêmica na adoção de medidas adequadas de prevenção ou remoção. Uma ocorrência isolada não é suficiente para caracterizar deficiência estrutural, pois a análise deve considerar o funcionamento das medidas implantadas pela plataforma diante de

padrões relevantes de circulação dos ilícitos especificados pela Corte (Supremo Tribunal Federal, 2026).

A noção de falha sistêmica amplia o foco tradicional da responsabilidade civil. Em lugar de observar apenas a decisão tomada diante de determinada postagem, passa-se a considerar também políticas internas, recursos técnicos, canais de denúncia, procedimentos de análise e mecanismos de revisão. Douek (2022) propõe perspectiva semelhante ao sustentar que a moderação em grandes plataformas precisa ser compreendida como sistema, visto que seus resultados decorrem de escolhas estruturais e não apenas de decisões individualizadas.

Por essa razão, a reformulação jurisprudencial alcançou deveres de organização e transparência. O STF passou a exigir estruturas destinadas ao recebimento e tratamento de notificações, canais de atendimento e mecanismos adequados de resposta, além de medidas relacionadas ao dever de cuidado. Os aperfeiçoamentos realizados em 2026 reforçaram que a responsabilidade contemporânea das plataformas não se limita à reação posterior a uma ordem judicial, podendo envolver também a qualidade dos procedimentos internos adotados para enfrentar determinados riscos (Supremo Tribunal Federal, 2026).

Essas obrigações, todavia, não instituem monitoramento absoluto. O sistema permanece baseado na diferenciação das hipóteses e procura evitar que a imposição de maior diligência converta as plataformas em garantidoras universais de todo conteúdo publicado. A preservação de critérios como dúvida razoável, tipo de serviço, gravidade do ilícito e forma de disseminação demonstra que a ampliação dos deveres deve permanecer compatível com a

liberdade de expressão e com as limitações técnicas da atividade (Supremo Tribunal Federal, 2026; Lefouili; Madio, 2022).

O resultado é um modelo graduado de responsabilidade. Serviços que exercem pouca interferência sobre o fluxo comunicacional recebem tratamento distinto daqueles que recomendam, impulsionam ou participam de mecanismos de amplificação, enquanto falhas estruturais relacionadas a conteúdos graves recebem disciplina própria. O julgamento deslocou, portanto, a responsabilidade de um modelo predominantemente uniforme para outro no qual a intensidade dos deveres se aproxima do grau de participação da plataforma sobre a circulação das informações (Lemos; Estima, 2026; Supremo Tribunal Federal, 2026).

A redefinição do artigo 19 não solucionou o problema mediante substituição de uma regra simples por outra igualmente abrangente. A responsabilidade pela não remoção passou a depender de fatores diversos, entre eles natureza do conteúdo, ciência do provedor, diligência, impulsionamento e adequação das medidas estruturais.

Essa diferenciação procura adaptar o regime às transformações ocorridas desde 2014 sem abandonar completamente a preocupação com censura privada e segurança jurídica que orientou a formulação original do Marco Civil (Lemos; Estima, 2026; Supremo Tribunal Federal, 2026).

4.2. Liberdade de Expressão, Desinformação e Discurso de Ódio

A responsabilidade pela permanência de conteúdo ilícito deve ser examinada em conjunto com a liberdade de expressão, uma vez que

decisões de retirada interferem diretamente na circulação de ideias, opiniões e informações.

A Constituição protege a manifestação do pensamento e a comunicação ao mesmo tempo em que tutela honra, imagem, intimidade e vida privada. Por isso, a resposta jurídica precisa considerar simultaneamente o direito daquele que se expressa e os direitos da pessoa eventualmente atingida pela manifestação (Brasil, 1988).

Essa tensão assume características próprias no ambiente digital porque conteúdos podem ultrapassar rapidamente o círculo inicial de destinatários. Compartilhamentos e sistemas de recomendação favorecem reprodução em larga escala e permanência da publicação, ampliando os efeitos de eventual dano. Assim, embora a autoria continue sendo relevante, as características técnicas de distribuição também influenciam a extensão alcançada pela manifestação e a possibilidade de sua contenção (Lemos; Estima, 2026).

A questão ganha maior importância diante de desinformação, discurso de ódio, fraudes e violações de direitos da personalidade. Gillespie (2018) demonstra que as plataformas não apenas hospedam, mas selecionam, organizam e promovem conteúdos segundo critérios próprios, o que contribui para compreender por que determinadas publicações podem obter elevada visibilidade. Ainda assim, a responsabilização não pode presumir que todo conteúdo falso, controverso ou ofensivo pertença à mesma categoria jurídica.

O uso da expressão fake news, por exemplo, exige cautela. Informação falsa pode decorrer de erro, integrar fraude, servir à manipulação ou produzir ofensa à reputação, situações que não recebem necessariamente o mesmo tratamento. A decisão do STF reforçou essa necessidade de diferenciação ao estabelecer regimes distintos conforme a natureza do conteúdo e preservar disciplina específica para ilícitos contra a honra. A qualificação jurídica depende, portanto, das circunstâncias concretas e não apenas do rótulo atribuído à manifestação (Supremo Tribunal Federal, 2025a; Supremo Tribunal Federal, 2026).

O discurso de ódio apresenta problemática própria quando manifestações dirigem hostilidade ou discriminação contra pessoas ou grupos. Santos et al. (2023) identificam que diferentes plataformas adotam definições, procedimentos de denúncia e sanções específicas para esse tipo de conteúdo, embora não exista uniformidade completa entre suas políticas. Isso demonstra que parte relevante dos limites práticos do discurso digital também é determinada por regras privadas elaboradas pelas empresas responsáveis pelos serviços.

A moderação privada antecede a própria revisão do artigo 19. Klonick (2018) demonstra que grandes plataformas desenvolveram sistemas internos de governança capazes de influenciar aquilo que pode circular em seus ambientes. Termos de uso podem inclusive restringir comportamentos que não constituem ilícitos jurídicos, fazendo com que moderação contratual e responsabilidade civil sejam categorias relacionadas, mas distintas. Uma publicação pode contrariar regras privadas sem gerar necessariamente dever indenizatório, assim como determinadas obrigações jurídicas independem da preferência empresarial.

Esse poder de regulação torna necessária atenção ao risco de censura privada. Fichtner (2022) destaca que a transferência de decisões sobre o discurso para empresas suscita questões de legitimidade, especialmente quando o risco de responsabilidade incentiva remoções preventivas. Por essa razão, maior exigência de atuação precisa ser acompanhada por transparência, procedimentos de contestação e critérios capazes de diferenciar conteúdo efetivamente ilícito de manifestação apenas impopular, controvertida ou contrária aos interesses comerciais da plataforma.

O novo regime brasileiro procura responder a essa tensão por meio de tratamentos diferenciados. Ilícitos graves, conflitos contra a honra, comunicação privada e formas de amplificação recebem soluções distintas, indicando que a liberdade de expressão permanece como limite da responsabilidade sem funcionar como justificativa para proteção uniforme dos provedores. A questão passa a ser determinar qual resposta é proporcional à natureza do conteúdo, à forma de circulação e à participação da plataforma no caso concreto (Supremo Tribunal Federal, 2026).

4.3. Moderação de Conteúdo e Deveres das Plataformas

A moderação não se resume à retirada de publicações. Plataformas podem reduzir alcance, suspender contas, limitar funcionalidades, impedir monetização, aplicar avisos ou encaminhar materiais para revisão. Santos et al. (2023) demonstram que essas práticas integram a governança cotidiana das redes e variam conforme o serviço, a natureza do conteúdo e as políticas adotadas. A responsabilidade civil surge em plano distinto, quando a atuação ou omissão do provedor passa a preencher os requisitos estabelecidos pelo ordenamento.

Essa distinção explica por que a capacidade técnica de remover conteúdo não significa, isoladamente, obrigação de indenizar sempre que uma publicação permaneça disponível, tendo o STJ reconhecido a legitimidade da moderação espontânea conforme regras internas, enquanto a responsabilidade pela não retirada depende dos pressupostos aplicáveis ao caso. O poder de agir e o dever jurídico de agir não são necessariamente equivalentes, embora possam aproximar-se em determinadas situações (Superior Tribunal de Justiça, 2024).

O provedor pode ser avaliado não apenas por sua resposta a uma ordem específica, mas também pela existência de canais de denúncia, procedimentos de análise, mecanismos de revisão e medidas adequadas diante de certas categorias de risco. A responsabilidade passa, portanto, a dialogar com a própria organização da atividade empresarial (Supremo Tribunal Federal, 2026).

Douek (2022) sustenta que a moderação em plataformas de grande escala deve ser analisada sistemicamente, pois seus resultados dependem de políticas, equipes, recursos técnicos e procedimentos institucionais. Essa perspectiva contribui para interpretar o dever de cuidado estabelecido pelo STF, que não transforma uma ocorrência isolada em falha estrutural, mas exige avaliação da qualidade das medidas adotadas diante de padrões relevantes de circulação de determinados ilícitos.

A forma como o provedor participa da distribuição também influencia a intensidade dos deveres. No impulsionamento remunerado, a plataforma presta serviço diretamente voltado à

ampliação do alcance, razão pela qual o STF estabeleceu regime mais rigoroso.

A diferenciação considera que simples hospedagem e amplificação econômica não representam o mesmo nível de participação, aproximando a responsabilidade do papel concreto desempenhado pela empresa na circulação do conteúdo (Supremo Tribunal Federal, 2026; Lemos; Estima, 2026).

Maior diligência, contudo, pode produzir incentivos à remoção excessiva. Lefouili e Madio (2022) demonstram que plataformas reagem economicamente às regras de responsabilidade e podem adotar políticas mais restritivas para reduzir exposição jurídica. Por essa razão, mecanismos de notificação e deveres preventivos precisam coexistir com procedimentos capazes de contestar decisões, reduzindo o risco de que a estratégia empresarial mais segura seja retirar qualquer conteúdo cuja licitude apresente alguma dúvida.

A transparência torna-se, nesse cenário, elemento relevante de controle. Canais de atendimento e procedimentos de contestação permitem questionar tanto a permanência de material possivelmente ilícito quanto a retirada considerada indevida.

A finalidade não é substituir o Poder Judiciário, mas evitar que decisões privadas com impactos significativos sobre direitos sejam tomadas sem critérios verificáveis ou possibilidade de revisão, preocupação incorporada à estrutura exigida pelo STF (Supremo Tribunal Federal, 2026).

A moderação juridicamente adequada exige, assim, proporcionalidade entre capacidade técnica, risco, participação da

plataforma e direitos envolvidos. A empresa não pode ser tratada como observadora neutra quando interfere de forma relevante nos fluxos informacionais, mas também não deve ser convertida em garantidora universal da inexistência de conteúdo ilícito. O regime contemporâneo procura posicionar seus deveres entre esses extremos (Lemos; Estima, 2026; Supremo Tribunal Federal, 2026).

4.4. Inteligência Artificial, Algoritmos e Moderação Automatizada

Ferramentas de inteligência artificial podem reconhecer padrões, comparar arquivos, classificar conteúdos e priorizar materiais para análise humana, oferecendo capacidade de processamento incompatível com fiscalização exclusivamente manual. Essa escala, entretanto, não elimina as dificuldades envolvidas na interpretação jurídica e contextual das manifestações (Gorwa; Binns; Katzenbach, 2020).

A identificação de conteúdo ilícito frequentemente depende de fatores que ultrapassam as palavras ou imagens isoladas. Ironia, contexto político, citações, linguagem figurada e diferenças culturais podem alterar o significado de uma manifestação, enquanto sistemas automatizados operam a partir de dados e critérios previamente definidos. Gorwa, Binns e Katzenbach (2020) ressaltam que essas limitações tornam problemática a expectativa de substituição integral da avaliação humana em decisões envolvendo expressão e contexto.

Crosset e Dupont (2022) também afastam a ideia de que a moderação algorítmica seja um processo puramente tecnológico, demonstrando que o funcionamento desses sistemas depende de interação entre ferramentas automatizadas, equipes humanas,

políticas empresariais e decisões sobre classificação dos conteúdos. O resultado atribuído ao algoritmo decorre, portanto, de escolhas anteriores sobre dados, categorias e níveis de tolerância, o que mantém a empresa responsável pela organização e supervisão do sistema que decidiu adotar.

Essa constatação é relevante para a responsabilidade civil porque o uso de inteligência artificial não elimina o dever de diligência, mas também não transforma cada erro de classificação em prova automática de negligência, e a avaliação precisa considerar como a tecnologia foi implementada, quais mecanismos de revisão existem e como a plataforma reage às falhas identificadas. O desafio está em aproveitar a capacidade de escala da automação sem presumir que seus resultados sejam tecnicamente infalíveis (Gorwa; Binns; Katzenbach, 2020; Crosset; Dupont, 2022).

A automação pode ainda produzir restrições indevidas à expressão. Cobbe (2021) destaca que sistemas algorítmicos interferem naquilo que usuários conseguem publicar ou visualizar e, por isso, ampliam o poder das empresas sobre a comunicação.

Quanto maior a utilização desses mecanismos para remover ou reduzir conteúdos, maior a necessidade de transparência e revisão, principalmente quando os critérios que orientam a decisão não são compreensíveis para a pessoa afetada.

Por outro lado, sistemas automatizados podem contribuir para identificar rapidamente materiais de elevado risco, reconhecer reproduções de conteúdos previamente conhecidos ou detectar padrões de comportamento. A questão não consiste em rejeitar sua utilização, mas em exigir combinação adequada entre tecnologia,

supervisão humana e procedimentos de correção. Vieira e Vieira (2024) relacionam essa expansão da capacidade técnica à necessidade de avaliar também responsabilidade, transparência e controle sobre os mecanismos empregados.

Chesney e Citron (2019) observam que representações falsas com aparência realista podem intensificar danos à imagem e à reputação e aumentar a dificuldade de identificação da falsidade. A mesma tecnologia capaz de auxiliar a moderação, portanto, também pode ampliar formas de produção e disseminação de conteúdos lesivos, demonstrando a ambivalência do desenvolvimento tecnológico.

A resposta jurídica não deve se limitar a exigir mais bloqueios automatizados. Quanto mais intensa for a automação, maior será a necessidade de observar falsos positivos, vieses e mecanismos de recurso, posto que sistemas tecnológicos ampliam as possibilidades de prevenção e resposta, mas não transformam plataformas em agentes capazes de garantir ausência absoluta de ilícitos. O parâmetro permanece relacionado à diligência, à adequação das ferramentas e à capacidade de corrigir falhas relevantes (Cobbe, 2021; Supremo Tribunal Federal, 2026).

4.5. Perspectiva Comparada da Responsabilidade das Plataformas Digitais

Diferentes ordenamentos adotam soluções distintas quanto ao papel das notificações, deveres preventivos, gestão de riscos e proteção da liberdade de expressão. A comparação não revela um modelo universal, mas permite identificar tendências e compreender de que forma outros sistemas enfrentam problemas semelhantes aos observados no Brasil (Brega, 2023).

Na União Europeia, o *Digital Services Act* estabeleceu obrigações de diligência, transparência e avaliação de riscos diferenciadas conforme as características e o porte dos serviços, e a disciplina alcança sistemas de recomendação, procedimentos internos e riscos sistêmicos relacionados às grandes plataformas, aproximando a regulação de uma perspectiva estrutural. Há pontos de contato com o modelo brasileiro posterior ao STF, embora o regime europeu decorra de legislação detalhada, enquanto no Brasil parte significativa da disciplina atual foi construída jurisprudencialmente (União Europeia, 2022; Supremo Tribunal Federal, 2026).

Brega (2023) demonstra que o NetzDG intensificou obrigações relacionadas a conteúdos ilegais, mas gerou preocupação com incentivos à sobre-remoção. Fichtner (2022) acrescenta que transferir decisões relevantes sobre expressão a empresas privadas exige mecanismos de legitimidade, transparência e contestação. Esses problemas são semelhantes à tensão brasileira entre necessidade de resposta rápida e risco de decisões privadas excessivamente restritivas.

O Reino Unido adotou abordagem centrada na segurança dos serviços por meio do *Online Safety Act* de 2023. A legislação demanda determinados provedores avaliação de riscos e adoção de medidas proporcionais destinadas a reduzir a circulação de conteúdos ilícitos, aproximando o foco regulatório da organização do serviço e de seus procedimentos preventivos, e não apenas da resposta oferecida depois de cada publicação individual (Reino Unido, 2023).

Nos Estados Unidos, a *Section 230* do *Communications Decency Act* consolidou proteção ampla aos intermediários em relação a

conteúdo produzido por terceiros e determinadas decisões de moderação. O modelo historicamente atribuiu maior espaço à autonomia das plataformas, embora discussões recentes tenham destacado os limites da intervenção estatal sobre suas escolhas editoriais.

Em Moody v. NetChoice, a Suprema Corte examinou restrições estaduais à moderação e reconheceu relevância constitucional das escolhas de seleção e organização realizadas por plataformas (Estados Unidos, 1996; Supreme Court of the United States, 2024).

A comparação evidencia que nenhuma das experiências se resume à alternativa entre completa imunidade e responsabilidade integral, em que a União Europeia enfatiza diligência e riscos sistêmicos, o Reino Unido adota deveres de segurança, os Estados Unidos conferem proteção mais ampla aos intermediários e o Brasil passou a construir um regime diferenciado após a revisão do artigo 19. Apesar das diferenças, permanece comum a dificuldade de equilibrar atuação das plataformas, proteção de direitos e liberdade de expressão (Brega, 2023; Fichtner, 2022).

Plataformas baseadas em sistemas de recomendação possuem riscos e possibilidades distintas de serviços voltados predominantemente à comunicação privada, princípio igualmente incorporado ao julgamento brasileiro.

A experiência comparada reforça, assim, a importância de proporcionalidade, transparência, avaliação de riscos e mecanismos de contestação, sem justificar a simples transposição de modelos estrangeiros para o ordenamento nacional (União Europeia, 2022; Supremo Tribunal Federal, 2026). O modelo brasileiro resultante dos

Temas 987 e 533 aproxima-se dessa tendência de maior atenção às estruturas internas e aos diferentes graus de participação dos provedores.

Sua particularidade está em combinar hipóteses nas quais permanece relevante a ordem judicial com outras em que diligência, amplificação ou falha sistêmica assumem maior importância. A comparação confirma que a responsabilidade contemporânea das plataformas ultrapassa a simples retirada de uma publicação e alcança a própria governança dos ambientes digitais (Supremo Tribunal Federal, 2026).

5. CONCLUSÃO

A ordem judicial específica deixa de ser o único critério de responsabilização civil das plataformas pela manutenção de conteúdos ilícitos. Em resposta ao problema de pesquisa, as plataformas respondem civilmente por conteúdos ilícitos de terceiros em hipóteses determinadas, independentemente de ordem judicial prévia. A hipótese nula é rejeitada e a hipótese alternativa é confirmada de forma qualificada, pois o artigo 19 permanece aplicável a hipóteses específicas, como os ilícitos contra a honra. O objetivo geral é atingido: a interpretação do STF redistribui os deveres dos provedores e mantém a liberdade de expressão como limite da responsabilização.

Os conteúdos ilícitos formam uma categoria heterogênea e exigem respostas jurídicas distintas. Os deveres de moderação e remoção passam a incluir diligência, transparência, dever de cuidado e prevenção de falhas sistêmicas. A inteligência artificial amplia a

capacidade de moderação e exige supervisão humana, revisão e transparência.

O regime resultante é graduado: a intensidade dos deveres acompanha o grau de participação da plataforma na circulação do conteúdo. A efetividade desse regime depende da interpretação de conceitos como dúvida razoável, diligência e falha sistêmica e de eventual disciplina legislativa. Pesquisas futuras podem examinar a aplicação desses critérios pelos tribunais.

REFERÊNCIAS BIBLIOGRÁFICAS

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil de 1988.** Brasília, DF: Presidência da República, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm. Acesso em: 4 jul. 2026.

_____. **Lei nº 12.965, de 23 de abril de 2014.** Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília, DF: Presidência da República, 2014. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 4 jul. 2026.

BREGA, Gabriel Ribeiro. A regulação de conteúdo nas redes sociais: uma breve análise comparativa entre o NetzDG e a solução brasileira. **Revista Direito GV,** São Paulo, v. 19, e2305, 2023. DOI: 10.1590/2317-6172202305. Disponível em: https://www.scielo.br/j/rdgv/a/qwwzmCyw5FmFQmTpRw3HCQh/. Acesso em: 26 ago. 2026.

CHESNEY, Robert; CITRON, Danielle Keats. Deep fakes: a looming challenge for privacy, democracy, and national security. **California Law Review,** Berkeley, v. 107, n. 6, p. 1753-1820, 2019. Disponível em: https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security. Acesso em: 2 nov. 2026.

COBBE, Jennifer. Algorithmic censorship by social platforms: power and resistance. **Philosophy & Technology,** v. 34, p. 739-766, 2021. DOI: 10.1007/s13347-020-00429-0. Disponível em: https://link.springer.com/article/10.1007/s13347-020-00429-0. Acesso em: 17 ago. 2026.

CROSSET, Valentine; DUPONT, Benoît. Cognitive assemblages: the entangled nature of algorithmic content moderation. **Big Data & Society,** v. 9, n. 2, p. 1-13, 2022. DOI: 10.1177/20539517221143361. Disponível em: https://journals.sagepub.com/doi/10.1177/20539517221143361. Acesso em: 18 ago. 2026.

DOUEK, Evelyn. Content moderation as systems thinking. **Harvard Law Review,** Cambridge, v. 136, n. 2, p. 526-607, 2022. Disponível em: https://harvardlawreview.org/wp-content/uploads/2022/11/136-Harv.-L.-Rev.-526.pdf. Acesso em: 11 ago. 2026.

ESTADOS UNIDOS. **United States Code. Title 47, § 230: protection for private blocking and screening of offensive material.** Washington, DC, 1996. Disponível em: https://www.law.cornell.edu/uscode/text/47/230. Acesso em: 29 ago. 2026.

FICHTNER, Laura. Content moderation and the quest for democratic legitimacy. **Weizenbaum Journal of the Digital Society,** v. 2, n. 2, 2022. DOI: 10.34669/wi.wjds/2.2.2. Disponível em: https://ojs.weizenbaum-institut.de/index.php/wjds/article/view/2_2_2. Acesso em: 22 ago. 2026.

FLUMIGNAN, Wévertton Gabriel Gomes. **Responsabilidade civil dos provedores no Marco Civil da Internet (Lei nº 12.965/14).** 2018. 167 f. Dissertação (Mestrado em Direito Civil) – Faculdade de Direito, Universidade de São Paulo, São Paulo, 2018. DOI: 10.11606/D.2.2018.tde-21082020-010522. Disponível em: https://www.teses.usp.br/teses/disponiveis/2/2131/tde-21082020-010522/. Acesso em: 8 jul. 2026.

_____. **Responsabilidade civil dos provedores de internet: análise teórica, prática e jurisprudencial anterior e posterior ao Marco Civil da Internet (Lei nº 12.965/14).** Londrina, PR: Thoth, 2024.

GILLESPIE, Tarleton. **Custodians of the Internet:** platforms, content moderation, and the hidden decisions that shape social media. New Haven: Yale University Press, 2018. Disponível em: https://doi.org/10.12987/9780300235029. Acesso em: 2 nov. 2026.

GORWA, Robert; BINNS, Reuben; KATZENBACH, Christian. Algorithmic content moderation: technical and political challenges in the automation of platform governance. **Big Data & Society,** v. 7, n. 1, 2020. DOI: 10.1177/2053951719897945. Disponível em: https://journals.sagepub.com/doi/10.1177/2053951719897945. Acesso em: 14 ago. 2026.

KLONICK, Kate. The new governors: the people, rules, and processes governing online speech. **Harvard Law Review,** Cambridge, v. 131, n.

6, p. 1598-1670, 2018. Disponível em: https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf. Acesso em: 13 jul. 2026.

LEFOUILI, Yassine; MADIO, Leonardo. The economics of platform liability. **European Journal of Law and Economics,** v. 53, n. 3, p. 319-351, 2022. DOI: 10.1007/s10657-022-09728-7. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC8857911/. Acesso em: 24 jul. 2026.

LEMOS, André; ESTIMA, Walmir. Responsabilidade civil das plataformas digitais no Brasil: a materialidade das plataformas na análise do julgamento do STF. **Galáxia,** São Paulo, v. 51, e73593, 2026. DOI: 10.1590/1982-2553202673593. Disponível em: https://www.scielo.br/j/gal/a/cf3htthYgh9t3LMXLqJMfLv/. Acesso em: 6 ago. 2026.

LIMA, Cíntia Rosa Pereira de. A responsabilidade civil dos provedores de aplicação de internet por conteúdo gerado por terceiro antes e depois do Marco Civil da Internet (Lei n. 12.965/14). **Revista da Faculdade de Direito, Universidade de São Paulo,** São Paulo, v. 110, p. 155-176, 2015. DOI: 10.11606/issn.2318-8235.v110i0p155-176. Disponível em: https://revistas.usp.br/rfdusp/article/view/115489. Acesso em: 10 jul. 2026.

REINO UNIDO. **Online Safety Act 2023.** 2023 c. 50. London: The Stationery Office, 2023. Disponível em: https://www.legislation.gov.uk/ukpga/2023/50/contents. Acesso em: 31 ago. 2026.

SANTOS, Luiza Carolina dos; TOMAZ, Renata; DIENSTBACH, Dalby; MATOS, Eurico; SANCHES, Danielle. Discurso de ódio on-line: uma

análise das políticas das plataformas digitais para moderação de conteúdo. **E-Compós,** v. 26, p. 1-22, 2023. DOI: 10.30962/ec.2709. Disponível em: https://www.e-compos.org.br/e-compos/article/view/2709. Acesso em: 8 ago. 2026.

SUPERIOR TRIBUNAL DE JUSTIÇA. **Jurisprudência em Teses, n. 222: Marco Civil da Internet – Lei n. 12.965/2014.** Brasília, DF: STJ, 2023. Disponível em: https://www.stj.jus.br/docs_internet/jurisprudencia/jurisprudenciaemteses/Jurisprudencia%20em%20Teses%20222%20-%20Marco%20Civil%20da%20Internet%20-%20Lei%20N%2012%20965%202014.pdf. Acesso em: 16 jul. 2026.

_____. **Recurso Especial nº 2.139.749/SP.** Terceira Turma. Relator: Ministro Ricardo Villas Bôas Cueva. Julgado em 27 ago. 2024. DJe, 30 ago. 2024. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?dt_publicacao=30%2F08%2F2024&num_registro=202300686600. Acesso em: 18 jul. 2026.

_____. **Recurso Especial nº 1.783.309/SP.** Quarta Turma. Relatora: Ministra Maria Isabel Gallotti. Julgado em 4 nov. 2025. DJe, 14 nov. 2025. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?dt_publicacao=14%2F11%2F2025&num_registro=201803175641. Acesso em: 2 set. 2026.

SUPREME COURT OF THE UNITED STATES. **Moody v. NetChoice, LLC, 603 U.S. 707.** Washington, DC, 1 July 2024. Disponível em: https://www.supremecourt.gov/opinions/23pdf/22-277new_8mjp.pdf. Acesso em: 3 set. 2026.

SUPREMO TRIBUNAL FEDERAL. **STF define parâmetros para responsabilização de plataformas por conteúdos de terceiros.** Brasília, DF, 26 jun. 2025. Disponível em: https://noticias.stf.jus.br/postsnoticias/stf-define-parametros-para-responsabilizacao-de-plataformas-por-conteudos-de-terceiros/. Acesso em: 25 jul. 2026.

_____. **Tema 533 da repercussão geral: dever de empresa hospedeira de sítio na internet fiscalizar o conteúdo publicado e de retirá-lo do ar quando considerado ofensivo, sem intervenção do Judiciário.** Leading case: RE nº 1.057.258. Relator: Ministro Luiz Fux. Brasília, DF, 2025. Disponível em: https://portal.stf.jus.br/jurisprudenciaRepercussao/tema.asp?num=533. Acesso em: 8 set. 2026.

_____. **Tema 987 da repercussão geral: discussão sobre a constitucionalidade do art. 19 da Lei nº 12.965/2014.** Leading case: RE nº 1.037.396/SP. Relator: Ministro Dias Toffoli. Brasília, DF, 2026. Disponível em: https://portal.stf.jus.br/jurisprudenciaRepercussao/tema.asp?num=987. Acesso em: 8 set. 2026.

UNIÃO EUROPEIA. **Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022, relativo a um mercado único para os serviços digitais e que altera a Diretiva 2000/31/CE (Regulamento dos Serviços Digitais).** Jornal Oficial da União Europeia, 27 out. 2022. Disponível em: https://eur-lex.europa.eu/eli/reg/2022/2065. Acesso em: 28 ago. 2026.

VIEIRA, Graziela Dutra; VIEIRA, Gabriela Dutra. O ciberespaço, inteligência artificial e responsabilidade civil: análise da

responsabilização civil dos provedores de conexão e de aplicação nos casos de conteúdo ilícito gerado por terceiro. **Revista de Direito da Empresa e dos Negócios,** v. 6, n. 1, p. 5-26, 2024. Disponível em: https://revistas.unisinos.br/index.php/rden/article/view/27926. Acesso em: 22 ago. 2026.

---

[1] Doutor em Direito Ambiental Internacional pela Unisantos. Docente do Curso Superior de Direito do Centro Universitário Santa Terezinha (Cest). E-mail: acesse o artigo original para visualizar o e-mail

[2] Discente do Curso Superior de Direito do Centro Universitário Santa Terezinha (Cest). E-mail: acesse o artigo original para visualizar o e-mail

[3] Discente do Curso Superior de Direito do Centro Universitário Santa Terezinha (Cest). E-mail: acesse o artigo original para visualizar o e-mail