

# LIBERDADE DE EXPRESSÃO E DISCURSO DE ÓDIO NAS REDES SOCIAIS: LIMITES CONSTITUCIONAIS E DESAFIOS DA REGULAÇÃO NO BRASIL

## **FREEDOM OF EXPRESSION AND HATE SPEECH ON SOCIAL MEDIA: CONSTITUTIONAL LIMITS AND REGULATORY CHALLENGES IN BRAZIL**

Artigo submetido em 21 de junho de 2026

Artigo aprovado em 22 de junho de 2026

Artigo publicado em 22 de junho de 2026

### **Scientia et Ratio**

Volume 6 – Número 10 – 2026

ISSN 2525-8532

### **Autores:**

Kaique Campos Duarte[1]

Wellen Cristina Oliveira Barbosa[2]

Thaís Maria Ferreira Simões[3]

**RESUMO:** Este estudo examina os limites constitucionais da liberdade de expressão no Brasil diante da expansão das redes sociais e das tensões com o discurso de ódio. Tem como

objetivo identificar critérios juridicamente consistentes para distinguir expressão legítima de manifestações discriminatórias e para fundamentar restrições compatíveis com o Estado Democrático de Direito. Analisa fundamentos constitucionais, vedação à censura prévia, função democrática da liberdade de expressão e sua relação com dignidade e igualdade, à luz da jurisprudência do Supremo Tribunal Federal, do Marco Civil da Internet, da Lei Geral de Proteção de Dados, do direito comparado e da moderação algorítmica em plataformas digitais. Adota abordagem qualitativa, de natureza bibliográfica, com análise doutrinária e jurisprudencial. Conclui que a liberdade de expressão, embora central, admite restrições proporcionais quando há lesão concreta a direitos fundamentais, exigindo ponderação, fundamentação rigorosa e maior transparência na atuação estatal e privada no ambiente digital.

**Palavras-chave:** liberdade de expressão; discurso de ódio; redes sociais; direitos fundamentais.

**ABSTRACT:** This study examines the constitutional limits of freedom of expression in Brazil in light of the expansion of social networks and the tensions arising from hate speech. Its objective is to identify legally consistent criteria to distinguish legitimate expression from discriminatory manifestations and to ground restrictions compatible with the Democratic Rule of Law. It analyzes the constitutional foundations of freedom of expression, the prohibition of prior censorship, the democratic function of free speech, and its relationship with dignity and equality, considering the jurisprudence of the Brazilian Supreme Federal Court, the Brazilian Internet Bill of Rights, the General Data Protection Law, comparative law, and algorithmic moderation on digital platforms. A qualitative, bibliographic approach is adopted, based on doctrinal and jurisprudential analysis. The study concludes that freedom of expression, although central, admits proportionate restrictions when there is concrete harm to fundamental rights, requiring balancing, rigorous justification, and greater transparency in both state and private actions in the digital environment.

**Keywords:** freedom of expression; hate speech; social networks; fundamental rights.

## INTRODUÇÃO

A expansão das redes sociais, em ritmo quase febril, redesenhou os modos de convivência humana e deslocou os limites do próprio espaço público. A palavra, outrora submetida a mediações institucionais relativamente estáveis, passou a percorrer trajetórias imprevisíveis, alcançando audiências difusas e dissolvendo distinções que antes pareciam minimamente asseguradas, como aquelas entre o íntimo e o coletivo, entre o impulso momentâneo e a elaboração refletida. Nesse cenário mutável, a liberdade de expressão, tradicionalmente alçada à condição de garantia elevada frente ao poder, revela novas camadas de complexidade, ao mesmo tempo em que deixa entrever impasses que o constitucionalismo clássico não fora instado a enfrentar com semelhante intensidade.

A Constituição Federal de 1988, ao erigir a liberdade de expressão à condição de direito fundamental, conferiu-lhe contornos amplos e plurais, abrangendo a exteriorização individual do pensamento e a circulação de ideias que sustenta a experiência democrática. Não obstante, o ambiente digital projeta sobre essa garantia uma incidência peculiar, ao amplificar enunciados que, embora revestidos da aparência de simples opinião, revelam-se aptos a operar como mecanismos de exclusão, silenciamento e violência simbólica. Diante disso, impõe-se uma indagação que escape ao conforto das formulações abstratas e exija enfrentamento direto: até que ponto a ordem constitucional brasileira admite a contenção de manifestações que, embora formalmente enquadradas como expressão, produzem efeitos concretos de marginalização e de corrosão da dignidade humana?

É nesse ponto que se delineia o problema de pesquisa que orienta o presente estudo. Busca-se compreender de que maneira é possível compatibilizar a proteção robusta da liberdade de expressão, tal como concebida na Constituição de 1988, com a necessidade de conter o discurso de ódio nas redes sociais, sem que tal movimento redunde em práticas de censura

incompatíveis com o Estado Democrático de Direito, seja na sua feição estatal, seja em sua vertente privada. Este é um desafio que demanda a reconstrução criteriosa dos parâmetros jurídicos capazes de distinguir o lícito do ilícito no âmbito do discurso público digital.

O objetivo geral consiste em examinar os limites constitucionais da liberdade de expressão no Brasil, com atenção particular às manifestações veiculadas em redes sociais, de modo a identificar critérios juridicamente consistentes para a delimitação do discurso de ódio e para a legitimação de eventuais restrições. Para tanto, propõe-se investigar os fundamentos constitucionais da liberdade de expressão, a vedação à censura prévia, sua dimensão democrática e os limites decorrentes da convivência com outros direitos fundamentais, bem como a construção jurisprudencial do Supremo Tribunal Federal sobre a matéria. Em paralelo, pretende-se aprofundar o exame do conceito de discurso de ódio, suas distinções em relação a outras formas de manifestação e os parâmetros adotados tanto no direito interno quanto em experiências estrangeiras e internacionais. Por fim, almeja-se refletir sobre os desafios regulatórios impostos pelas plataformas digitais, com a proposição de critérios analíticos que favoreçam decisões mais transparentes, proporcionais e compatíveis com a ordem constitucional.

A justificativa do estudo ancora-se na percepção de que o debate contemporâneo acerca da liberdade de expressão já não se limita ao domínio da teoria constitucional, passando a manifestar implicações concretas de elevada sensibilidade social e institucional. A carência de balizas nítidas propicia respostas oscilantes, por vezes condescendentes com conteúdos discriminatórios, por vezes excessivamente restritivas, o que fragiliza simultaneamente a proteção de grupos vulneráveis e a própria vitalidade do debate público.

Soma-se a isso a crescente centralidade das plataformas digitais na mediação do discurso coletivo, circunstância que introduz um elemento adicional de complexidade, tendo em vista a atuação de agentes privados dotados de significativo poder sobre a visibilidade e a circulação de ideias, frequentemente desprovidos dos deveres de fundamentação e controle

que se impõem ao Estado.

Diante de tal quadro, a investigação proposta pretende contribuir para o adensamento teórico e prático da temática, sem abdicar de uma sensibilidade voltada às ambiguidades que marcam a realidade concreta. Em última instância, busca-se interrogar as condições de possibilidade de uma liberdade de expressão que consiga coexistir, ainda que de forma permanentemente desafiada, com o compromisso constitucional de promoção da igualdade e de respeito à dignidade de todas as pessoas.

## 1. LIBERDADE DE EXPRESSÃO COMO DIREITO FUNDAMENTAL

A concepção da liberdade de expressão remonta à organização das pólis gregas, nas quais se desenvolveram as primeiras estruturas de participação política, justiça distributiva e diálogo público. Para Mendes et al (2020), essa manifestação do pensamento tornou-se parte constitutiva do exercício democrático, sendo entendida como direito coletivo de influência nas decisões sociais.

Posteriormente, durante a formação dos Estados modernos, com o progressivo abandono das estruturas feudais, a normatização jurídica passou a contemplar novamente tal prerrogativa, integrando-a aos primeiros textos constitucionais. Com isso, o discurso público deixou de ser privilégio das elites e transformou-se em prerrogativa vinculada à cidadania e ao ordenamento legal (Mendes et al., 2020).

Bravo (2021) considera que a liberdade de expressão é o fundamento dos demais direitos e liberdades. Ausente a liberdade de expressão, inviabiliza-se a organização, a informação, o alerta ou a mobilização na defesa dos direitos humanos. O adequado funcionamento dos sistemas políticos dos Estados democráticos é impossibilitado diante da falta de garantias da liberdade de expressão, exigindo meios de comunicação caracterizados pela independência.

Nesse contexto, Barroso (2019) considera que a liberdade de expressão ocupa posição

central na configuração dos direitos fundamentais, sendo condição estruturante da democracia constitucional, mas passível de restrição quando conflita com outros valores igualmente protegidos, como a dignidade humana. Essa compreensão impõe a necessidade de ponderação entre garantias, considerando que a supremacia de uma liberdade comunicacional ilimitada não se sustenta diante de violações aos direitos de personalidade.

No plano internacional, a Declaração Universal dos Direitos Humanos de 1948 (DUDH/48), em seus arts. 7º, 12, 18, 19 e 29, estabelece como direito humano a liberdade de expressão, vinculando-a à responsabilidade social e à proteção da honra, da imagem e da reputação pessoal. Com fundamento semelhante, a Convenção Americana sobre Direitos Humanos de 1969 (CADH/69), em seus arts. 11, 13 e 14, reconhece a liberdade comunicativa com admissibilidade de restrições legais nas hipóteses de risco à ordem democrática ou à integridade moral de terceiros.

A Declaração Universal dos Direitos Humanos representa a afirmação histórica da dignidade como valor universal, ainda que construída a partir de experiências específicas de violência e exclusão. Esse documento, segundo o autor, traduz o esforço coletivo de superação da barbárie, sendo fruto de um pacto ético entre as nações. Entretanto, a força normativa da Declaração não decorre de sua forma jurídica, mas do consenso ético que a sustenta como horizonte civilizatório da humanidade (Comparato, 2022).

A legitimidade desses direitos decorre da sua efetivação no plano normativo, e não da sua proclamação simbólica, o que exige vigilância institucional permanente (Bobbio, 2024). Assim, o Direito à manifestação do pensamento, guardados os limites éticos e legais, situa-se no bojo dos direitos humanos.

Conforme Mendes et al. (2020), o exercício de qualquer liberdade fundamental, ainda que constitucionalmente assegurado, não possui caráter absoluto, devendo ser limitado quando ultrapassa os marcos do respeito à pessoa. A vedação à censura prévia coexiste com a

possibilidade de responsabilização posterior por manifestações abusivas, especialmente quando configuradas como discurso de ódio. A hermenêutica constitucional exige, nesse ponto, uma leitura sistemática que preserve o núcleo essencial dos direitos em colisão, sem sacrificar o princípio da dignidade.

Bobbio (2024) considera que os direitos humanos não possuem fundamento metafísico estável, sendo construções históricas resultantes da dinâmica política e das lutas emancipatórias que marcam os distintos períodos da modernidade. O deslocamento conceitual proposto por esse autor insere-se na crítica à ideia de universalidade abstrata, ao sustentar que tais prerrogativas derivam de processos sociais e não de enunciados teóricos transcendentais.

O direito deve ser compreendido como uma prática interpretativa orientada por valores morais, sendo a integridade um critério normativo que exige coerência entre decisões passadas e julgamentos presentes. A liberdade de expressão, nesse contexto, não pode ser tratada como um fim em si, mas como parte de uma rede de valores que inclui igualdade, respeito e reconhecimento (Dworkin, 2023). Especialmente considerando a imprescindibilidade do respeito no plano da livre expressão, compreende-se que a superação de práticas que se contrapõem à conduta civilizada, carece de instrumentos normativos efetivos.

Ressalta-se, inclusive, que a discussão no campo da liberdade de expressão conquista maior complexidade quando incorporada ao cenário de desenvolvimento tecnológico, mormente das tecnologias de comunicação e informação, considerando a maior abrangência da comunicação a partir destes recursos. Pinheiro (2022) afirma que o advento das tecnologias digitais implicou o redesenho das fronteiras normativas da liberdade de expressão, sobretudo quanto à responsabilização de intermediários e usuários.

A expansão das tecnologias digitais tem provocado mudanças estruturais nos modos de

produção, circulação e apropriação das informações. Segundo Lévy (2020) esse ambiente informacional é caracterizado por uma multiplicidade de plataformas interativas que possibilitam a comunicação e a reconfiguração dos processos sociais e cognitivos, o que amplia os riscos de manipulação discursiva e disseminação de conteúdos ilícitos.

O ambiente virtual ampliou a incidência de discursos ofensivos, exigindo mecanismos jurídicos que assegurem a convivência entre o pluralismo comunicativo e a proteção da dignidade. A normatividade digital, portanto, deve ser orientada pela proporcionalidade e pela vedação do anonimato como instrumento de violação a direitos fundamentais (Pinheiro, 2022).

Essa relação entre mídias digitais e liberdade de expressão apresenta-se sob tensão constante. Conforme Fiss (2020), embora o direito de manifestar pensamentos, opiniões e informações seja juridicamente protegido, ele não se reveste de caráter absoluto, devendo coexistir com outros direitos fundamentais, como a proteção à honra, à privacidade e à segurança coletiva.

## 2. DISCURSO DE ÓDIO E A PROTEÇÃO DE DIREITOS FUNDAMENTAIS

O discurso de ódio, no interior do debate jurídico contemporâneo, não se deixa reduzir a um simples desdobramento da liberdade de expressão, sob pena de se esvaziar a própria consistência normativa dos direitos fundamentais. Cuida-se de manifestação que ultrapassa o plano da opinião e se projeta como mecanismo de negação simbólica e concreta de indivíduos ou coletividades. Sua configuração costuma envolver incitação à discriminação, à hostilidade ou à violência, dirigida especialmente a grupos historicamente vulnerabilizados, a partir de marcadores como raça, religião, gênero, orientação sexual ou origem. Murari (2023) observa que tais manifestações traduzem uma visão de mundo excludente, com capacidade de reforçar desigualdades já sedimentadas e de conferir aparência de legitimidade a práticas discriminatórias.



A literatura jurídica recente tem se dedicado a delimitar esse fenômeno com maior precisão conceitual, buscando evitar expansões interpretativas que acabem por fragilizar a própria liberdade de expressão. Rodrigues Junior (2022) destaca que o discurso de ódio se diferencia por sua aptidão de produzir efeitos que extrapolam a esfera individual, alcançando a coletividade e afetando o próprio equilíbrio do espaço democrático. Em linha semelhante, Silva (2026) enfatiza o caráter performativo dessas manifestações, aptas a reiterar padrões sociais de exclusão, o que impõe sua análise para além de uma leitura estritamente formal do direito à expressão.

Essa compreensão impõe a necessidade de distinções mais refinadas entre opinião, ofensa e discurso discriminatório. Manifestações ásperas, incômodas ou mesmo moralmente reprováveis não se confundem automaticamente com discurso de ódio, sob pena de se instaurar um ambiente de restrição incompatível com a pluralidade que sustenta a vida democrática. A liberdade de expressão também resguarda enunciados que causam desconforto ou resistência, desde que não avancem para a negação da dignidade de terceiros. Bonavides (2020) já advertia que a convivência com o dissenso integra a própria estrutura da democracia, sendo a supressão indiscriminada de ideias uma forma de empobrecimento do debate público.

Por outro lado, há manifestações que, embora formalmente revestidas de opinião, incorporam conteúdo discriminatório relevante, dirigido à inferiorização de grupos determinados. Nesses casos, a análise deve considerar seu conteúdo, seu alcance e seu impacto social. Denicol (2024) ressalta que a efetivação dos direitos fundamentais impõe a identificação de práticas que, mesmo sob a aparência de liberdade, operam como obstáculos à igualdade material. A distinção, portanto, exige uma leitura substancial das manifestações.

A problemática do discurso de ódio conduz, de maneira quase inevitável, ao exame da colisão entre direitos fundamentais, sobretudo quando se deparam a liberdade de expressão, a dignidade da pessoa humana e o princípio da igualdade. O ordenamento constitucional

brasileiro não organiza tais direitos em uma hierarquia rígida, o que desloca ao intérprete a responsabilidade de construir soluções que preservem, na medida possível, a integridade de cada um deles sem esvaziamento recíproco. Abreu (2022) observa que os direitos fundamentais se desenvolvem em permanente inter-relação, dependentes de um arranjo sensível entre distintas pretensões normativas.

A colisão, longe de representar uma falha do sistema, revela a complexidade das relações sociais que o direito busca regular. Bonavides (2020) aponta que a Constituição de 1988, ao adotar um modelo aberto de direitos fundamentais, admite conflitos internos que devem ser resolvidas por meio de critérios racionais de ponderação. Nesse cenário, a dignidade humana assume função orientadora, funcionando como parâmetro de contenção de práticas que promovem a desumanização do outro.

A jurisprudência do Supremo Tribunal Federal tem desempenhado papel decisivo na delimitação desses contornos, especialmente ao enfrentar casos envolvendo manifestações discriminatórias. No julgamento do HC 82.424/RS, a Corte firmou entendimento no sentido de que a publicação de obras com conteúdo antissemita não se encontra protegida pela liberdade de expressão, reconhecendo o racismo como prática incompatível com a ordem constitucional. A decisão reafirmou a gravidade do discurso discriminatório e sinalizou que a liberdade não pode ser instrumentalizada para legitimar a exclusão.

Em outro momento paradigmático, ao julgar a ADO 26 e o MI 4733, o Tribunal reconheceu a omissão legislativa na proteção de pessoas LGBTQIA+ e equiparou a homofobia ao crime de racismo. Tal posicionamento evidencia uma leitura evolutiva da Constituição, capaz de incorporar novas demandas sociais sem perder de vista os compromissos fundantes com a igualdade e a dignidade.

Já no âmbito do Inquérito 4781, a Corte passou a enfrentar fenômenos mais recentes, relacionados à disseminação de conteúdos potencialmente lesivos no ambiente digital. Ainda

que o debate sobre seus limites e implicações permaneça aberto, o julgamento sinaliza a preocupação institucional com práticas discursivas que possam comprometer a integridade das instituições democráticas e a segurança de indivíduos e grupos.

Essas decisões revelam uma orientação jurisprudencial que, sem negar a centralidade da liberdade de expressão, reconhece a necessidade de sua limitação em face de manifestações que atentem contra direitos fundamentais. A atuação do Supremo Tribunal Federal, nesse particular, contribui para a construção de parâmetros interpretativos capazes de orientar o debate público e a atuação dos demais poderes.

A análise do direito comparado e internacional amplia o horizonte de compreensão acerca do tratamento jurídico do discurso de ódio. No âmbito da Corte Interamericana de Direitos Humanos, observa-se uma preocupação constante com a preservação da liberdade de expressão, sem desconsiderar a necessidade de proteção de grupos vulneráveis. A Corte adota uma abordagem que valoriza o contexto e o impacto das manifestações, reconhecendo que determinadas formas de discurso podem comprometer a própria democracia que a liberdade busca proteger.

Em linha próxima, a Corte Europeia de Direitos Humanos construiu jurisprudência que admite restrições ao discurso de ódio, sobretudo quando há incitação à violência ou estímulo à exclusão de minorias. A Alemanha, por sua vez, adota um modelo mais rigoroso, fortemente influenciado por sua experiência histórica, com previsão de responsabilização mais severa em casos de manifestações discriminatórias. Já os Estados Unidos preservam tradição liberal mais expansiva, conferindo ampla proteção à liberdade de expressão e restringindo-a apenas em situações excepcionalmente graves (Knick, 2024).

A comparação entre esses modelos revela a inexistência de uma solução uniforme, substituída por arranjos normativos moldados por trajetórias históricas, escolhas políticas e sensibilidades culturais próprias. Ainda assim, emerge um eixo comum que atravessa essas

diferentes experiências, qual seja, a recusa em admitir a plena proteção do discurso de ódio sob o argumento de liberdade irrestrita. A dignidade da pessoa humana desponta, nesse cenário, como parâmetro transversal de contenção, ainda que materializado por vias institucionais diversas.

A limitação legítima do discurso de ódio, portanto, representa uma afirmação de seus próprios fundamentos. Ao excluir do âmbito de proteção manifestações que atentem contra direitos fundamentais de terceiros, o ordenamento jurídico preserva a integridade do espaço público, impedindo que ele se converta em arena de violência simbólica. Murari (2023) sustenta que a restrição dessas manifestações encontra respaldo na própria lógica dos direitos fundamentais, cuja finalidade se estende à promoção de uma convivência social minimamente equânime.

A aplicação de tais limites, contudo, não pode prescindir da observância de critérios rigorosos, especialmente aqueles associados ao princípio da proporcionalidade e à vedação da censura prévia. A intervenção estatal deve ser cuidadosamente calibrada, de modo a evitar tanto a permissividade excessiva quanto o cerceamento indevido do debate público. Configura-se como uma tarefa que exige sensibilidade institucional e compromisso com os valores constitucionais.

É a partir dessa encruzilhada que se impõe a reflexão sobre os mecanismos concretos de regulação do discurso no ambiente digital, particularmente no que se refere às redes sociais, cuja atuação tem redefinido, em escala inédita, as condições de circulação da palavra e os critérios de sua contenção.

### 3. REGULAÇÃO DAS REDES SOCIAIS E CRITÉRIOS PARA DISTINÇÃO ENTRE EXPRESSÃO E DISCURSO DE ÓDIO

A disciplina jurídica das redes sociais no Brasil encontra seu eixo normativo inaugural no Marco Civil da Internet, instituído pela Lei nº 12.965/2014, cuja arquitetura revela uma opção

deliberada por um modelo de responsabilidade que privilegia a liberdade de expressão sem abandonar, por completo, a tutela de direitos fundamentais. A lógica adotada afasta, como regra, a responsabilização automática dos provedores de aplicação por conteúdos gerados por terceiros, condicionando-a, em grande medida, ao descumprimento de ordem judicial específica que determine a remoção do material reputado ilícito. Tal desenho normativo decorre de uma tentativa de evitar que plataformas digitais se convertam em instâncias privadas de controle prévio do discurso, o que comprometeria a abertura do espaço público digital.

A exigência de intervenção jurisdicional como requisito para a remoção de conteúdos, ressalvadas hipóteses excepcionais previstas em lei, como aquelas relacionadas à divulgação não autorizada de imagens íntimas, revela uma preocupação com a preservação do debate livre, ainda que por vezes conflituoso. Conforme observa Kallajian (2019), a construção de um ambiente informacional democrático exige cautela diante de soluções que, sob o pretexto de proteção, possam ensejar formas indiretas de censura. A responsabilização posterior, nesse sentido, emerge como mecanismo de equilíbrio, ainda que não imune a críticas quanto à sua efetividade em contextos de rápida disseminação de conteúdos.

A Lei Geral de Proteção de Dados Pessoais, instituída pela Lei nº 13.709/2018, inaugura um campo de reflexão que desloca o olhar jurídico para além da superfície imediata das manifestações no ambiente digital. Ao estabelecer princípios como finalidade, adequação e necessidade, e ao conferir tutela reforçada aos dados pessoais sensíveis, o diploma normativo alarga a percepção dos riscos que atravessam a circulação informacional nas redes sociais. A regulação do discurso, sob tal enfoque, deixa de se restringir ao enunciado explícito das mensagens e passa a alcançar também os mecanismos discretos de coleta, tratamento e organização de dados que, silenciosamente, moldam percepções e orientam escolhas.

A proteção contra usos discriminatórios de dados revela-se particularmente relevante, uma

vez que práticas algorítmicas podem reforçar padrões de exclusão sem que haja manifestação discursiva explícita. Kallajian (2019) assinala que a interseção entre privacidade, informação e liberdade de expressão impõe desafios inéditos ao direito, exigindo a construção de respostas que ultrapassem categorias tradicionais. A transparência no tratamento de dados e a possibilidade de controle por parte dos titulares emergem, assim, como elementos indispensáveis à contenção de práticas que, embora silenciosas, produzem efeitos significativos sobre o espaço público.

No campo penal, a intervenção estatal se apresenta com maior densidade normativa, especialmente diante da necessidade de repressão a condutas discriminatórias. O ordenamento jurídico brasileiro estabelece limites objetivos à liberdade de expressão ao tipificar determinadas manifestações como ilícitas quando atentam contra a dignidade de indivíduos ou grupos. A Lei nº 7.716/1989, ao prever o crime de racismo, e o art. 140, §3º, do Código Penal, ao disciplinar a injúria racial, estruturam um espaço normativo em que a palavra, quando convertida em instrumento de discriminação, perde a proteção conferida pela liberdade de expressão e passa a ser juridicamente sancionada.

A intervenção penal, contudo, não pode ser compreendida de forma simplista, como se qualquer manifestação ofensiva devesse ser automaticamente criminalizada. Souza (2022) adverte que a tutela penal dos discursos potencialmente ofensivos exige critérios rigorosos, sob pena de se comprometer a própria liberdade que se pretende proteger. A identificação do conteúdo discriminatório, seu contexto e seu potencial lesivo constituem elementos indispensáveis à delimitação do ilícito, evitando-se tanto a banalização da punição quanto a tolerância com práticas que violam direitos fundamentais.

A moderação de conteúdo pelas plataformas digitais introduz um elemento adicional de complexidade ao debate, na medida em que transfere a agentes privados parcela significativa do controle sobre a circulação de ideias. A remoção de publicações, o bloqueio de perfis e a desmonetização de conteúdos configuram práticas que, embora

frequentemente justificadas pela necessidade de combater abusos, suscitam questionamentos acerca de sua legitimidade e de seus limites. Salvador (2023) observa que as redes sociais passaram a desempenhar função análoga à de arenas públicas, o que torna particularmente sensível a atuação de seus administradores.

Essa atuação, por não se submeter, em regra, aos mesmos parâmetros de controle e fundamentação exigidos do poder público, pode dar ensejo ao fenômeno da chamada censura privada. A ausência de critérios transparentes e a opacidade dos algoritmos que orientam decisões de moderação contribuem para um ambiente de incerteza, no qual usuários se veem sujeitos a restrições cuja justificativa nem sempre é claramente apresentada. Murari (2023) aponta que a regulação do discurso por plataformas digitais exige a construção de mecanismos que assegurem maior *accountability*, sob pena de se consolidar um poder difuso e pouco controlável.

O risco de censura privada se apresenta como realidade que se manifesta na prática cotidiana das redes sociais. O conflito entre a necessidade de coibir discursos ilícitos e a preservação da liberdade de expressão revela-se particularmente aguda quando decisões de remoção são tomadas sem critérios claros ou sem possibilidade efetiva de contestação. Harff (2022), ao examinar o direito comparado, evidencia que diferentes ordenamentos enfrentam dificuldades semelhantes, ainda que adotem soluções diversas.

A ausência de transparência nas decisões de moderação compromete a confiança dos usuários e a própria legitimidade das plataformas como mediadoras do discurso público. A definição de padrões mais claros, aliada à possibilidade de revisão das decisões, mostra-se indispensável para a construção de um ambiente digital que não se afaste dos parâmetros constitucionais. A regulação, portanto, não pode limitar-se à imposição de deveres formais, devendo alcançar também as práticas concretas que moldam a circulação da palavra.

A distinção entre liberdade de expressão e discurso de ódio demanda a construção de

critérios analíticos capazes de orientar decisões de forma consistente. Entre tais critérios, destacam-se a intenção do emissor, o contexto da manifestação, o potencial de dano, a identificação de grupos vulneráveis atingidos e o impacto social da mensagem. Silva (2026) sustenta que a análise isolada de qualquer desses elementos revela-se insuficiente, sendo necessário um exame integrado que considere a complexidade das interações sociais.

A adoção de parâmetros objetivos contribui para a redução de arbitrariedades e para a promoção de maior segurança jurídica. Murari (2023) ressalta que a identificação do discurso de ódio exige sensibilidade para captar o conteúdo explícito da mensagem e seus efeitos sobre a dignidade dos destinatários.

A aplicação do princípio da proporcionalidade, com seus subprincípios de adequação, necessidade e proporcionalidade em sentido estrito, constitui instrumento indispensável para o controle das restrições à liberdade de expressão, permitindo avaliar se a medida adotada é apta a atingir o fim pretendido, se não há alternativa menos gravosa e se o benefício obtido justifica o sacrifício imposto ao direito restringido.

A formulação teórica desenvolvida por Alexy (2026) oferece base metodológica de elevada consistência para a construção de um modelo normativo voltado à delimitação do discurso de ódio sem recorrer a práticas de censura prévia. Ao conceber os direitos fundamentais como princípios dotados de caráter *prima facie* e submetidos à ponderação, o autor fornece um instrumental apto a orientar decisões em situações de colisão normativa, nas quais a solução não se apresenta de forma imediata ou unívoca. A restrição de um direito, sob essa ótica, exige demonstração argumentativa rigorosa de sua conformidade com critérios racionais e proporcionais.

A incorporação de parâmetros objetivos, somada à existência de controle jurisdicional efetivo, contribui para reduzir margens de arbitrariedade e conferir maior estabilidade às decisões que envolvem a limitação do discurso de ódio. Souza (2022) ressalta que a atuação



estatal, quando chamada a intervir, deve ser cuidadosamente fundamentada, de modo a não comprometer o núcleo essencial da liberdade de expressão. Nesse arranjo, a dignidade da pessoa humana assume função orientadora, funcionando como eixo valorativo que sustenta a construção de um espaço público mais inclusivo e juridicamente responsável.

A articulação entre teoria e prática revela-se, assim, indispensável à construção de um modelo regulatório que contribua para a formação de uma cultura jurídica comprometida com a responsabilidade no uso da palavra. A superação de soluções simplificadoras exige a combinação de instrumentos normativos, mecanismos institucionais e reflexão crítica permanente.

A análise desenvolvida permite perceber que a regulação das redes sociais envolve a compreensão de dinâmicas sociais complexas e em constante transformação. A liberdade de expressão, longe de se apresentar como valor isolado, insere-se em um campo de divergências que exige respostas cuidadosamente construídas, sob pena de se comprometer tanto a proteção de direitos fundamentais quanto a vitalidade do debate público.

## METODOLOGIA

O presente estudo estrutura-se a partir de uma abordagem qualitativa, de natureza eminentemente bibliográfica, orientada pela análise crítica de produções doutrinárias, normativas e jurisprudenciais relacionadas à liberdade de expressão, ao discurso de ódio e à regulação das redes sociais. A opção metodológica implica um exercício interpretativo atento às nuances argumentativas de cada autor mobilizado, buscando identificar convergências, conflitos e lacunas no tratamento jurídico da matéria. Nesse sentido, foram selecionadas obras de referência no campo do direito constitucional e da teoria dos direitos fundamentais, bem como estudos específicos sobre discurso de ódio e ambiente digital, com vistas a construir um panorama teórico consistente e articulado.

A investigação desenvolve-se por meio de leitura analítica e sistematização crítica das

fontes, com especial atenção à interlocução entre doutrina e jurisprudência, notadamente a produção do Supremo Tribunal Federal. A escolha dos autores considerou a relevância acadêmica, a atualidade das obras e a pertinência temática em relação ao objeto da pesquisa. A análise buscou ultrapassar a descrição dos institutos jurídicos, privilegiando a problematização dos conceitos e a reconstrução dos critérios interpretativos aplicáveis à distinção entre liberdade de expressão e discurso de ódio, especialmente no contexto das redes sociais.

## ANÁLISE E DISCUSSÃO

A compreensão contemporânea da liberdade de expressão, quando deslocada para o ambiente digital, mostra-se indissociável da própria expansão histórica dos direitos fundamentais. Abreu (2022) assinala que a concretização dessas garantias se inscreve em um campo atravessado por tensões persistentes, no qual pretensões normativas distintas se entrecruzam em busca de reconhecimento e validade. O advento das redes sociais intensifica tais dinâmicas, ampliando a circulação de discursos que, em outros períodos, permaneceriam confinados a círculos mais restritos de enunciação.

A formulação teórica de Alexy (2026) fornece um instrumental particularmente expressivo para a leitura desses conflitos, ao conceber os direitos fundamentais como princípios submetidos a operações de ponderação quando em situação de colisão. A liberdade de expressão, sob essa ótica, não se apresenta como um comando absoluto, mas como uma garantia cuja aplicação exige a consideração do peso relativo dos valores em confronto. Essa construção permite compreender a possibilidade de restrição de determinadas manifestações, ainda que formalmente protegidas, sempre que se revele a lesão a bens jurídicos de igual envergadura, como a dignidade de terceiros.

A tradição constitucional brasileira, conforme destaca Bonavides (2020), atribui à liberdade de expressão posição de relevo na conformação do regime democrático. Tal centralidade,

contudo, não se sustenta sem a presença de balizas normativas aptas a resguardar a própria coerência do ordenamento constitucional. O enfrentamento do discurso de ódio, nessa linha, exige uma leitura que não dissocie a liberdade comunicativa da proteção de valores estruturantes, entre os quais se destacam a igualdade e a dignidade da pessoa humana, cuja preservação condiciona a legitimidade do exercício das liberdades públicas.

A contribuição de Chadid (2015) revela-se profícua ao permitir a apreensão da historicidade dos direitos fundamentais, demonstrando que sua conformação não se dá de maneira estanque, mas antes acompanha as mutações sociais e políticas que atravessam diferentes épocas. A emergência do discurso de ódio no ambiente das redes sociais pode ser compreendida como manifestação que tensiona categorias jurídicas consolidadas, impondo leituras mais atentas à densidade do meio digital, em que a circulação de conteúdos se acelera e adquire alcance ampliado.

Ao tratar dos entraves à efetivação dos direitos fundamentais, Denicol (2024) destaca a permanência de práticas que, embora aparentem compatibilidade com a liberdade de expressão, acabam por produzir efeitos de exclusão e assimetria no espaço social. Tal constatação conduz à percepção de que a análise jurídica não se esgota na literalidade normativa, exigindo atenção aos desdobramentos concretos das manifestações no convívio coletivo. O discurso de ódio, sob tal prisma, inscreve-se em práticas que reforçam desigualdades historicamente sedimentadas.

A tutela jurisdicional dos direitos fundamentais, conforme enfatiza Fontela (2024), assume função relevante na contenção de excessos e na salvaguarda do espaço público democrático. A atuação do Supremo Tribunal Federal, ao enfrentar controvérsias envolvendo liberdade de expressão e manifestações discriminatórias, expõe a complexidade inerente às decisões que exigem a articulação de valores constitucionais em tensão. A análise desses julgados evidencia um esforço constante de acomodação interpretativa, por vezes marcado por divergências que espelham a própria densidade dos temas enfrentados.

A intersecção entre privacidade, informação e liberdade de expressão, tal como examinada por Kallajian (2019), adquire feições particularmente delicadas quando transposta para o universo das redes sociais. A circulação contínua de dados, associada à atuação opaca de sistemas algorítmicos, interfere na formação de percepções e opiniões de modo por vezes sutil, quase silencioso. Esse quadro impõe ao pensamento jurídico a necessidade de deslocar o foco exclusivamente normativo, alcançando também os modos de funcionamento que estruturam a visibilidade e a difusão das mensagens.

A leitura do discurso de ódio como limite ao exercício da liberdade de expressão, desenvolvida por Murari (2023), evidencia que determinadas manifestações não se ajustam à ideia de exercício legítimo de um direito fundamental. Em certas situações, tais práticas se orientam por dinâmicas que corroem a própria racionalidade dos direitos fundamentais, ao produzir exclusões reiteradas e formas de desumanização de grupos sociais específicos. Daí decorre a exigência de uma delimitação conceitual rigorosa, que evite tanto a ampliação indevida quanto o esvaziamento de sua incidência.

A perspectiva comparada apresentada por Harff (2022) evidencia que diferentes sistemas jurídicos lidam com problemas semelhantes, ainda que por caminhos normativos distintos. Essa diversidade revela a inexistência de soluções universais, já que cada resposta jurídica se vincula a trajetórias históricas e arranjos culturais próprios. Ainda assim, nota-se a recorrência de uma preocupação comum com a salvaguarda da dignidade humana, que opera como referência na contenção de discursos que ultrapassam os limites do dissenso socialmente aceitável.

A análise proposta por Kınık (2024) acrescenta densidade filosófica ao debate ao situar o discurso de ódio no campo da teoria política. Essa abordagem permite entrever que a regulação dessas manifestações envolve escolhas normativas que projetam determinada concepção de convivência social. A definição dos limites entre liberdade e restrição deixa de ser mera operação técnica e passa a refletir opções valorativas sobre o próprio desenho da

esfera pública.

A distinção entre liberdade de expressão e discurso de ódio, conforme sublinha Rodrigues Junior (2022), demanda critérios que levem em consideração forma, conteúdo e efeitos das manifestações. A avaliação do contexto em que a fala se insere, das intenções atribuídas ao emissor e da potencialidade de lesão decorrente do discurso mostra-se indispensável para evitar decisões arbitrárias ou excessivamente amplas. Essa compreensão se articula com a reflexão de Silva (2026), que propõe uma abordagem capaz de apreender a complexidade das interações comunicativas, sem reduzir a análise jurídica a esquemas simplificadores.

Por fim, a tutela penal dos discursos potencialmente ofensivos, analisada por Souza (2022), evidencia os limites e as possibilidades da intervenção estatal. A criminalização de determinadas deve ser cuidadosamente delimitada, de modo a preservar a liberdade de expressão sem tolerar práticas que violem direitos fundamentais. A análise desenvolvida sugere que o enfrentamento do discurso de ódio exige uma reflexão contínua sobre os valores que orientam a convivência democrática.

## CONSIDERAÇÕES FINAIS

As reflexões desenvolvidas ao longo deste estudo conduzem a uma percepção que se afasta de simplificações fáceis e, ao mesmo tempo, resiste à tentação de respostas absolutas. A liberdade de expressão, tal como consagrada na ordem constitucional brasileira, se revela como uma experiência viva, atravessada por conflitos que se intensificam no ambiente digital. A palavra, ao circular com amplitude inédita, deixa de ser apenas veículo de ideias para assumir, em determinadas circunstâncias, a capacidade de produzir efeitos concretos sobre a posição social de indivíduos e grupos.

Não se ignora que a tradição constitucional, ancorada na Constituição de 1988, atribui à liberdade de expressão um estatuto privilegiado, precisamente por reconhecer nela uma das condições de possibilidade da vida democrática. Ainda assim, essa posição de destaque não

autoriza sua compreensão como espaço imune a qualquer forma de contenção. A convivência entre direitos fundamentais, longe de se dar em harmonia espontânea, exige constante trabalho interpretativo, sobretudo quando a manifestação do pensamento se converte em instrumento de negação da dignidade alheia.

O exame do discurso de ódio, sob essa perspectiva, evidencia que determinadas manifestações operam como mecanismos de exclusão simbólica, reforçando desigualdades e naturalizando práticas discriminatórias. A dificuldade, contudo, reside em traçar a linha que separa o dissenso legítimo da violência discursiva. Uma linha que exige atenção ao contexto, à intenção e, sobretudo, aos efeitos produzidos pela linguagem no espaço social.

A atuação do Supremo Tribunal Federal, ao longo dos últimos anos, revela um esforço de construção de parâmetros que não sacrifiquem a liberdade em nome de uma proteção abstrata, nem tolerem, sob o pretexto de neutralidade, práticas que corroem o próprio fundamento dos direitos fundamentais. A jurisprudência indica que a contenção de determinadas manifestações pode representar sua reafirmação em termos mais densos, ao impedir que ela seja instrumentalizada como meio de opressão.

No âmbito das redes sociais, esse desafio adquire contornos ainda mais complexos. A presença de agentes privados na mediação do discurso público introduz um elemento de incerteza que não pode ser ignorado. A ausência de transparência, aliada à velocidade de circulação das informações, amplia o risco tanto de permissividade em relação a conteúdos lesivos quanto de restrições indevidas. O direito, ao buscar responder a esse cenário, vê-se diante da tarefa de equilibrar a proteção da liberdade com a necessidade de estabelecer balizas que impeçam a degradação do espaço público.

A utilização de critérios como a proporcionalidade, a análise contextual e a identificação do potencial de dano revela-se, nesse horizonte, não como uma solução definitiva, mas como um esforço de racionalização das decisões que envolvem a limitação do discurso. A

construção de tais critérios não elimina a dimensão trágica que, por vezes, acompanha a colisão entre direitos fundamentais, mas permite que ela seja enfrentada com maior rigor argumentativo e menor margem para arbitrariedades.

Diante de todo o percurso analítico empreendido, torna-se possível enfrentar, com maior precisão, o problema de pesquisa que orientou este trabalho. A ordem constitucional brasileira admite a contenção de manifestações que, embora formalmente enquadradas como expressão, produzem efeitos concretos de marginalização e de corrosão da dignidade humana, desde que tal contenção não se realize por meio de censura prévia e observe critérios rigorosos de justificação.

A resposta, portanto, se encontra em um movimento contínuo de interpretação, no qual a liberdade de expressão é preservada em sua dimensão mais autêntica, ao mesmo tempo em que se impede sua degeneração em mecanismo de exclusão. O desafio permanece aberto, como expressão das próprias contradições de uma sociedade que, ao buscar garantir a liberdade de todos, precisa enfrentar os modos pelos quais ela pode ser utilizada para negar a humanidade de alguns.

## REFERÊNCIAS

ABREU, Mariana Freitas de. **Direitos fundamentais sociais em crise: desequilíbrio constitucional e onerosidade municipal na prestação de serviço público**. São Paulo: Dialética, 2022.

ALEXY, Robert. **Teoria dos direitos fundamentais**. Tradução Virgílio Afonso da Silva. 3. ed. São Paulo: Malheiros, 2026.

BONAVIDES, Paulo. **Curso de direito Constitucional**. 35. ed. São Paulo: Malheiros Editores Ltda. 2020.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Disponível em:

[http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm). Acesso em: 2 maio 2026.

BRASIL. Supremo Tribunal Federal. **Arguição de Descumprimento de Preceito Fundamental n. 130/DF**. Relator: Min. Carlos Ayres Britto. Julgamento em: 30 abr. 2009. Disponível em: <http://portal.stf.jus.br/processos/detalhe.asp?incidente=12837>. Acesso em: 2 maio 2026.

BRASIL. Supremo Tribunal Federal. **Recurso Extraordinário com Agravo n. 660.861/RJ (Tema 786)**. Relator: Min. Luís Roberto Barroso. Julgamento em: 10 jun. 2015. Disponível em: <https://portal.stf.jus.br/processos/detalhe.asp?incidente=4175529>. Acesso em: 2 maio 2026.

CHADID, Ronaldo. Direitos fundamentais: origem, evolução, precursores doutrinários e seu perfil geral. **Revista Direito UFMS**, Campo Grande, n. 1, v. 1, p. 87-111, 2015.

DENICOL, Karina Albuquerque. **Direitos fundamentais sociais**: elementos delimitadores e obstáculos à efetivação. Curitiba: Paco Editorial, 2024.

FONTELA, Carlos Alberto Ely. **A Proteção Judicial dos Direitos Sociais no Brasil**. Curitiba: Juruá, 2024.

HARFF, Graziela. **Discurso de ódio no direito comparado**: um enfoque sobre o tratamento jurídico nos Estados Unidos, Alemanha e Brasil. Belo Horizonte: Foco, 2022.

KALLAJIAN, Manuela Cibim. **Privacidade, Informação e Liberdade de Expressão**. Curitiba: Juruá, 2019.

Kinik, Burcu Nur Binbuğa. **O discurso de ódio do ponto de vista da teoria política e da regulamentação jurídica**. Belo Horizonte: Nosso Conhecimento, 2024.

MURARI, Georgia Anastácia Campana. **O Discurso do Ódio como Limitação à Liberdade**



**de Expressão no Ordenamento Jurídico Brasileiro.** Curitiba: Juruá, 2023.

Pompeu, G. M.; Coelho, R. J. P. Liberdades de expressão e de informação em face da igualdade e dignidade humana: o caso Siegfried Ellwanger. **Revista Justiça Do Direito**, v. 31, n. 2, p. 329-357, 2017.

RODRIGUES JUNIOR, Álvaro. **Liberdade de Expressão e Liberdade de Informação.** 2 ed. Curitiba: Juruá, 2022.

SALVADOR, João Pedro Favaretto. **Discurso de ódio e redes sociais.** São Paulo: Almedina, 2023.

SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais.** 14. ed. Porto Alegre: Livraria do Advogado, 2024.

SILVA, Robert Antonio Brito da. **O discurso de ódio e a liberdade de expressão:** aspectos filosóficos e jurídicos. Joinville: Guedes Jus Editora, 2026. eBook Kindle.

SOUZA, Renan de Araujo de. **Discurso de ódio e direito penal:** a tutela penal dos discursos potencialmente ofensivos. São Paulo: Dialética, 2022.

---

[1] Advogado, Professor da Faculdade Carajás – Grupo Educacional Carajás, Mestre em Direitos Fundamentais pela Universidade da Amazônia – UNAMA. Especialista em Direito Penal e Processual penal pela Universidade Estácio de Sá – UNESA e Direito Constitucional Faculdade Damásio Educacional – DAMÁSIO. Graduado em Direito pela Faculdade Ideal – Facci | Wyden e Ciências Sociais, Universidade Cidade de São Paulo – UNICID. E-mail: [kaique.ma1507@gmail.com](mailto:kaique.ma1507@gmail.com).

[2] Advogado, Graduado em Direito pela Faculdade Carajás – Grupo Educacional Carajás.

[3] Advogado, Graduada em Direito pela Faculdade Carajás – Grupo Educacional Carajás.