

XXXII CONGRESSO NACIONAL DO CONPEDI SÃO PAULO - SP

DIREITOS E GARANTIAS FUNDAMENTAIS II

FERNANDO DE BRITO ALVES

VIVIANE GRASSI

EDINILSON DONISETE MACHADO

BRUNA AZZARI PUGA

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

Diretoria - CONPEDI

Presidente - Profa. Dra. Samyra Haydée Dal Farra Naspolini - FMU - São Paulo

Diretor Executivo - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

Vice-presidente Norte - Prof. Dr. Jean Carlos Dias - Cesupa - Pará

Vice-presidente Centro-Oeste - Prof. Dr. José Querino Tavares Neto - UFG - Goiás

Vice-presidente Sul - Prof. Dr. Leonel Severo Rocha - Unisinos - Rio Grande do Sul

Vice-presidente Sudeste - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio - Rio de Janeiro

Vice-presidente Nordeste - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

Representante Discente: Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

Conselho Fiscal:

Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão

Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais

Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte

Prof. Dr. Fernando Passos - UNIARA - São Paulo

Prof. Dr. Ednilson Donisete Machado - UNIVEM/UENP - São Paulo

Secretarias

Relações Institucionais:

Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná

Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia

Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal

Comunicação:

Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina

Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul

Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe

Relações Internacionais para o Continente Americano:

Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul

Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão

Prof. Dr. Felipe Chiarello de Souza Pinto - UPM - São Paulo

Relações Internacionais para os demais Continentes:

Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará

Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul

Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI - Santa Catarina

Educação Jurídica

Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicuritiba - PR

Prof. Dr. Rubens Beçak - USP - SP

Profa. Dra. Livia Gaigher Bosio Campello - UFMS - MS

Eventos:

Prof. Dr. Yuri Nathan da Costa Lannes - FDF - São Paulo

Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina

Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

Comissão Especial

Prof. Dr. João Marcelo de Lima Assafim - UFRJ - RJ

Profa. Dra. Maria Creusa De Araújo Borges - UFPB - PB

Prof. Dr. Antônio Carlos Diniz Murta - Fumec - MG

Prof. Dr. Rogério Borba - UNIFACVEST - SC

D597

Direitos e garantias fundamentais II[Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Fernando De Brito Alves, Viviane Grassi, Ednilson Donisete Machado, Bruna Azzari Puga – Florianópolis: CONPEDI, 2025.

Inclui bibliografia

ISBN: 978-65-5274-301-5

Modo de acesso: www.conpedi.org.br em publicações

Tema: Os Caminhos Da Internacionalização E O Futuro Do Direito

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Nacionais. 2. Direitos. 3. Garantias fundamentais. XXXII Congresso Nacional do CONPEDI São Paulo - SP (4: 2025: Florianópolis, Brasil).

CDU: 34

XXXII CONGRESSO NACIONAL DO CONPEDI SÃO PAULO - SP

DIREITOS E GARANTIAS FUNDAMENTAIS II

Apresentação

Direitos e Garantias Fundamentais II

É com grande satisfação que apresentamos a produção acadêmica debatida no Grupo de Trabalho de Direitos e Garantias Fundamentais, no âmbito do XXXII Congresso Nacional do CONPEDI. A presente coletânea reflete a vitalidade da pesquisa jurídica brasileira, reunindo investigações que não apenas reafirmam a centralidade da dogmática constitucional, mas que, sobretudo, enfrentam as tensões contemporâneas de uma sociedade em rede e em constante transformação.

A pauta dos trabalhos apresentados revela a preocupação dos pesquisadores com os novos contornos da esfera pública digital. O GT aprofundou-se no que se denominou vetor jurídico da infodemia, dissecando a colisão entre liberdade de comunicação e desinformação. A judicialização do conflito entre fake news e liberdade de expressão, a responsabilidade digital frente ao discurso de ódio e os limites do humor — exemplificados no debate sobre o caso Léo Lins — demonstram a urgência de balizas hermenêuticas para o ambiente virtual. Neste eixo, destacam-se ainda as análises sobre a aplicação da LGPD, a interface entre Inteligência Artificial e a proteção de crianças e adolescentes, e as inovadoras propostas de um direito à vida analógica e à desconexão sob uma perspectiva garantista.

Não obstante o foco tecnológico, o Grupo de Trabalho manteve firme o olhar sobre a materialidade da vida e a justiça social. Foram intensos os debates acerca da função social da propriedade, da usucapião e do direito à moradia adequada sob o paradigma do PIDESC. Questões sensíveis como a relativização da impenhorabilidade do salário, a mitigação do mínimo existencial e a proporcionalidade nas sanções políticas tributárias (IPTU) evidenciaram a busca por um equilíbrio entre a eficácia econômica e a dignidade humana.

A proteção de grupos vulnerabilizados ocupou lugar de destaque. As pesquisas trouxeram à luz a violência estrutural contra a mulher e a luta pela autonomia privada feminina — seja em interpretações dworkinianas, seja na contestação de barreiras em concursos militares. No espectro da infância e juventude, os artigos transitaram da evolução das políticas de acolhimento às inovações legislativas recentes. O GT também acolheu críticas contundentes sobre a exclusão social, abordando desde a inclusão de pessoas com sofrimento mental até a supressão de direitos no sistema prisional e o "estado de coisas" da dignidade encarcerada.

Por fim, a densidade teórica do evento se revelou nas discussões bioéticas e de filosofia do direito. O direito à morte digna, a recusa terapêutica e a governança médica foram analisados par e passo com reflexões sobre a biopolítica e a "vida nua". A teoria constitucional foi revisitada através das lentes de Günther Teubner e Thomas Vesting, discutindo a fragmentação constitucional e o Estado em rede, bem como o debate sobre o direito ao esquecimento na reforma civilista.

Os textos aqui reunidos são o resultado de um diálogo profícuo e rigoroso. Convidamos a comunidade acadêmica a debruçar-se sobre estas páginas, que representam um retrato fiel e desafiador do estado da arte da pesquisa em Direitos Fundamentais no Brasil.

São Paulo, primavera de 2025.

Prof. Dr. Fernando de Brito Alves - UENP

Profa. Dra. Viviane Grassi - UNIFACVEST

Prof. Dr. Ednilson Donisete Machado - UENP

Profa. Dra. Bruna Azzari Puga - UPM

LIBERDADE DE EXPRESSÃO E DISCURSO DE ÓDIO NAS REDES SOCIAIS: ENTRE A PROTEÇÃO CONSTITUCIONAL E A RESPONSABILIDADE DIGITAL

FREEDOM OF EXPRESSION AND HATE SPEECH ON SOCIAL MEDIA: BETWEEN CONSTITUTIONAL PROTECTION AND DIGITAL RESPONSIBILITY

Frank Aguiar Rodrigues ¹

Resumo

Este artigo examina a tensão entre a liberdade de expressão e o discurso de ódio nas redes sociais, com enfoque no ordenamento brasileiro. A partir de análise teórico-jurídica e revisão bibliográfica, fundamentada na obra *Liberdade de expressão e o discurso de ódio nas redes* (FAM, 2025), são discutidos (i) o conceito e o alcance da liberdade de expressão; (ii) a delimitação do discurso de ódio e seus impactos; e (iii) a responsabilidade das plataformas digitais e os desafios contemporâneos de moderação de conteúdo. O trabalho dialoga com o marco constitucional de 1988, com a legislação infraconstitucional (v.g., Marco Civil da Internet) e com a jurisprudência recente do STF, notadamente o debate sobre o art. 19 do Marco Civil (ADIs 6.949, 6.960, 7.030 e ADPF 403). Conclui-se que o direito de se expressar, embora nuclear à democracia, não é absoluto e deve ser compatibilizado com a dignidade humana, impondo deveres correlatos aos usuários e às plataformas.

Palavras-chave: Liberdade de expressão, Discurso de ódio, Redes sociais, Responsabilidade das plataformas, Direitos fundamentais

Abstract/Resumen/Résumé

This article examines the tension between freedom of expression and hate speech on social media within the Brazilian legal framework. Based on a theoretical-legal analysis and literature review—grounded in the book *Liberdade de expressão e o discurso de ódio nas redes* (FAM, 2025)—it discusses (i) the scope of freedom of expression; (ii) the definition and impacts of hate speech; and (iii) platform liability and content-moderation challenges. It engages with Brazil's 1988 Constitution, the Internet Civil Framework, and recent Supreme Court case law, particularly the debate over Article 19 of the Internet Civil Framework (ADIs 6.949, 6.960, 7.030; ADPF 403). The paper argues that freedom of expression, while central to democracy, is not absolute and must be balanced with human dignity, imposing correlative duties on users and platforms.

Keywords/Palabras-claves/Mots-clés: Freedom of expression, Hate speech, Social networks, Platform liability, Fundamental rights

¹ Doutor em Direito, ITE. Mestre em Direito, UNIVEM. Especialista em Direito e Processo do Trabalho, Universidade Anhanguera. Bacharel em Direito – FAI. Prof Efetivo UESPI.

INTRODUÇÃO

A expansão das redes sociais transformou a esfera pública, viabilizando circulação massiva de opiniões, dados e narrativas em velocidade inédita. Nesse ambiente, a liberdade de expressão continua sendo um pilar democrático, porém submetido a fricções intensas quando se confronta com conteúdos que incitam discriminação, hostilidade ou violência — o chamado discurso de ódio. O livro que serve de base a este artigo sistematiza esse debate em doze capítulos, cobrindo do conceito de liberdade de expressão ao papel das plataformas, decisões paradigmáticas do STF, responsabilidade civil e penal, moderação de conteúdo, ética comunicacional e diretrizes para uma cultura digital responsável.

No plano conceitual, o discurso de ódio é compreendido como comunicação que incita violência ou preconceito contra indivíduos ou grupos em razão de características como raça, religião, etnia, orientação sexual ou deficiência; sua aplicação prática é contextual, sujeita a ambiguidades e disputas interpretativas. Essa definição — recorrente em documentos e iniciativas internacionais — orienta a análise jurídica e sociopolítica do tema, iluminando efeitos reais sobre a dignidade e a participação de minorias na esfera pública digital.

A premissa normativa do trabalho é que a liberdade de expressão não é absoluta; ela convive, em regime de ponderação, com outros direitos fundamentais (dignidade, igualdade, honra, imagem), sendo legítima a imposição de limites proporcionais em situações de abuso de direito e lesão a bens constitucionalmente protegidos. O material de referência enfatiza esse equilíbrio e destaca que a proteção jurídica deve ser acompanhada de políticas de moderação e de educação digital que promovam convivência plural e debate público inclusivo.

No Brasil, a discussão contemporânea ganhou relevo com o julgamento, pelo STF, do regime de responsabilidade das plataformas à luz do art. 19 do Marco Civil da Internet. O debate contrapõe, de um lado, um modelo que condiciona a responsabilização à ordem judicial específica e, de outro, entendimentos que reconhecem deveres de cuidado proativos em hipóteses notórias de ilicitude (v.g., símbolos nazistas, incitação à violência, racismo), com possível responsabilização pela inércia da plataforma. Esse contencioso — ventilado nas ADIs 6.949, 6.960, 7.030 e na ADPF 403 — aponta mudança paradigmática: as plataformas deixam de ser vistas apenas como meios neutros para se afirmarem como agentes reguladores de seus ambientes, com deveres correlatos de prevenção e resposta.

Objetivamente, este artigo busca responder às seguintes questões: (i) quais limites normativos balizam a liberdade de expressão frente ao discurso de ódio? (ii) como a jurisprudência do STF tem ajustado esse balanço na era digital? (iii) quais são as consequências

jurídicas e éticas para usuários e plataformas? Para tanto, adota-se abordagem teórico-analítica: (a) revisão do conceito e do arcabouço jurídico da liberdade de expressão; (b) delimitação do discurso de ódio e de seus efeitos; (c) exame dos deveres e responsabilidades na moderação de conteúdo, considerando impactos práticos (cancelamento digital, responsabilização civil e penal) apresentados na obra de referência.

Do ponto de vista metodológico, trabalha-se com análise bibliográfica e jurisprudencial. Além do livro base (ALBUQUERQUE et al., 2025), mobiliza-se literatura constitucional e de teoria dos direitos fundamentais (v.g., Barroso; Canotilho; Bobbio) e decisões paradigmáticas do STF relacionadas a liberdade de imprensa, controle de abuso e discurso de ódio. A própria obra sintetiza exemplos e impactos sociais decorrentes da retórica hostil, reforçando que a ética da comunicação e a responsabilidade compartilhada são dimensões indissociáveis da tutela jurídica.

LIBERDADE DE EXPRESSÃO E SEU AR CABOUÇO JURÍDICO

A liberdade de expressão constitui-se em um dos pilares centrais da democracia moderna. Reconhecida expressamente como direito fundamental na Constituição Federal de 1988, em dispositivos como o art. 5º, incisos IV, IX e XIV, garante a livre manifestação do pensamento, a vedação ao anonimato, bem como a livre comunicação de ideias e informações. Essa consagração constitucional representou um marco decisivo no processo de redemocratização do Brasil, após o longo período de censura e repressão durante a ditadura militar (1964–1985). O texto constitucional, ao proteger de forma ampla o direito de expressão, procurou romper com a experiência autoritária, garantindo que o debate público passasse a ser regido pela pluralidade de vozes e pelo respeito ao dissenso.

Contudo, como bem registra a obra-base, a liberdade de expressão não se constitui em direito absoluto. Tanto a tradição constitucional quanto a jurisprudência consolidada do Supremo Tribunal Federal reconhecem que o exercício desse direito fundamental deve ser compatibilizado com outros bens igualmente tutelados, como a honra, a intimidade, a vida privada e, sobretudo, a dignidade da pessoa humana, inscrita no art. 1º, inciso III, da Constituição Federal. Nesse ponto, o livro enfatiza que a ordem constitucional brasileira estabelece um verdadeiro “tecido interligado” de direitos fundamentais, em que a livre expressão encontra limites na proteção da integridade moral e existencial dos indivíduos.

Essa concepção é reforçada pelo STF em diversos julgados paradigmáticos. O Tribunal tem reiteradamente afirmado que a liberdade de expressão ocupa posição preferencial

na hierarquia dos direitos fundamentais, sendo essencial para a preservação do regime democrático. Todavia, quando utilizada para fomentar práticas discriminatórias, incitar a violência ou atacar grupos vulneráveis, deixa de ser um exercício legítimo e passa a configurar abuso de direito. A célebre decisão proferida no HC 82.424/RS (Caso Ellwanger) consolidou essa orientação: o discurso racista não encontra guarida sob o manto da liberdade de expressão, pois viola frontalmente a dignidade da pessoa humana e ameaça a própria coesão social.

No plano doutrinário, essa interpretação também encontra respaldo. Luís Roberto Barroso (2020) ressalta que a liberdade de expressão é “condição da democracia”, na medida em que permite a circulação de ideias, o confronto de opiniões e o controle social das instituições. No entanto, alerta que tal liberdade não pode ser compreendida como salvo-conduto para manifestações que promovam intolerância, ódio ou exclusão. Em perspectiva semelhante, José Joaquim Gomes Canotilho (2003) adverte que os direitos fundamentais não se apresentam em termos absolutos, mas sim como garantias que coexistem em permanente tensão e que demandam um exercício constante de ponderação constitucional.

O livro também recupera a relevância histórica da consagração da liberdade de expressão em 1988. Ao encerrar um ciclo de censura e perseguições políticas, a Assembleia Constituinte buscou construir um espaço público em que a diversidade e o pluralismo fossem valorizados. Essa abertura democrática, contudo, não significou a eliminação de responsabilidades: ao mesmo tempo em que protegeu a livre manifestação, a Constituição instituiu mecanismos de responsabilização civil e penal para manifestações abusivas. Dessa forma, a estrutura normativa brasileira consagra o princípio de que não existe liberdade sem responsabilidade.

Outro aspecto enfatizado na obra é que a liberdade de expressão deve ser compreendida dentro de um modelo de democracia substancial, que não se limita a assegurar a formalidade do debate, mas que garante a todos os indivíduos condições de participação igualitária. Isso significa que manifestações que visem silenciar ou intimidar determinados grupos sociais — como discursos racistas, misóginos, homofóbicos ou xenofóbicos — não podem ser toleradas como simples expressões de opinião, pois desestabilizam a igualdade de condições no espaço público. Nesse sentido, a vedação ao anonimato, prevista no art. 5º, inciso IV, CF, reforça a ideia de responsabilização e contribui para coibir abusos.

A jurisprudência brasileira, acompanhando tendências internacionais, tem destacado a necessidade de ponderação entre a liberdade de expressão e outros valores constitucionais. O STF, em decisões recentes, reafirmou que a proteção desse direito fundamental não alcança manifestações criminosas, tais como apologia ao nazismo,

negacionismo do Holocausto ou incitação à violência contra minorias. A experiência comparada, especialmente no âmbito europeu, demonstra que a defesa da dignidade humana é princípio estruturante da limitação da liberdade de expressão.

Assim, a estrutura normativa brasileira legitima a restrição de manifestações que ultrapassem o campo do debate público legítimo e passem a configurar abusos de direito, notadamente nos casos de incitação à violência, racismo, nazismo ou ataques discriminatórios. A Constituição Federal, ao mesmo tempo em que consagra a livre manifestação, exige que seu exercício seja compatível com o respeito à integridade e à dignidade de todos os cidadãos.

Portanto, pode-se concluir que a liberdade de expressão no Brasil é simultaneamente ampla e limitada: ampla porque constitui direito fundamental preferencial, indispensável à democracia e à cidadania; limitada porque não se sobreponha à dignidade da pessoa humana, nem pode ser invocada como escudo para práticas de intolerância e violência. A obra analisada evidencia que esse equilíbrio é condição essencial para a manutenção de uma esfera pública plural e democrática, capaz de assegurar tanto a diversidade de ideias quanto a proteção contra discursos que corroem a convivência social.

O DISCURSO DE ÓDIO NAS REDES: CONCEITO, MANIFESTAÇÕES E IMPACTOS

O discurso de ódio é um fenômeno social complexo que se manifesta de forma crescente nas redes sociais, constituindo um desafio significativo para a convivência democrática e para a proteção dos direitos fundamentais. De modo geral, entende-se por discurso de ódio qualquer manifestação comunicativa que promova preconceito, hostilidade, discriminação ou violência contra indivíduos ou grupos em razão de características identitárias, como raça, etnia, religião, gênero, orientação sexual, deficiência ou qualquer condição que configure vulnerabilidade histórica ou social. Esse tipo de manifestação não apenas atenta contra a dignidade e a igualdade dos indivíduos, mas também compromete o próprio tecido social, corroendo a confiança mútua e a capacidade de diálogo presente na esfera pública.

Organizações internacionais, como a Organização das Nações Unidas (ONU), reconhecem que o discurso de ódio representa uma ameaça concreta à democracia, à paz social e à convivência harmônica entre diferentes grupos sociais. A ONU enfatiza que manifestações discriminatórias ou hostis podem gerar efeitos cumulativos, impactando de forma significativa a vida de indivíduos e coletividades. A União Europeia compartilha dessa perspectiva, apontando que, em ambientes digitais, o discurso de ódio encontra terreno fértil para sua

expansão devido à velocidade de disseminação, à possibilidade de anonimato e ao efeito multiplicador dos algoritmos que priorizam conteúdos com maior engajamento, independentemente de seu caráter ético ou jurídico. Nesse contexto, os meios digitais não apenas reproduzem preconceitos existentes, mas também contribuem para a sua amplificação, tornando o discurso de ódio mais intenso e difícil de controlar.

A doutrina brasileira aponta que o discurso de ódio não se limita a agressões diretas ou explícitas, como xingamentos, ameaças e insultos, mas também se manifesta de forma mais sutil, através de ironias, piadas, estereótipos naturalizados e conteúdos que reforçam desigualdades históricas. Memes, narrativas falsas ou distorcidas e campanhas de desinformação podem funcionar como veículos de discriminação, consolidando preconceitos e estigmatizando grupos sociais específicos. Essa forma de comunicação, embora muitas vezes disfarçada de humor ou crítica, tem efeitos reais sobre indivíduos e comunidades, perpetuando estereótipos e reforçando hierarquias sociais injustas.

Os impactos do discurso de ódio sobre o indivíduo são amplos e profundos. No plano psicológico, as vítimas frequentemente experimentam sentimentos de humilhação, medo, ansiedade e isolamento, que comprometem diretamente a saúde mental e o bem-estar. O silenciamento imposto pelo receio de novas agressões é outro efeito relevante, uma vez que limita a liberdade de expressão do próprio alvo e reduz sua participação ativa na esfera pública. Além disso, a exposição contínua a mensagens hostis ou discriminatórias pode gerar traumas cumulativos, afetando a autoestima e a percepção de segurança pessoal. Esses efeitos não se restringem à vítima individual, reverberando em famílias, grupos de apoio e comunidades inteiras, criando um ambiente de insegurança e desconfiança generalizada.

Os efeitos coletivos do discurso de ódio também são significativos, uma vez que interferem na qualidade do debate público e na coesão social. A polarização social é um dos resultados mais evidentes, na medida em que a circulação de mensagens hostis contribui para a criação de bolhas ideológicas, onde diferentes grupos têm contato limitado com perspectivas divergentes. Essa dinâmica dificulta o diálogo democrático e promove a fragmentação da sociedade, tornando mais complexa a construção de consensos e soluções coletivas para problemas comuns. Além disso, o discurso de ódio pode funcionar como catalisador de comportamentos violentos, incentivando a radicalização política e o engajamento em práticas criminosas ou extremistas. Nesse sentido, o fenômeno não se limita à esfera virtual, mas frequentemente transborda para o mundo físico, potencializando conflitos e ameaçando a segurança coletiva.

No Brasil, a legislação tem buscado enfrentar o discurso de ódio conciliando a proteção à liberdade de expressão com a necessidade de combater manifestações discriminatórias. A Lei 7.716/1989 tipifica crimes resultantes de discriminação ou preconceito de raça ou cor, estabelecendo penas para condutas que atentem contra a igualdade e a dignidade humana. Posteriormente, a Lei 9.459/1997 ampliou as hipóteses de discriminação, incluindo gênero e outras características pessoais, reforçando a proteção de grupos historicamente vulneráveis. Mais recentemente, a Lei 14.532/2023 incorporou a injúria racial como forma de racismo, consolidando a tendência de ampliar a abrangência da legislação antidiscriminatória. A jurisprudência do Supremo Tribunal Federal também reforça essa linha de proteção. No julgamento do HC 82.424/RS, conhecido como Caso Ellwanger, o Tribunal afirmou de forma categórica que a liberdade de expressão não protege manifestações de caráter racista ou que neguem a dignidade humana, consolidando o entendimento de que o discurso de ódio é incompatível com o núcleo axiológico da Constituição, especialmente com os princípios do pluralismo, da dignidade e da igualdade.

A delimitação jurídica do discurso de ódio demanda equilíbrio e cautela, pois é necessário evitar censura indevida e garantir que a liberdade de expressão continue sendo um pilar da democracia. Ao mesmo tempo, é imprescindível que a legislação e a jurisprudência assegurem proteção efetiva contra manifestações que promovam discriminação, violência ou exclusão social. Para tanto, a interpretação das normas deve considerar o contexto social e tecnológico em que os conteúdos circulam, os efeitos concretos sobre as vítimas e a sociedade, bem como a necessidade de educação digital e conscientização cidadã sobre o impacto do discurso de ódio. A análise detalhada do fenômeno evidencia que seu enfrentamento não pode se restringir ao campo jurídico, sendo fundamental também o desenvolvimento de políticas públicas, ações educativas e mecanismos tecnológicos de moderação, que promovam um ambiente digital seguro e inclusivo.

O enfrentamento do discurso de ódio nas redes sociais é, portanto, um desafio multidimensional, que exige compreensão profunda de seus múltiplos efeitos e da complexidade de sua propagação. Trata-se de uma questão que envolve direitos fundamentais, dinâmicas sociais, tecnologia e cultura digital, sendo necessário articular ações jurídicas, educativas e tecnológicas para mitigar seus impactos. A legislação brasileira, aliada à jurisprudência do Supremo Tribunal Federal, constitui um marco importante nesse processo, orientando a construção de limites claros para a liberdade de expressão em casos que atentem contra a dignidade humana e a igualdade. O debate sobre o discurso de ódio é, em última análise, essencial para a preservação da democracia, da coesão social e do pluralismo,

contribuindo para a criação de um ambiente digital que respeite a diversidade, a dignidade e os direitos de todos os cidadãos.

PLATAFORMAS DIGITAIS, RESPONSABILIDADE E MODERAÇÃO DE CONTEÚDO

O avanço das plataformas digitais trouxe consigo um conjunto de desafios complexos, tanto do ponto de vista jurídico quanto ético, particularmente no que se refere à circulação de discursos de ódio, desinformação e conteúdos ilícitos. Um dos grandes debates contemporâneos envolve a definição precisa dos limites da responsabilidade civil e penal das plataformas diante do vasto universo de informações geradas por usuários. Nesse cenário, a circulação de conteúdos ofensivos, discriminatórios ou potencialmente prejudiciais coloca em evidência a necessidade de equilibrar liberdade de expressão, inovação tecnológica e proteção aos direitos fundamentais. O Marco Civil da Internet, instituído pela Lei nº 12.965 de 2014, representou um marco regulatório essencial nesse contexto, estabelecendo princípios, garantias, direitos e deveres para o uso da internet no Brasil. Em particular, o artigo 19 do referido diploma legal determina que os provedores de aplicação — ou seja, as plataformas digitais — somente podem ser responsabilizados civilmente por conteúdos publicados por terceiros após ordem judicial específica. Tal regra busca proteger a liberdade de expressão na internet, evitando que provedores pratiquem censura privada e garantindo um ambiente propício à inovação e ao desenvolvimento tecnológico.

A lógica adotada pelo Marco Civil da Internet apresenta inspiração no modelo norte-americano, notadamente na Seção 230 do Communications Decency Act, que estabelece imunidade para provedores em relação a conteúdos de terceiros, salvo em situações excepcionais. No entanto, como destacado na obra consultada, essa abordagem também gerou um vazio de responsabilização em situações de violação grave de direitos, nas quais conteúdos notoriamente ilícitos — como pornografia infantil, incitação à violência, apologia ao racismo ou a práticas discriminatórias — circulam livremente, colocando em risco direitos fundamentais e a própria ordem social. Essa lacuna evidencia um desafio regulatório significativo: como equilibrar proteção à liberdade de expressão, inovação digital e responsabilização efetiva das plataformas diante de abusos flagrantes.

A jurisprudência do Supremo Tribunal Federal tem avançado nesse debate. No ano de 2023, ao julgar as Ações Diretas de Inconstitucionalidade 6.949, 6.960, 7.030, bem como a Arguição de Descumprimento de Preceito Fundamental 403, o Tribunal analisou a aplicação

do artigo 19 do Marco Civil da Internet, modulando seu alcance e delimitando situações de responsabilização das plataformas digitais. Embora o STF tenha mantido a regra geral de que provedores só respondem após ordem judicial específica, admitiu exceções relevantes: em casos de conteúdos manifestamente ilícitos, a inércia das plataformas pode gerar responsabilização imediata. Entre os exemplos de conteúdos ilícitos, destacam-se a pornografia infantil, a incitação ao nazismo e a apologia ao racismo. Esse posicionamento reposiciona as plataformas digitais não apenas como intermediárias neutras, mas como agentes ativos na regulação do conteúdo, com o dever de adotar mecanismos eficientes de prevenção, monitoramento e moderação. A decisão reforça a ideia de que a responsabilidade das plataformas não é absoluta, mas se ajusta de acordo com a gravidade e notoriedade da violação, buscando compatibilizar liberdade de expressão e proteção de direitos fundamentais.

Além do aspecto jurídico, surge uma dimensão ética fundamental nesse debate. O fenômeno conhecido como “cancelamento digital” exemplifica o dilema ético enfrentado pelas plataformas: denúncias legítimas de condutas ofensivas podem rapidamente se transformar em linchamentos virtuais, com efeitos desproporcionais sobre indivíduos ou grupos. A velocidade de disseminação de informações, combinada com algoritmos de engajamento e anonimato relativo, cria um ambiente no qual julgamentos públicos e punitivos muitas vezes substituem procedimentos legais ou princípios de justiça. Dessa forma, o desafio ético consiste em estabelecer limites claros entre a moderação necessária para proteger a dignidade humana e a liberdade de expressão, evitando que mecanismos de denúncia e exclusão de conteúdo sejam utilizados de forma arbitrária ou excessiva.

No cenário internacional, observa-se uma tendência crescente de maior regulação das plataformas digitais. A União Europeia, por meio da Digital Services Act, estabeleceu exigências rigorosas para as empresas, incluindo transparência algorítmica, procedimentos robustos de moderação, obrigações de prestação de contas e mecanismos para remoção rápida de conteúdos ilícitos. O objetivo dessas normas é reduzir a circulação de material prejudicial sem comprometer a pluralidade de opiniões e o debate público, promovendo um modelo de cogovernança digital em que direito, ética e tecnologia se articulam de forma integrada. No Brasil, tramita o Projeto de Lei nº 2.630 de 2020, popularmente conhecido como “PL das Fake News”, que busca estabelecer deveres mais claros às plataformas digitais, impondo responsabilidades adicionais em relação à circulação de informações falsas, discurso de ódio e práticas que atentem contra direitos fundamentais. Esses movimentos regulatórios indicam que a discussão sobre moderação de conteúdo é global e envolve desafios comuns: como proteger

usuários vulneráveis, responsabilizar plataformas e, ao mesmo tempo, garantir liberdade de expressão e inovação tecnológica.

A moderação de conteúdo, portanto, emerge como um espaço de cogovernança digital, no qual o direito, a ética e a autorregulação das empresas devem convergir. As plataformas são chamadas a adotar políticas claras e transparentes, criar canais acessíveis para denúncias, empregar inteligência tecnológica e humana na análise de conteúdos e fornecer informações claras sobre critérios de remoção e suspensão de material. Ao mesmo tempo, cabe à legislação estabelecer parâmetros que definam limites objetivos para a atuação das plataformas, evitando que medidas de moderação se transformem em censura privada ou em instrumentos de controle indevido do debate público. A obra consultada destaca que esse equilíbrio é delicado, uma vez que decisões equivocadas podem gerar efeitos significativos, seja pela omissão diante de conteúdos ilícitos ou pela remoção injustificada de conteúdos legítimos.

O desenvolvimento de normas nacionais e internacionais reflete a complexidade desse cenário. O Marco Civil da Internet, a jurisprudência do STF e projetos como o PL das Fake News representam tentativas de construir um sistema regulatório que contemple inovação tecnológica, liberdade de expressão e proteção aos direitos fundamentais. No plano ético, a moderação de conteúdo deve respeitar princípios de proporcionalidade, transparência e accountability, evitando excessos e garantindo que usuários tenham acesso a mecanismos de recurso e revisão. O papel das plataformas digitais, nesse contexto, transcende a função de simples intermediários de informação, exigindo que atuem como co-responsáveis na preservação de um ambiente digital seguro, inclusivo e democrático.

Além das questões jurídicas e éticas, a moderação de conteúdo envolve impactos sociais significativos. A circulação de discursos de ódio e desinformação afeta diretamente a coesão social, a confiança nas instituições e o respeito à diversidade. Ao mesmo tempo, a moderação inadequada ou insuficiente pode contribuir para a naturalização de preconceitos, marginalização de grupos vulneráveis e erosão do debate público. Assim, políticas eficazes de moderação são essenciais para a promoção de uma cultura digital que valorize a dignidade humana, a igualdade e o pluralismo, sem comprometer a liberdade de expressão e a circulação de ideias.

A convergência entre direito, ética e tecnologia indica que a regulação das plataformas digitais não é um processo isolado, mas exige a participação ativa de múltiplos atores: legisladores, órgãos reguladores, sociedade civil, empresas de tecnologia e usuários. A implementação de boas práticas de moderação, aliada a mecanismos legais claros e à educação

digital, constitui um caminho para enfrentar os desafios da era digital, protegendo direitos fundamentais sem inviabilizar a inovação e a liberdade de expressão. Nesse contexto, a moderação de conteúdo surge como um instrumento de cogovernança digital, capaz de equilibrar interesses conflitantes e garantir que a internet continue sendo um espaço de circulação de ideias, aprendizado, interação social e construção democrática.

CANCELAMENTO DIGITAL E SEUS EFEITOS

O cancelamento digital é um fenômeno contemporâneo que emergiu em paralelo ao crescimento das redes sociais e à intensificação da comunicação mediada por plataformas digitais. Trata-se de uma prática caracterizada pela retirada de apoio, boicote ou condenação social dirigida a indivíduos, organizações ou instituições em virtude de ações, declarações ou comportamentos percebidos como ofensivos, inadequados ou contrários a normas sociais e éticas vigentes. Essa dinâmica não se restringe a manifestações pontuais de desaprovação, mas constitui um mecanismo coletivo de responsabilização que, por vezes, adquire proporções significativas, transcendendo a esfera individual e gerando impactos profundos na vida social e digital.

O fenômeno do cancelamento digital insere-se em um contexto social mais amplo, permeado por demandas por justiça, equidade e responsabilização. Historicamente, a busca por responsabilização social está vinculada a movimentos por direitos civis, igualdade e reconhecimento de vulnerabilidades de grupos marginalizados. No ambiente digital, entretanto, essa busca adquire características peculiares, dado o caráter instantâneo da comunicação, a amplitude da disseminação de informações e a amplificação proporcionada por algoritmos que priorizam conteúdos de maior engajamento, independentemente de sua veracidade ou proporcionalidade. Assim, o cancelamento digital revela-se como um mecanismo de mobilização coletiva, capaz de pressionar indivíduos e instituições a responderem por suas ações, ao mesmo tempo em que levanta questões relativas à proporcionalidade e à justiça aplicada.

Do ponto de vista social, o cancelamento digital apresenta efeitos multifacetados. A prática pode atuar como instrumento de responsabilização, incentivando a reflexão sobre comportamentos considerados inadequados e promovendo mudanças sociais relevantes. Entretanto, a mesma prática pode se tornar excessiva e desproporcional, extrapolando o campo da crítica construtiva e configurando um mecanismo de ostracismo, silenciamento e punição pública. O impacto psicológico sobre aqueles que são alvo do cancelamento digital é

significativo, podendo gerar estresse, ansiedade, isolamento social e comprometimento da saúde mental. Além disso, a visibilidade e a rapidez das interações digitais potencializam os efeitos do cancelamento, tornando-o mais intenso e de difícil mitigação.

O debate acadêmico sobre o cancelamento digital destaca que a prática envolve uma tensão inerente entre liberdade de expressão e responsabilização. Enquanto a crítica social é um componente legítimo do debate democrático, o cancelamento digital pode transitar para uma forma de censura informal, na qual a condenação coletiva substitui procedimentos formais de avaliação ou sanção. Nesse sentido, torna-se imperativo analisar os critérios que distinguem críticas legítimas de ataques desproporcionais, considerando o impacto sobre a dignidade, os direitos fundamentais e a integridade pessoal. A responsabilidade digital não é apenas uma questão ética, mas também jurídica, na medida em que envolve o equilíbrio entre o direito à liberdade de expressão e a proteção contra violação de direitos, difamação ou assédio.

As plataformas digitais desempenham um papel central nesse fenômeno, na medida em que funcionam como mediadoras das interações e amplificadoras de conteúdos. A responsabilidade dessas plataformas na moderação de conteúdos envolve desafios complexos, pois é necessário conciliar o dever de prevenir danos com a preservação da liberdade de expressão. Algoritmos de moderação, políticas internas e mecanismos de denúncia devem ser desenhados de forma transparente, proporcional e fundamentada em critérios objetivos, de modo a evitar decisões arbitrárias que possam gerar efeitos punitivos desnecessários ou injustificados. A literatura especializada indica que a efetividade da moderação depende não apenas de medidas técnicas, mas também de práticas éticas e regulatórias que orientem a atuação das plataformas e assegurem mecanismos de contestação e revisão.

No campo jurídico, o cancelamento digital suscita reflexões acerca da responsabilidade civil e da proteção de direitos individuais. O fenômeno levanta questões sobre reputação, honra, imagem e liberdade de expressão, exigindo análise criteriosa para definir os limites entre críticas admissíveis e linchamento virtual. A jurisprudência e a legislação brasileiras ainda estão em processo de consolidação de parâmetros claros sobre essas questões, mas observa-se uma tendência de maior atenção à proteção de direitos fundamentais frente às dinâmicas digitais. Nesse contexto, o cancelamento digital desafia a concepção tradicional de responsabilização, demandando instrumentos que articulem prevenção, mediação e reparação de danos, tanto no nível individual quanto coletivo.

Do ponto de vista social e cultural, o cancelamento digital reflete tensões contemporâneas relacionadas a valores, normas e percepções de justiça. Ele atua como um indicador da exigência de responsabilização social, revelando sensibilidades sobre

comportamentos percebidos como inadequados e reforçando padrões coletivos de conduta. No entanto, também evidencia fragilidades na comunicação digital, como a polarização, a amplificação de julgamentos precipitados e a dificuldade de diálogo construtivo. O fenômeno demonstra a necessidade de desenvolver competências digitais, educação para a cidadania e práticas de empatia, de modo a equilibrar a responsabilização com a manutenção de um ambiente de interação saudável e inclusivo.

A análise acadêmica do cancelamento digital enfatiza ainda sua dimensão ética. A prática coloca em evidência a importância da reflexão crítica sobre os efeitos das ações coletivas no espaço digital e sobre a responsabilidade individual diante da propagação de julgamentos e críticas. A ética digital, nesse sentido, envolve a ponderação entre a busca por justiça e a preservação da dignidade humana, estabelecendo limites para que a mobilização social não se transforme em mecanismo de opressão ou exclusão. O equilíbrio entre responsabilização e empatia é, portanto, central para compreender a complexidade do fenômeno e orientar políticas e práticas sociais mais justas.

Além disso, o cancelamento digital apresenta efeitos estruturais sobre a sociedade, influenciando a forma como os indivíduos percebem e exercem a liberdade de expressão. A ameaça de reprovação massiva pode induzir autocensura, restringindo o debate público e a circulação de ideias, com repercussões diretas sobre a pluralidade e a diversidade de opiniões. Por outro lado, a existência de mecanismos de responsabilização contribui para reforçar normas sociais, estimular comportamentos responsáveis e promover maior conscientização sobre os impactos de ações individuais no espaço coletivo. Assim, o cancelamento digital emerge como um fenômeno complexo, no qual convergem dimensões psicológicas, sociais, éticas e jurídicas.

Em síntese, o cancelamento digital constitui um fenômeno multifacetado que demanda abordagem interdisciplinar e reflexão crítica. A análise científica do fenômeno evidencia a necessidade de compreender suas raízes sociais, implicações psicológicas, repercussões jurídicas e desafios éticos. A construção de respostas adequadas requer não apenas mecanismos de moderação e regulação, mas também a promoção de cultura digital, educação para a cidadania e desenvolvimento de práticas de empatia e diálogo. A compreensão do cancelamento digital como fenômeno social e jurídico é fundamental para fomentar uma sociedade digital mais equilibrada, na qual a responsabilização coexiste com a proteção de direitos e o respeito à dignidade humana, garantindo um espaço de debate plural, seguro e inclusivo.

DESAFIOS DE MODERAÇÃO DE CONTEÚDO EM PLATAFORMAS DIGITAIS

A moderação de conteúdo em plataformas digitais constitui um dos desafios mais complexos da sociedade contemporânea, especialmente considerando a escala e a velocidade com que informações são compartilhadas nas redes sociais. A tarefa de regular postagens, interações e publicações envolve múltiplas dimensões, que vão desde aspectos técnicos e operacionais até questões éticas, jurídicas e sociais. A natureza massiva e dinâmica dessas plataformas exige estratégias que permitam conciliar a liberdade de expressão com a proteção de indivíduos e grupos contra discursos de ódio, assédio e desinformação.

A prática da moderação envolve decisões contínuas sobre a aceitabilidade de conteúdos, considerando a subjetividade inerente à comunicação humana. O que é percebido como inofensivo por determinado grupo social pode ser interpretado como ofensivo ou prejudicial por outro, demonstrando a complexidade cultural e contextual que permeia a avaliação de conteúdos digitais. Nesse sentido, a moderação não se limita a uma operação técnica de remoção de informações, mas requer compreensão de fatores socioculturais, sensibilidade ética e critérios claros de justiça e proporcionalidade.

Os algoritmos têm sido empregados como ferramentas essenciais na moderação de conteúdo, permitindo a identificação e a classificação de postagens com base em padrões previamente definidos. Entretanto, a tecnologia apresenta limitações significativas na interpretação de nuances linguísticas, contextos implícitos e interações subjetivas. Assim, apesar de sua eficiência em processar grandes volumes de dados, os sistemas automatizados não substituem a necessidade de intervenção humana, especialmente em casos que demandam julgamento contextualizado e sensível às particularidades de diferentes comunidades digitais.

A escassez de recursos humanos qualificados para exercer a moderação é outro fator que agrava a complexidade dessa atividade. Moderadores são responsáveis por analisar conteúdos potencialmente prejudiciais, tomar decisões rápidas e lidar com consequências que se estendem além do ambiente digital, afetando a percepção, a reputação e a integridade de usuários. A natureza repetitiva e intensa da atividade, aliada à exposição constante a conteúdos sensíveis, torna o trabalho emocionalmente desgastante e exige medidas de apoio, capacitação e supervisão adequadas.

A moderação de conteúdo apresenta um dilema central entre a proteção dos usuários e a preservação da liberdade de expressão. Por um lado, a regulação e a remoção de conteúdos nocivos são necessárias para garantir um ambiente seguro, livre de discriminação, assédio e

incitação à violência. Por outro lado, a aplicação excessiva de medidas de controle pode restringir vozes legítimas, dificultar o debate público e gerar formas indiretas de censura. A subjetividade envolvida na interpretação de conteúdos aumenta a dificuldade de estabelecer critérios uniformes e imparciais, requerendo políticas claras, transparentes e fundamentadas em princípios éticos, sociais e legais.

As políticas internas das plataformas digitais refletem, em grande medida, os valores das comunidades que elas servem, bem como as pressões externas, incluindo demandas sociais, regulamentações e expectativas legais. Essas diretrizes, frequentemente revisadas em resposta a eventos específicos ou mudanças culturais, buscam equilibrar interesses diversos e concorrentes, como segurança, diversidade, inclusão e liberdade de expressão. Entretanto, a eficácia dessas políticas depende não apenas da definição de regras, mas também de sua aplicação consistente, da capacidade de revisão e do monitoramento contínuo de impactos sobre os usuários.

O dilema entre liberdade de expressão e proteção nas plataformas digitais também envolve dimensões éticas e sociais. A moderação deve considerar o impacto das decisões sobre grupos vulneráveis e a responsabilidade coletiva de preservar a dignidade e os direitos individuais. Além disso, a prática de moderação influencia a dinâmica do debate público, podendo promover um ambiente de diálogo seguro ou, inversamente, induzir autocensura, polarização e fragmentação social. A avaliação das consequências das decisões de moderação é, portanto, essencial para garantir que o equilíbrio entre proteção e expressão seja sustentável e equitativo.

A complexidade da moderação de conteúdo também está associada à diversidade cultural, à variação de normas sociais e à interpretação subjetiva de linguagem, símbolos e contextos. Um conteúdo considerado aceitável em determinado contexto pode ser ofensivo em outro, exigindo sensibilidade intercultural e critérios de proporcionalidade que transcendam decisões automatizadas. Nesse sentido, a moderação não deve ser entendida como mera aplicação de regras, mas como um processo reflexivo que envolve análise contextual, ponderação de valores e mitigação de danos potenciais.

A literatura científica e jurídica enfatiza que a moderação de conteúdo é um espaço de cogovernança digital, no qual convergem interesses da sociedade, da legislação e da ética empresarial. A atuação das plataformas deve articular prevenção de danos, mecanismos de denúncia, transparência algorítmica e revisão humana, de modo a equilibrar proteção, liberdade e responsabilidade. Além disso, a governança digital exige a construção de políticas que

promovam diálogo construtivo, respeitem direitos fundamentais e fomentem a inclusão social, sem sacrificar a diversidade de opiniões e a pluralidade democrática.

A análise do fenômeno revela que a moderação de conteúdo envolve riscos e desafios inerentes à digitalização das interações sociais. A dificuldade de interpretar nuances, a subjetividade da comunicação, a pressão por decisões rápidas e a necessidade de conciliar interesses divergentes tornam essa tarefa intrinsecamente complexa. Ao mesmo tempo, a moderação representa uma oportunidade de aprimorar práticas digitais, fortalecer a cidadania, proteger direitos individuais e coletivos e promover ambientes online mais seguros, inclusivos e responsáveis.

Em síntese, a moderação de conteúdo em plataformas digitais é um fenômeno multidimensional que exige abordagem interdisciplinar, combinando tecnologia, direito, ética e sociologia. A construção de políticas eficazes depende da compreensão das nuances culturais, da implementação de mecanismos de responsabilidade e transparência e do equilíbrio entre liberdade de expressão e proteção. A reflexão sobre os desafios da moderação contribui para o desenvolvimento de uma governança digital que valorize a dignidade humana, o diálogo plural e a convivência ética, constituindo-se em elemento central para a sustentabilidade de ambientes digitais saudáveis e democráticos.

CONSIDERAÇÕES FINAIS

A análise desenvolvida evidencia que a liberdade de expressão, embora constituindo direito fundamental e condição de possibilidade para a democracia, não possui caráter absoluto. A Constituição Federal de 1988 a consagra de forma ampla, mas em permanente diálogo com outros valores igualmente fundantes da ordem constitucional, como a dignidade da pessoa humana, a igualdade e a vedação à discriminação. A liberdade de expressão, portanto, deve ser interpretada de maneira contextualizada, reconhecendo suas limitações diante de manifestações que possam comprometer outros direitos e interesses coletivos.

O fenômeno do cancelamento digital e os desafios de moderação de conteúdo demonstram que a esfera digital é um espaço no qual esses limites se tornam particularmente visíveis. O cancelamento digital, ao operar como mecanismo de responsabilização social, evidencia o impacto das práticas coletivas sobre a vida de indivíduos e grupos, destacando que a expressão, quando acompanhada de reação imediata e massiva, pode gerar consequências

desproporcionais. A experiência contemporânea demonstra que, embora a busca por justiça social seja legítima, a reação exacerbada às falhas ou opiniões controversas de indivíduos pode comprometer a própria função democrática do debate público.

O discurso de ódio, em suas diversas manifestações, representa um limite claro à liberdade de expressão. Sua capacidade de promover exclusão, violência simbólica e real, bem como perpetuar desigualdades históricas, o coloca fora da proteção conferida a opiniões divergentes ou críticas construtivas. A jurisprudência do Supremo Tribunal Federal, como exemplificado pelo Caso Ellwanger (HC 82.424/RS), reforça que manifestações racistas, misóginas, homofóbicas ou xenofóbicas não encontram respaldo constitucional, reafirmando a necessidade de um equilíbrio entre liberdade de expressão e respeito aos direitos fundamentais de terceiros.

No ambiente digital, a complexidade se intensifica. A velocidade de difusão das informações, o alcance potencialmente global das publicações e a possibilidade de anonimato favorecem a propagação de conteúdos nocivos, dificultando a aplicação de mecanismos de responsabilização. Nesse cenário, as plataformas digitais emergem como atores centrais na governança do espaço virtual. O Marco Civil da Internet (Lei nº 12.965/2014) instituiu um modelo de responsabilização condicional à ordem judicial, mas decisões recentes do STF (ADIs 6.949, 6.960, 7.030 e ADPF 403) reconheceram a necessidade de exceções em casos de conteúdos manifestamente ilícitos, reposicionando as plataformas como agentes ativos de prevenção, denúncia e moderação.

A moderação de conteúdo, portanto, assume um papel estratégico na preservação da convivência democrática e na mitigação dos efeitos negativos do cancelamento digital. Trata-se de um processo complexo, que combina decisões humanas e algoritmos tecnológicos, considerando aspectos culturais, éticos e jurídicos. A subjetividade inerente à interpretação de postagens, memes e mensagens digitais evidencia que a tecnologia, isoladamente, não é capaz de compreender todas as nuances da comunicação humana. Assim, a atuação responsável das plataformas, aliada à capacitação de moderadores e à transparência nos critérios adotados, torna-se essencial para a eficácia de qualquer política de governança digital.

O desafio contemporâneo é construir um equilíbrio dinâmico: preservar a liberdade de expressão como espaço de crítica, diálogo e criatividade, enquanto se combatem discursos de ódio e práticas de cancelamento desproporcionais que ameaçam a coesão social e a democracia. A resposta não se limita à normatização legal, sendo igualmente necessária uma abordagem ética e educativa, que fomente a cultura digital de respeito, empatia e responsabilidade compartilhada entre usuários, plataformas e legisladores.

Conclui-se, portanto, que a efetividade da liberdade de expressão no século XXI depende da convergência de três dimensões complementares: (i) a existência de um marco jurídico claro, proporcional e capaz de orientar condutas e responsabilidades; (ii) a atuação proativa, transparente e ética das plataformas digitais, garantindo mecanismos de prevenção, moderação e responsabilização; e (iii) a conscientização dos cidadãos quanto ao uso responsável, ético e crítico da palavra no ambiente digital. A conjugação dessas dimensões permite não apenas que a liberdade de expressão seja exercida como direito individual, mas também que seja consolidada como valor coletivo indispensável à manutenção da vida democrática, da diversidade e da dignidade humana.

Em síntese, o cancelamento digital e os desafios de moderação de conteúdo não configuram apenas problemas técnicos ou jurídicos, mas representam questões éticas, sociais e políticas de grande relevância. Garantir a liberdade de expressão, sem negligenciar a proteção contra abusos e discriminações, é uma tarefa contínua e complexa, que exige reflexão crítica, políticas equilibradas e compromisso coletivo. A construção de ambientes digitais seguros, plurais e democráticos depende da integração entre direito, ética, educação digital e governança responsável, assegurando que a expressão humana seja instrumento de diálogo, aprendizagem e fortalecimento da cidadania, e não de exclusão ou silenciamento.

REFERÊNCIAS

- ALBUQUERQUE, Alexandre Bento Bernardes de; PINTO, Anderson de Sousa; SOUSA, Andreza Alexandra Soares; RODRIGUES, Frank Aguiar; LUZ, Júlio César de Moura; RIOS, Tácita Pereira. Liberdade de expressão e o discurso de ódio nas redes. Piauí: FAM, 2025.
- BARROSO, Luís Roberto. Liberdade de expressão, democracia e Constituição. 2. ed. São Paulo: Fórum, 2020.
- BOBBIO, Norberto. A era dos direitos. Rio de Janeiro: Elsevier, 2004.
- BRASIL. Constituição da República Federativa do Brasil de 1988. Brasília: Senado Federal, 1988.
- BRASIL. Lei nº 7.716, de 5 de janeiro de 1989. Define os crimes resultantes de preconceito de raça ou de cor. Diário Oficial da União, Brasília, 6 jan. 1989.
- BRASIL. Lei nº 12.965, de 23 de abril de 2014. Marco Civil da Internet. Diário Oficial da União, Brasília, 24 abr. 2014.
- BRASIL. Lei nº 14.532, de 11 de janeiro de 2023. Altera a Lei nº 7.716/1989 para equiparar a injúria racial ao crime de racismo. Diário Oficial da União, Brasília, 12 jan. 2023.

BRASIL. Supremo Tribunal Federal. Habeas Corpus 82.424/RS. Rel. Min. Moreira Alves, Tribunal Pleno, j. 17 set. 2003.

BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade 6.949. Rel. Min. Dias Toffoli, Tribunal Pleno, j. 21 jun. 2023.

BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade 6.960. Rel. Min. Dias Toffoli, Tribunal Pleno, j. 21 jun. 2023.

BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade 7.030. Rel. Min. Dias Toffoli, Tribunal Pleno, j. 21 jun. 2023.

BRASIL. Supremo Tribunal Federal. Arguição de Descumprimento de Preceito Fundamental 403. Rel. Min. Edson Fachin, Tribunal Pleno, j. 21 jun. 2023.

CANOTILHO, José Joaquim Gomes. Direito Constitucional e Teoria da Constituição. 7. ed. Coimbra: Almedina, 2003.

ONU. Organização das Nações Unidas. Plano de Ação contra o Discurso de Ódio. Nova Iorque: ONU, 2019.

SARLET, Ingo Wolfgang; MARINONI, Luiz Guilherme; MITIDIERO, Daniel. Curso de Direito Constitucional. 4. ed. São Paulo: RT, 2020.

UNIÃO EUROPEIA. Digital Services Act. Bruxelas: Parlamento Europeu, 2022.