

João Paulo de Sousa Curvêlo
Maria Emilia Camargo



PROTEÇÃO AOS DIREITOS HUMANOS FRENTE AO DISCURSO DE ÓDIO NAS REDES SOCIAIS E À LIBERDADE DE EXPRESSÃO



Shift

SÃO PAULO | 2025

João Paulo de Sousa Curvêlo
Maria Emilia Camargo



PROTEÇÃO AOS DIREITOS HUMANOS FRENTE AO DISCURSO DE ÓDIO NAS REDES SOCIAIS E À LIBERDADE DE EXPRESSION



Shift

SÃO PAULO | 2025

1.^a edição

João Paulo de Sousa Curvêlo
Maria Emilia Camargo

**PROTEÇÃO AOS DIREITOS HUMANOS FRENTE AO DISCURSO DE
ÓDIO NAS REDES SOCIAIS E À LIBERDADE DE EXPRESSÃO**

ISBN 978-65-6054-241-9



João Paulo de Sousa Curvêlo
Maria Emilia Camargo

PROTEÇÃO AOS DIREITOS HUMANOS FRENTE AO DISCURSO DE
ÓDIO NAS REDES SOCIAIS E À LIBERDADE DE EXPRESSÃO

1.^a edição

SÃO PAULO
EDITORA ARCHÉ
2025

Copyright © dos autores e das autoras.

Todos os direitos garantidos. Este é um livro publicado em acesso aberto, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que sem fins comerciais e que o trabalho original seja corretamente citado. Este trabalho está licenciado com uma Licença *Creative Commons Internacional* (CC BY- NC 4.0).



**Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)**

C975p

Curvêlo, João Paulo de Sousa.

Proteção aos direitos humanos frente ao discurso de ódio nas redes sociais e à liberdade de expressão [livro eletrônico] / João Paulo de Sousa Curvêlo, Maria Emilia Camargo. – 1. ed. – São Paulo, SP : Editora Arché, 2025.
206 p.

Formato: PDF

Requisitos de sistema: Adobe Acrobat Reader

Modo de acesso: World Wide Web

Inclui bibliografia.

ISBN 978-65-6054-241-9

1. Liberdade de expressão – Brasil. 2. Discurso de ódio – Aspectos jurídicos – Brasil. 3. Direitos humanos – Brasil. 4. Redes sociais – Regulação. 5. Plataformas digitais – Responsabilidade civil. I. Camargo, Maria Emilia. II. Título.

CDD 323.44

Elaborado por Maurício Amormino Júnior – CRB6/2422

Revista REASE cancelada pela Editora Arché.

São Paulo- SP

Telefone: +55 55(11) 5107-0941

<https://periodicorease.pro.br>

contato@periodicorease.pro.br

1ª Edição- *Copyright*© 2025 dos autores.

Direito de edição reservado à Revista REASE.

O conteúdo de cada capítulo é de inteira e exclusiva responsabilidade do (s) seu(s) respectivo (s) autor (es).

As normas ortográficas, questões gramaticais, sistema de citações e referenciais bibliográficos são prerrogativas de cada autor (es).

Endereço: Av. Brigadeiro Faria de Lima n.º 1.384 — Jardim Paulistano.

CEP: 01452 002 — São Paulo — SP.

Tel.: 55(11) 5107-0941

<https://periodicorease.pro.br/rease>

contato@periodicorease.pro.br

Editora: Dra. Patrícia Ribeiro

Produção gráfica e direção de arte: Ana Cláudia Néri Bastos

Assistente de produção editorial e gráfica: Talita Tainá Pereira Batista, Cintia Milena Gonçalves Rolim

Projeto gráfico: Ana Cláudia Néri Bastos

Ilustrações: Ana Cláudia Néri Bastos, Talita Tainá Pereira Batista, Cintia Milena Gonçalves Rolim

Revisão: Ana Cláudia Néri Bastos e Talita Tainá Pereira Batista, Cintia Milena Gonçalves Rolim

Tratamento de imagens: Ana Cláudia Néri Bastos

EQUIPE DE EDITORES

EDITORA- CHEFE

Dra. Patrícia Ribeiro, Universidade de Coimbra- Portugal

CONSELHO EDITORIAL

Doutoranda Silvana Maria Aparecida Viana Santos- Facultad Interamericana de Ciências Sociais - FICS

Doutorando Alberto da Silva Franqueira-Facultad Interamericana de Ciencias Sociales (FICS)

Me. Ubiranilze Cunha Santos- Corporación Universitaria de Humanidades Y Ciências Sociales de Chile

Doutorando Allysson Barbosa Fernandes- Facultad Interamericana de Ciencias Sociales (FICS)

Doutor. Avaetê de Lunetta e Rodrigues Guerra- Universidad del Sol do Paraguai- PY

Me. Victorino Correia Kinhama- Instituto Superior Politécnico do Cuanza Sul-Angola

Me. Andrea Almeida Zamorano- SPSIG

Esp. Ana Cláudia N. Bastos- PUCRS

Dr. Alfredo Oliveira Neto, UERJ, RJ

PhD. Diogo Vianna, IEPA

Dr. José Fajardo- Fundação Getúlio Vargas

PhD. Jussara C. dos Santos, Universidade do Minho

Dra. María V. Albaronedo, Universidad Nacional del Comahue, Argentina

Dra. Uaiana Prates, Universidade de Lisboa, Portugal

Dr. José Benedito R. da Silva, UFSCar, SP

PhD. Pablo Guadarrama González, Universidad Central de Las Villas, Cuba

Dra. Maritza Montero, Universidad Central de Venezuela, Venezuela

Dra. Sandra Moitinho, Universidade de Aveiro-Portugal

Me. Eduardo José Santos, Universidade Federal do Ceará,

Dra. Maria do Socorro Bispo, Instituto Federal do Paraná, IFPR

Cristian Melo, MEC

Dra. Bartira B. Barros, Universidade de Aveiro-Portugal

Me. Roberto S. Maciel- UFBA

Dra. Francisne de Souza, Universidade de Aveiro-Portugal

Dr. Paulo de Andrada Bittencourt – MEC

PhD. Aparecida Ribeiro, UFG

Dra. Maria de Sandes Braga, UFTM

DECLARAÇÃO DOS AUTORES

Os autores se responsabilizam publicamente pelo conteúdo desta obra, garantindo que o mesmo é de autoria própria, assumindo integral responsabilidade diante de terceiros, quer de natureza moral ou patrimonial, em razão de seu conteúdo, declarando que o trabalho é original, livre de plágio acadêmico e que não infringe quaisquer direitos de propriedade intelectual de terceiros. Os autores declaram não haver qualquer interesse comercial ou irregularidade que comprometa a integridade desta obra.

DECLARAÇÃO DA EDITORA

A Editora Arché declara, para os devidos fins de direito, que: 1. A presente publicação constitui apenas transferência temporária dos direitos autorais, direito sobre a publicação, inclusive não constitui responsabilidade solidária na criação dos manuscritos publicados, nos termos previstos na Lei sobre direitos autorais (Lei 9610/98), no art. 184 do Código Penal e no art.º 927 do Código Civil; 2. Autoriza e incentiva os autores a assinarem contratos com repositórios institucionais, com fins exclusivos de divulgação da obra, desde que com o devido reconhecimento de autoria e edição e sem qualquer finalidade comercial; 3. Todos os e-book são *open access*, desta forma não os comercializa em seu site, sites parceiros, plataformas de *ecommerce*, ou qualquer outro meio virtual ou físico, portanto, está isenta de repasses de direitos autorais aos autores; 4. Não cede, comercializa ou autoriza a utilização dos nomes e e-mails dos autores, bem como nenhum outro dado dos mesmos, para qualquer finalidade que não o escopo da divulgação desta obra.

DEDICATÓRIA

Esta Dissertação de Mestrado é dedicada a memória do meu pai Américo e da minha mãe Cristina, aos quais devo a minha vida em todas suas dimensões, pois sempre me guiaram pelo bom caminho e graças ao carinho, ao cuidado e ao exemplo deles, consegui chegar até aqui e espero sinceramente poder honrá-los sempre!

AGRADECIMENTOS

Agradeço, em primeiro lugar, a Deus, por iluminar meu caminho e me dar forças durante toda esta jornada.

Expresso minha profunda gratidão à minha orientadora, a professora Maria Emília Camargo, pela sua dedicação, paciência e valiosa orientação. Sua atenção e prestatividade foram fundamentais para a concretização deste trabalho.

Agradeço aos professores Aprígio Teles Mascarenhas Neto, Mariane Camargo Priesnitz e Eloy Pereira Lemos Júnior, membros da banca examinadora. Agradeço não apenas pelos elogios, mas, sobretudo, pelas sugestões criteriosas que foram essenciais para o aperfeiçoamento final desta dissertação.

Meu reconhecimento se estende aos docentes do curso de Mestrado em Ciências Jurídicas da Veni Creator Christian University, por seu apoio, sabedoria e humanidade, que tornaram o aprendizado uma experiência rica e exitosa.

Por fim, sou grato a todos que, de alguma forma, fizeram parte da minha vida e me apoiaram nesta caminhada. Deixo um agradecimento especial aos colegas de mestrado, pelo companheirismo, carinho e convívio que me ajudaram a crescer e a superar os momentos desafiadores.

RESUMO

Este livro digital examina o dilema jurídico no Brasil de conciliar a liberdade de expressão com a necessidade de combater o discurso de ódio nas redes sociais. A análise é baseada nos princípios constitucionais e nos compromissos internacionais de direitos humanos do país. A pesquisa parte da hipótese de que a efetiva proteção aos direitos humanos demanda a construção de parâmetros jurídicos mais claros para limitar a liberdade de expressão quando esta se converte em discurso de ódio, exigindo normas específicas e maior responsabilização das plataformas digitais. Metodologicamente, adota-se uma abordagem qualitativa com enfoque dedutivo, utilizando métodos documental, bibliográfico, comparado e jurisprudencial, com análise crítica pautada na hermenêutica dos direitos fundamentais. Os resultados evidenciam que o discurso de ódio se configura como forma de violência simbólica e estrutural que marginaliza grupos vulneráveis e mina os alicerces da democracia plural. Constatou-se que a legislação brasileira, apesar de contar com instrumentos como a Lei Caó e o Marco Civil da Internet, carece de mecanismos específicos para lidar com a complexidade do fenômeno no ambiente digital. A jurisprudência do STF e da Corte Interamericana de Direitos Humanos tem evoluído no sentido de reconhecer que a repressão ao discurso de ódio, quando devidamente balizada, constitui condição essencial para a efetividade universal da própria liberdade de expressão. Conclui-se propondo um modelo regulatório equilibrado que contemple: educação digital crítica e alfabetização midiática; revisão do regime de responsabilidade das plataformas digitais; tipificação penal específica do discurso de ódio com base em critérios objetivos; parâmetros consistentes para a atuação judicial; e modelos de correção com participação da sociedade civil. O trabalho contribui para o avanço do debate jurídico sobre um tema de crescente relevância social, política e constitucional, oferecendo subsídios teóricos e práticos para a construção de um modelo que proteja efetivamente tanto a liberdade de expressão quanto a dignidade humana no ambiente digital.

Palavras-chave: Liberdade de expressão. Discurso de ódio. Redes sociais. Direitos fundamentais. Regulação de plataformas digitais.

ABSTRACT

This e-book examines the legal dilemma in Brazil of balancing freedom of expression with the need to combat hate speech on social media. The analysis is based on the country's constitutional principles and international human rights commitments. The research is based on the hypothesis that effective protection of human rights requires the construction of clearer legal parameters to limit freedom of expression when it becomes hate speech, requiring specific rules and greater accountability of digital platforms. Methodologically, a qualitative approach with a deductive focus is adopted, using documentary, bibliographical, comparative and jurisprudential methods, with a critical analysis based on the hermeneutics of fundamental rights. The results show that hate speech is a form of symbolic and structural violence that marginalizes vulnerable groups and undermines the foundations of plural democracy. It was found that Brazilian legislation, despite having instruments such as the Caó Law and the Internet Civil Rights Framework, lacks specific mechanisms to deal with the complexity of the phenomenon in the digital environment. The jurisprudence of the Brazilian Supreme Court and the Inter-American Court of Human Rights has evolved towards recognizing that the repression of hate speech, when properly guided, is an essential condition for the universal effectiveness of freedom of expression itself. The paper concludes by proposing a balanced regulatory model that includes: critical digital education and media literacy; review of the liability regime of digital platforms; specific criminal classification of hate speech based on objective criteria; consistent parameters for judicial action; and co-regulation models with the participation of civil society. The work contributes to the advancement of the legal debate on a topic of increasing social, political and constitutional relevance, offering theoretical and practical support for the construction of a model that effectively protects both freedom of expression and human dignity in the digital environment.

Keywords: Freedom of expression. Hate speech. Social media. Fundamental rights. Regulation of digital platforms.

"A tolerância ilimitada leva ao desaparecimento da tolerância. Se estendermos a tolerância ilimitada mesmo aos intolerantes, e se não estivermos preparados para defender a sociedade tolerante do assalto da intolerância, então, os tolerantes serão destruídos e a tolerância com eles.—Nessa formulação, não insinuo, por exemplo, que devemos sempre suprimir a expressão de filosofias intolerantes; desde que possamos combatê-las com argumentos racionais e mantê-las em xeque frente à opinião pública, suprimi-las seria, certamente, imprudente. Mas devemos-nos reservar o direito de suprimi-las, se necessário, mesmo que pela força; pode ser que eles não estejam preparados para nos encontrar nos níveis dos argumentos racionais, ao começar por criticar todos os argumentos e proibindo seus seguidores de ouvir argumentos racionais, porque são enganadores, e ensiná-los a responder aos argumentos com punhos ou pistolas. Devemos-nos, então, reservar, em nome da tolerância, o direito de não tolerar o intolerante. Devemos exigir que qualquer movimento que pregue a intolerância fique à margem da lei e que qualquer incitação à intolerância e perseguição seja considerada criminosa, da mesma forma que no caso de incitação ao homicídio, sequestro de crianças ou revivescência do tráfico de escravos". (Karl Popper)

SUMÁRIO

CAPÍTULO 01	15
INTRODUÇÃO	
CAPÍTULO 02	24
REFERENCIAL TEÓRICO	
CAPÍTULO 03	153
METODOLOGIA	
CAPÍTULO 04	163
RESULTADOS E DISCUSSÕES	
CONSIDERAÇÕES FINAIS	176
REFERÊNCIAS	184
ÍNDICE REMISSIVO	202

CAPÍTULO 01

INTRODUÇÃO

1 INTRODUÇÃO

O advento e a popularização das redes sociais e plataformas digitais nas últimas décadas transformaram profundamente as dinâmicas comunicacionais, políticas e sociais em escala global. O que inicialmente se apresentava como um promissor espaço de democratização do acesso à informação, de ampliação das possibilidades de expressão e de fortalecimento da participação cidadã no debate público, gradualmente revelou também sua face sombria: um ambiente propício à disseminação acelerada e massiva de discursos de ódio, desinformação e conteúdo que atentam contra a dignidade humana e os fundamentos do Estado Democrático de Direito. Esse fenômeno, longe de ser uma questão meramente tecnológica, configura-se como um complexo desafio jurídico, político e social que demanda respostas multidimensionais e interdisciplinares.

No contexto brasileiro, essa problemática adquire contornos ainda mais sensíveis e urgentes. Um país marcado por profundas desigualdades estruturais, por um histórico de violência e discriminação contra grupos minoritários e vulnerabilizados, e por uma democracia ainda em processo de consolidação após um longo período ditatorial, enfrenta o desafio de equilibrar a proteção robusta à liberdade de expressão – pilar fundamental de qualquer regime democrático e garantia constitucional expressa – com a necessidade igualmente imperiosa de coibir manifestações que promovam o ódio, a intolerância e a exclusão social. Como observa Sarlet (2021, p. 221), "a liberdade de expressão, embora essencial à democracia,

não pode servir de escudo para práticas que negam a própria dignidade humana e o pluralismo que caracteriza uma sociedade aberta".

A Constituição Federal de 1988, conhecida como "Constituição Cidadã", estabeleceu um compromisso inequívoco com a proteção da dignidade da pessoa humana, com a construção de uma sociedade livre, justa e solidária, e com a promoção do bem de todos, sem preconceitos ou discriminações de qualquer natureza (art. 1º, III e art. 3º, I e IV). Ao mesmo tempo, consagrou a liberdade de manifestação do pensamento (art. 5º, IV), a liberdade de expressão da atividade intelectual, artística, científica e de comunicação (art. 5º, IX) e a liberdade de informação jornalística (art. 220), como direitos fundamentais invioláveis. Esse aparente paradoxo – a necessidade de proteger simultaneamente a liberdade de expressão e a dignidade humana – constitui o cerne do debate jurídico contemporâneo sobre os limites do discurso no espaço público, especialmente no ambiente digital.

As redes sociais, com sua arquitetura algorítmica orientada primordialmente à maximização do engajamento e do tempo de permanência dos usuários, frequentemente amplificam conteúdos extremos, polarizadores e emocionalmente carregados, criando o que Sunstein (2017) denomina "câmaras de eco" e Pariser (2012) chama de "filtros-bolha" – ambientes digitais onde os indivíduos são expostos predominantemente a informações e opiniões que reforçam suas crenças preexistentes, com pouco ou nenhum contato com perspectivas divergentes. Esse fenômeno, aliado à velocidade de disseminação de conteúdos nas plataformas digitais e à possibilidade de anonimato ou

pseudonimato, cria condições particularmente favoráveis à proliferação do discurso de ódio online.

Segundo dados da pesquisa TIC Domicílios 2024, realizada pelo Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (CETIC.BR, 2024), 97% dos brasileiros com mais de 10 anos são usuários de internet, e destes, 81% utilizam redes sociais regularmente. Esse elevado grau de penetração das plataformas digitais na vida cotidiana dos cidadãos brasileiros amplifica tanto o potencial democrático quanto os riscos associados a esses espaços. Um levantamento realizado pela SaferNet Brasil (2023) registrou um aumento de 41% nas denúncias de discurso de ódio nas redes sociais em comparação ao ano anterior, com destaque para manifestações racistas (na forma de discriminação baseada na raça, a partir do falso entendimento de que haveria raças humanas superiores a outras) (DIAS, [2025]), xenofóbicas (hostilidade e ódio manifestado contra pessoas por elas serem estrangeiras ou por serem enxergadas como estrangeiras) (SILVA, [2024]), homofóbicas (atitudes e sentimentos negativos, como aversão, repugnância, medo e agressão, contra indivíduos que se identificam como lésbicas, gays, bissexuais, transexuais e travestis) (MENDES, [2025]) e misóginas (ódio ou aversão às mulheres, manifestando-se através de atitudes e comportamentos que as desvalorizam, discriminam e violentam) (CAMPOS, [2025]).

A relevância do tema se evidencia ainda mais quando observamos os impactos concretos que o discurso de ódio online produz na vida de indivíduos e grupos historicamente marginalizados. Estudos recentes, como o de Recuero e Soares (2021), demonstram correlações significativas

entre picos de discurso de ódio nas redes sociais e o aumento de casos de violência física contra minorias. Além disso, o fenômeno tem implicações diretas para a qualidade do debate democrático e para a própria estabilidade institucional, como ficou evidente em episódios recentes da história brasileira, nos quais narrativas de ódio e desinformação contribuíram para a polarização extrema e para tentativas de deslegitimação de instituições democráticas (RECUERO; SOARES, 2021).

No plano jurídico, o Brasil enfrenta o desafio de atualizar seu arcabouço normativo para responder adequadamente a essa nova realidade. O Marco Civil da Internet (Lei nº 12.965/2014), que foi criado em um contexto de crescente popularização da internet no Brasil, com a necessidade de regulamentar minimamente o uso da rede e garantir direitos aos usuários, apesar de ter surgido a partir de debates e consultas públicas que envolveram a sociedade civil, o governo e o setor privado, embora tenha representado um avanço significativo ao estabelecer princípios, garantias, direitos e deveres para o uso da internet no país, tem se mostrado insuficiente diante da complexidade e da dinamicidade dos problemas relacionados ao discurso de ódio online. Em particular, o regime de responsabilidade civil dos provedores de aplicações de internet previsto em seu artigo 19 – que condiciona a responsabilização à desobediência de ordem judicial específica – tem sido objeto de intenso debate acadêmico, legislativo e jurisprudencial.

Paralelamente, o Supremo Tribunal Federal (STF) tem sido chamado a se manifestar sobre casos emblemáticos envolvendo o conflito

entre liberdade de expressão e outros direitos fundamentais, como no Habeas Corpus (HC) 82.424/RS (Caso Ellwanger), na Arguição de Descumprimento de Preceito Fundamental (ADPF) 130/DF (Lei de Imprensa), na Ação Direta de Inconstitucionalidade por Omissão (ADO) 26/DF (criminalização da LGBTfobia) e no Inquérito 4.781/DF (Inquérito das Fake News). Essas decisões, embora fundamentais para a construção de parâmetros interpretativos, ainda não consolidaram um entendimento uniforme e abrangente sobre os limites da liberdade de expressão no ambiente digital e sobre a caracterização jurídica do discurso de ódio.

No cenário internacional, diversos países e blocos regionais têm adotado abordagens regulatórias mais assertivas em relação às plataformas digitais e ao discurso de ódio online. A União Europeia, com o Digital Services Act (DSA) e o Code of Conduct on Countering Illegal Hate Speech Online, e a Alemanha, com o Netzwerkdurchsetzungsgesetz (NetzDG), são exemplos de iniciativas que buscam estabelecer obrigações mais claras para as plataformas em termos de moderação de conteúdo e transparência algorítmica. Essas experiências internacionais oferecem importantes referências para o debate brasileiro, embora seja fundamental considerar as especificidades do contexto nacional na formulação de respostas regulatórias.

Diante desse cenário complexo e desafiador, esta dissertação se propõe a investigar o seguinte problema de pesquisa: Como o direito brasileiro pode equilibrar a liberdade de expressão com o combate ao discurso de ódio nas redes sociais, respeitando os princípios constitucionais e os compromissos internacionais de proteção aos direitos

humanos?

A hipótese que orienta esta investigação é a de que a proteção aos direitos humanos no Brasil demanda a construção de parâmetros jurídicos mais claros para a limitação da liberdade de expressão quando esta se converte em discurso de ódio, especialmente nas redes sociais, exigindo não só normas mais específicas, mas também maior responsabilização das plataformas digitais. Parte-se do pressuposto de que é possível – e necessário – desenvolver um modelo normativo, institucional e jurisprudencial que concilie a ampla liberdade de expressão garantida pela Constituição Federal de 1988 com a vedação a manifestações que, sob o manto dessa liberdade, promovam a exclusão, a discriminação e a negação da dignidade humana.

Quanto à estrutura, a dissertação está organizada em Introdução, Referencial Teórico, Metodologia, Resultados e Discussões e Considerações Finais.

Destaco o conteúdo do Referencial Teórico, que aborda o conceito de direitos humanos, princípio da dignidade da pessoa humana e direito à não discriminação, a evolução histórica e filosófica da liberdade de expressão, sua configuração no sistema constitucional brasileiro, as especificidades da liberdade de informação, comunicação e imprensa, e os limites constitucionais e cláusulas de vedação à liberdade de expressão. Traz à lume também o estudo do discurso de ódio, explorando seu conceito jurídico e filosófico, as diferenças entre discurso ofensivo, discurso de ódio e discurso discriminatório, a dignidade da pessoa humana como limite à manifestação do pensamento, a tipificação penal brasileira, e os parâmetros

dos sistemas interamericano e europeu de direitos humanos. Segue com a análise do papel das redes sociais como arenas públicas contemporâneas, os desafios relacionados aos algoritmos, à viralização e à responsabilidade das plataformas, o Marco Civil da Internet, a Lei Geral de Proteção de Dados (LGPD) e a liberdade de expressão, os dilemas jurídicos da moderação de conteúdo, e casos emblemáticos no Brasil. Por fim, o Referencial Teórico discute a proteção de grupos vulneráveis como dever do Estado, o discurso de ódio como forma de violência simbólica e estrutural, a jurisprudência do STF e da Corte Interamericana, propostas para um modelo normativo e jurisprudencial equilibrado, e o papel da educação digital, da regulação democrática e da cultura de paz no enfrentamento ao discurso de ódio.

Ao longo deste percurso investigativo, busca-se contribuir para o avanço do debate jurídico sobre um tema de crescente relevância social, política e constitucional, oferecendo subsídios teóricos e práticos para a construção de um modelo regulatório que proteja efetivamente tanto a liberdade de expressão quanto a dignidade humana no ambiente digital.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Analisar criticamente os fundamentos jurídicos, os parâmetros normativos e os desafios práticos relacionados à regulação do discurso de ódio nas plataformas digitais no Brasil, considerando as tensões entre a liberdade de expressão e a dignidade humana.

1.1.2 Objetivos Específicos

- a) Examinar a evolução histórica e filosófica do conceito de liberdade de expressão e sua configuração no sistema constitucional brasileiro;
- b) Analisar o conceito jurídico de discurso de ódio, suas características e suas distinções em relação a outras formas de manifestação controversa;
- c) Investigar o papel das redes sociais como arenas públicas contemporâneas e os desafios específicos que suas características tecnológicas e econômicas impõem à regulação do discurso;
- d) Avaliar criticamente o atual marco normativo brasileiro e as propostas legislativas em discussão sobre o tema;
- e) Analisar a jurisprudência do STF e da Corte Interamericana de Direitos Humanos sobre liberdade de expressão e discurso de ódio;
- f) Propor diretrizes para um modelo normativo e jurisprudencial equilibrado, que concilie a proteção à liberdade de expressão com o combate efetivo ao discurso de ódio.

CAPÍTULO 02

REFERENCIAL TEÓRICO

2 REFERENCIAL TEÓRICO

2.1 DIREITOS HUMANOS

Os direitos humanos constituem o alicerce fundamental sobre o qual se erguem as sociedades democráticas e justas. Representam um conjunto intrínseco de prerrogativas e faculdades que são inerentes a cada ser humano, desde o seu nascimento, independentemente de quaisquer distinções como raça, cor, sexo, idioma, religião, opinião política ou de outra natureza, origem nacional ou social, riqueza, nascimento ou qualquer outra condição. Essa universalidade, inalienabilidade, indivisibilidade e interdependência são os pilares que sustentam a dignidade e a liberdade de cada indivíduo, conferindo-lhes um valor intrínseco e inestimável (BOBBIO, 2004). A compreensão aprofundada de sua natureza, evolução histórica e dos princípios que os regem é não apenas fundamental, mas indispensável para a análise de qualquer fenômeno que envolva a proteção da pessoa humana, como o complexo e multifacetado desafio do combate ao discurso de ódio no ambiente digital (SARLET, 2021).

Neste capítulo, exploraremos os aspectos históricos que moldaram a concepção moderna dos direitos humanos, as diferentes dimensões que refletem sua progressiva ampliação e os princípios basilares que os sustentam, com especial atenção à dignidade da pessoa humana e ao direito à não-discriminação. A relevância desses conceitos para a temática da dissertação reside na sua capacidade de fornecer o arcabouço teórico e axiológico necessário para compreender os limites da liberdade de expressão e a imperatividade da proteção contra manifestações que

atentem contra a dignidade e a igualdade de todos os indivíduos (PIOVESAN, 2021).

2.1.1 Aspectos históricos

A trajetória dos direitos humanos é uma narrativa rica e complexa, pontuada por avanços e retrocessos, mas sempre impulsionada pela busca incessante por justiça e reconhecimento da dignidade inerente a cada ser humano. Embora a formalização e a universalização desses direitos sejam características da modernidade, a semente da ideia de direitos inatos ao indivíduo pode ser rastreada até as civilizações antigas. Códigos como o de Hamurabi (século XVIII a.C.) e as filosofias gregas e romanas já continham preceitos de justiça e moralidade que, de alguma forma, vislumbravam a proteção de certos interesses individuais, ainda que restritos a grupos específicos da sociedade (COMPARATO, 2003).

No Ocidente, marcos históricos significativos começaram a delinear o caminho para a concepção contemporânea dos direitos humanos. A Magna Carta Libertatum (1215), imposta pelos barões ingleses ao Rei João Sem Terra, é frequentemente citada como um dos primeiros documentos a limitar o poder arbitrário do monarca e a garantir certas liberdades aos súditos, como o direito a um julgamento justo e a proteção contra a prisão arbitrária. Embora seu alcance fosse restrito à nobreza, ela estabeleceu um precedente necessário para a ideia de que o poder governamental não é absoluto e está sujeito a leis (INGLATERRA, 1215). Séculos mais tarde, a Revolução Gloriosa na Inglaterra (1688) e a subsequente Declaração de Direitos (Bill of Rights) de 1689 consolidaram

o poder do Parlamento sobre a Coroa e reafirmaram direitos civis e políticos, como a liberdade de expressão no Parlamento e o direito de petição. Esses eventos foram fundamentais para o desenvolvimento do constitucionalismo e para a afirmação de que os direitos dos cidadãos devem ser protegidos por lei (INGLATERRA, 1689).

O século XVIII testemunhou o florescimento do Iluminismo, um movimento filosófico que enfatizou a razão, a ciência e os direitos individuais. Essas ideias, que defendiam a liberdade, a igualdade e a separação de poderes, influenciaram profundamente as revoluções que se seguiram. A Revolução Americana, culminando na Declaração de Independência dos Estados Unidos (1776), proclamou que todos os homens são criados iguais e dotados de certos direitos inalienáveis, como a vida, a liberdade e a busca da felicidade. Este documento, embora ainda não universal em sua aplicação (mantendo a escravidão, por exemplo), foi um passo ousado na afirmação da soberania popular e dos direitos individuais (ESTADOS UNIDOS, 1776). Contudo, foi a Revolução Francesa (1789) que, com sua Declaração dos Direitos do Homem e do Cidadão, estabeleceu um marco ainda mais abrangente. Inspirada nos ideais de liberdade, igualdade e fraternidade, a Declaração francesa proclamou direitos universais a todos os cidadãos, independentemente de sua posição social, e serviu de inspiração para inúmeras constituições e movimentos por direitos em todo o mundo. No entanto, mesmo com esses avanços, o alcance dos direitos ainda era limitado, excluindo mulheres, escravos e outras minorias, o que demonstra a natureza progressiva e contínua da luta pelos direitos humanos (FRANÇA, 1789).

O verdadeiro divisor de águas para a universalização dos direitos humanos ocorreu no século XX, em resposta às atrocidades e violações massivas de direitos cometidas durante as duas Guerras Mundiais, especialmente o Holocausto. A necessidade de criar um sistema de proteção global tornou-se evidente. A fundação da Organização das Nações Unidas (ONU) em 1945 foi um passo decisivo nesse sentido. O ápice desse movimento foi a adoção da Declaração Universal dos Direitos Humanos (DUDH) em 10 de dezembro de 1948. A DUDH, ao proclamar que "todos os seres humanos nascem livres e iguais em dignidade e em direitos", estabeleceu um ideal comum a ser alcançado por todos os povos e nações. Ela não é um tratado vinculante por si só, mas sua influência moral e política é imensa, servindo de base para a elaboração de inúmeros tratados e convenções internacionais de direitos humanos, como os Pactos Internacionais de Direitos Civis e Políticos e de Direitos Econômicos, Sociais e Culturais, que conferiram força jurídica a muitos dos direitos ali enunciados (ONU, 1948). Desde então, o sistema internacional de direitos humanos tem se expandido, com a criação de órgãos de monitoramento, tribunais regionais e mecanismos de denúncia, consolidando a ideia de que os direitos humanos são uma preocupação legítima da comunidade internacional e não apenas uma questão interna dos Estados. A história dos direitos humanos, portanto, é uma história de constante evolução, reflexo das transformações sociais, políticas e tecnológicas, e da persistente busca pela efetivação da dignidade humana para todos.

2.1.2. Breves considerações acerca das dimensões de direitos humanos

A evolução histórica dos direitos humanos é frequentemente analisada através da lente das "dimensões" ou "gerações" de direitos. Essa categorização, embora didática e útil para compreender a progressiva ampliação do escopo dos direitos, não implica uma hierarquia ou que uma dimensão anule ou substitua a outra. Pelo contrário, os direitos humanos são, em sua essência, indivisíveis, interdependentes e inter-relacionados, o que significa que a plena realização de um direito muitas vezes depende da realização de outros (ALEXY, 2015).

Primeira Dimensão (Direitos Cívicos e Políticos): Estes direitos emergiram com as revoluções liberais do século XVIII, notadamente a Americana e a Francesa, e são frequentemente associados à ideia de liberdade individual e à proteção do indivíduo contra o poder arbitrário do Estado. São considerados direitos de "não-intervenção" ou "liberdades negativas", pois exigem do Estado uma abstenção, ou seja, que não interfira na esfera de autonomia do indivíduo. Incluem uma vasta gama de prerrogativas, como o direito à vida, à liberdade pessoal, à propriedade, à segurança, à liberdade de pensamento, de consciência e de religião, à liberdade de expressão, de associação e de reunião pacífica. Além disso, englobam os direitos políticos, como o direito de votar e ser votado, e o direito a um julgamento justo e imparcial. Autores como Norberto Bobbio (2004) destacam que esses direitos foram determinantes para a consolidação do Estado de Direito e para a limitação do absolutismo (BOBBIO, 2004).

Segunda Dimensão (Direitos Econômicos, Sociais e Culturais -

DESC): Desenvolvidos a partir do século XIX e consolidados no século XX, especialmente após as revoluções industriais, as lutas sociais e as duas Guerras Mundiais, estes direitos buscam garantir a igualdade material e o bem-estar social. Ao contrário da primeira dimensão, os DESC são considerados direitos de "prestação" ou "liberdades positivas", pois exigem uma atuação ativa e propositiva do Estado para sua efetivação, por meio de políticas públicas e investimentos sociais. Incluem o direito ao trabalho, a condições justas e favoráveis de trabalho, à saúde, à educação, à moradia adequada, à previdência social, à alimentação, à cultura e à participação na vida cultural da comunidade. A Constituição Federal de 1988 do Brasil é um exemplo notável de um diploma legal que consagra amplamente esses direitos, refletindo o compromisso com a construção de uma sociedade mais justa e solidária, onde a igualdade não é apenas formal, mas também material (BRASIL, 1988). A interdependência entre a primeira e a segunda dimensão é evidente: de que adianta a liberdade de expressão se o indivíduo não tem acesso à educação para formar sua opinião ou a condições mínimas de existência que lhe permitam participar plenamente do debate público? (CANOTILHO, 2003).

Terceira Dimensão (Direitos de Fraternidade ou Solidariedade): Esta dimensão surge na segunda metade do século XX, em resposta a desafios que transcendem as fronteiras dos Estados-nação e a capacidade de um único indivíduo de resolvê-los. São direitos de titularidade difusa ou coletiva, que enfatizam a solidariedade e a cooperação entre os povos e as nações. Exemplos proeminentes incluem o direito à paz, ao desenvolvimento sustentável, a um meio ambiente ecologicamente

equilibrado, à autodeterminação dos povos e ao patrimônio comum da humanidade. A preocupação com as gerações futuras e com a sustentabilidade do planeta é um traço marcante desta dimensão. A globalização e a interconexão mundial tornaram esses direitos cada vez mais relevantes, exigindo uma abordagem transnacional e a colaboração entre Estados, organizações internacionais e a sociedade civil para sua proteção e promoção. A busca por um mundo mais justo e equitativo, onde os recursos são compartilhados e os desafios globais são enfrentados coletivamente, é o cerne desta dimensão (MAZZUOLI, 2019).

Quarta Dimensão (Direitos da Globalização e da Tecnologia): Esta é a dimensão mais recente e ainda em fase de consolidação e intenso debate acadêmico. Ela é impulsionada pelos avanços exponenciais da tecnologia da informação e da comunicação, da biotecnologia e da globalização, que criam novos desafios e dilemas éticos e jurídicos. Inclui direitos relacionados à bioética (como o direito à integridade genética, à não-manipulação do genoma humano e à autodeterminação informacional), à democracia digital (direito de acesso à internet como um direito fundamental, participação em processos decisórios online), à proteção de dados pessoais e à privacidade no ambiente digital, e aos impactos da inteligência artificial e da robótica. Alguns autores, como Paulo Bonavides (2004), já se referem a essa dimensão como fundamental para a adaptação dos direitos humanos às novas realidades da era digital, onde a informação, a conectividade e a tecnologia desempenham um papel central na vida dos indivíduos e das sociedades (BONAVIDES, 2004). A discussão sobre o discurso de ódio online, por exemplo, insere-se diretamente nesta

dimensão, pois envolve a colisão de direitos fundamentais no ambiente virtual e a necessidade de regulamentação de novas tecnologias (BARROSO, 2010).

É fundamental reiterar que, apesar da categorização em dimensões, os direitos humanos formam um todo coeso e interdependente. A violação de um direito, em qualquer dimensão, pode comprometer a efetivação dos demais. A interconexão entre o direito à liberdade de expressão (primeira dimensão), o direito à educação (segunda dimensão) e o direito a um ambiente digital seguro (quarta dimensão) é um exemplo claro dessa indivisibilidade, especialmente no contexto do combate ao discurso de ódio. A proteção efetiva dos direitos humanos exige uma abordagem holística que reconheça e promova a inter-relação entre todas as suas dimensões (MENDES; BRANCO, 2020).

2.1.3 Princípio da dignidade da pessoa humana

O princípio da dignidade da pessoa humana não é meramente um conceito jurídico, mas o pilar ético e moral que sustenta todo o edifício dos direitos humanos. No ordenamento jurídico brasileiro, sua centralidade é inquestionável, sendo elevado à condição de fundamento da República Federativa do Brasil, conforme expresso no artigo 1º, inciso III, da Constituição Federal de 1988. Essa posição privilegiada não é acidental; ela reflete a compreensão de que a dignidade não é um direito entre outros, mas a fonte e o fim de todos os direitos, conferindo a cada indivíduo um valor intrínseco, inestimável e incondicional, que deve ser reconhecido, respeitado e protegido em todas as circunstâncias, sem qualquer tipo de

relativização ou hierarquização (SARLET, 2021).

A essência desse princípio reside na ideia de que o ser humano jamais pode ser instrumentalizado, ou seja, tratado como um simples meio ou objeto para a consecução de fins alheios, sejam eles estatais, sociais, econômicos ou políticos. Pelo contrário, cada pessoa deve ser sempre considerada como um fim em si mesma, dotada de autonomia, racionalidade e valor próprio. Essa concepção, profundamente enraizada na filosofia moral de Immanuel Kant, que defendia que a humanidade deve ser tratada sempre como um fim e nunca meramente como um meio, serve como um limite material intransponível para a atuação tanto do Estado quanto dos particulares (KANT, 1986). Qualquer ação que viole a integridade física, moral, psicológica ou existencial do indivíduo, ou que o reduza a uma condição de objeto, desprovido de vontade e autodeterminação, é frontalmente incompatível com a dignidade da pessoa humana (BARROSO, 2010).

No contexto do discurso de ódio, a dignidade da pessoa humana assume um papel de baluarte inexpugnável e de critério delimitador da liberdade de expressão. É precisamente em nome da proteção desse princípio que se justifica a limitação da liberdade de expressão quando esta se desvirtua para manifestações que visam à desumanização, à humilhação, à incitação à violência, à discriminação ou à estigmatização contra qualquer pessoa ou grupo. O discurso de ódio, ao atacar a essência da dignidade de um indivíduo ou coletividade, nega a própria premissa de que todos são iguais em valor e merecem respeito, promovendo a exclusão e a intolerância (WALDRON, 2012). Assim, a dignidade não é um

obstáculo à liberdade de expressão, mas o seu limite intrínseco, garantindo que a liberdade de um não se transforme na opressão do outro e que o exercício de um direito não anule a dignidade alheia (ALEXY, 2015).

O Supremo Tribunal Federal (STF) no Brasil tem reiteradamente afirmado a centralidade da dignidade da pessoa humana em suas decisões, utilizando-a como critério interpretativo para a aplicação de diversos direitos fundamentais e para a ponderação de conflitos entre princípios constitucionais. A dignidade é vista como o núcleo essencial dos direitos humanos, irradiando seus efeitos por todo o ordenamento jurídico e servindo como vetor para a construção de uma sociedade justa, solidária e livre de preconceitos. Decisões que criminalizam a homofobia e a transfobia, por exemplo, são exemplos claros da aplicação da dignidade como fundamento para a proteção de grupos vulneráveis contra o discurso de ódio (BRASIL, 2019). A proteção da dignidade, portanto, não é apenas um dever do Estado, mas um imperativo ético para toda a sociedade, que deve zelar para que cada indivíduo seja tratado com o respeito e a consideração que lhe são devidos (COMPARATO, 2003).

2.1.4 Direito à não-discriminação

O direito à não-discriminação, como já mencionado, não é apenas um preceito legal, mas um imperativo ético e moral que emana diretamente do princípio da dignidade da pessoa humana. Ele se configura como um dos pilares inabaláveis do sistema internacional e nacional de direitos humanos, garantindo que todas as pessoas sejam tratadas com igual respeito e consideração, independentemente de suas características

individuais ou de grupo. A essência desse direito reside na proibição de qualquer tratamento diferenciado ou prejudicial que não seja justificado por critérios objetivos e razoáveis, e que tenha por fim ou efeito anular ou comprometer o reconhecimento, gozo ou exercício, em igualdade de condições, de direitos e liberdades fundamentais (DWORKIN, 2002).

A Declaração Universal dos Direitos Humanos (DUDH), adotada em 1948, é o documento seminal que estabelece a base para o direito à não-discriminação. Seu artigo 2º proclama de forma inequívoca que "todo ser humano tem capacidade para gozar todos os direitos e as liberdades estabelecidos nesta Declaração, sem distinção de qualquer espécie, seja de raça, cor, sexo, língua, religião, opinião política ou de outra natureza, origem nacional ou social, riqueza, nascimento, ou qualquer outra condição" (ONU, 1948). Essa disposição fundamental serve de base para inúmeros tratados e convenções internacionais que buscam erradicar a discriminação em suas diversas manifestações, como a Convenção Internacional sobre a Eliminação de Todas as Formas de Discriminação Racial (1965), a Convenção sobre a Eliminação de Todas as Formas de Discriminação contra a Mulher (1979), e a Convenção sobre os Direitos das Pessoas com Deficiência (2006), que demonstram o compromisso da comunidade internacional em proteger grupos historicamente marginalizados (ONU, 1965, 1979, 2006).

No Brasil, o direito à não-discriminação é amplamente protegido pela Constituição Federal de 1988, que se destaca por seu caráter inclusivo e garantista. O artigo 3º, inciso IV, estabelece como um dos objetivos fundamentais da República "promover o bem de todos, sem preconceitos

de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação". Adicionalmente, o artigo 5º, caput, garante a igualdade de todos perante a lei, sem distinção de qualquer natureza, e pelo inciso XLI do mesmo artigo, que prevê a punição de qualquer discriminação atentatória dos direitos e liberdades fundamentais (BRASIL, 1988).

A não-discriminação é essencial para a efetivação de todos os demais direitos humanos. A discriminação, em suas múltiplas formas – seja ela direta ou indireta, intencional ou estrutural – impede que indivíduos e grupos acessem e usufruam plenamente de suas prerrogativas. Ela cria barreiras que limitam o acesso à educação, ao trabalho, à saúde, à justiça, à moradia e à participação política, perpetuando desigualdades e marginalizando grupos inteiros da sociedade. Ao negar a plena cidadania e o reconhecimento de sua dignidade, a discriminação mina os fundamentos de uma sociedade justa e equitativa (PIOVESAN, 2021).

No contexto do discurso de ódio, o direito à não-discriminação é direta e gravemente violado. O discurso de ódio se manifesta precisamente através da segregação, do preconceito, da estigmatização e da incitação à aversão, à hostilidade ou à violência contra grupos específicos com base em suas características identitárias, como raça, etnia, religião, orientação sexual, identidade de gênero, deficiência, entre outras. Ao desumanizar, inferiorizar e incitar a violência, o discurso de ódio nega a igualdade inerente a todos os seres humanos e fomenta um ambiente de intolerância que pode escalar para a violência física, a perseguição e até mesmo genocídios (GOMES et al, 2020).

O combate ao discurso de ódio, portanto, não é apenas uma questão

de proteger a honra ou a imagem de indivíduos, mas uma medida imperativa para garantir o direito fundamental à não-discriminação e para assegurar que a dignidade da pessoa humana seja respeitada em todas as suas dimensões, inclusive no ambiente digital. A proteção contra a discriminação é um pilar para a construção de uma sociedade verdadeiramente inclusiva, plural e democrática, onde a diversidade é valorizada e todos podem coexistir em respeito mútuo e com plenas condições de exercer seus direitos e liberdades (SARLET, 2021).

2.2 FUNDAMENTOS CONSTITUCIONAIS DA LIBERDADE DE EXPRESSÃO

2.2.1 Evolução Histórica e Filosófica da Liberdade e Expressão

A liberdade de expressão, hoje amplamente reconhecida como um dos pilares essenciais e inegociáveis do Estado Democrático de Direito, possui raízes filosóficas profundas e uma evolução histórica complexa, marcada por intensas lutas contra o arbítrio estatal e clerical, e em favor da autonomia individual e do florescimento do conhecimento. Embora as civilizações clássicas, notadamente a democracia ateniense, já contassem com rudimentos significativos desse princípio: como a valorização do livre debate na ágora, expressa nos conceitos de *isegoria* (igualdade no direito de falar na assembleia) e *parrhesia* (o direito e dever de falar francamente, com coragem, mesmo que a verdade seja desconfortável ao poder). Sua concepção moderna, tal como a compreendemos hoje, está intrinsecamente associada ao surgimento do pensamento liberal, à Reforma Protestante e, de forma decisiva, ao Iluminismo europeu.

Foucault (2006), ao analisar a *parrhesia*, destaca que não se tratava apenas de liberdade, mas de uma prática ética e política de coragem na enunciação da verdade, um elemento vital para a cidadania ativa na pólis grega.

Entre os precursores fundamentais do pensamento moderno sobre a liberdade de expressão, destaca-se a figura de John Milton, cujo eloquente discurso *Areopagítica* (1644), dirigido ao Parlamento inglês em pleno contexto das guerras civis, permanece como um marco na defesa da liberdade de imprensa. Nessa obra seminal, Milton insurge-se veementemente contra as licenças prévias e a censura, argumentando que “quem mata um bom livro, mata a razão em si mesma” (MILTON, 2003, p. 14). Sua crítica transcende a mera defesa da publicação; ele sinaliza que a censura atinge não apenas a obra individual, mas o próprio espírito humano, a capacidade coletiva de buscar a verdade. Para Milton, a liberdade de imprimir e circular ideias, mesmo as controversas ou tidas como errôneas, seria a condição indispensável para o florescimento da verdade e da virtude cívica, devendo o Estado confiar na capacidade de discernimento dos cidadãos e no poder autocorretivo do debate aberto. A verdade, segundo ele, não necessita de estratégias ou licenças para vencer; basta que o campo esteja livre para o confronto de ideias (MILTON, 2003).

O pensamento de John Locke (2000), outro pilar do liberalismo, também influenciou profundamente a concepção moderna de liberdade, fornecendo bases filosóficas robustas. Em sua *Carta sobre a tolerância* (1689), Locke defende a separação entre Estado e Igreja ao afirmar categoricamente que "ninguém pode ser forçado a aceitar uma crença

contra sua vontade" (LOCKE, 2000, p. 39).

Ao posicionar-se contra o autoritarismo religioso e político e delimitar o campo legítimo da atuação estatal à proteção dos direitos civis (vida, liberdade, propriedade), Locke estabelece a tolerância como um princípio fundamental. Essa ideia de tolerância, embora focada inicialmente na esfera religiosa, constitui um alicerce indispensável da liberdade de expressão, especialmente quando relacionada à liberdade de consciência, pensamento e opinião. A autonomia da consciência individual, para Locke (2000), é inviolável pelo poder político, sendo a base para a participação cidadã em uma sociedade livre.

De forma ainda mais sistemática e influente, John Stuart Mill, em sua obra *On Liberty* (1859), desenvolveu uma das mais robustas e duradouras defesas do direito à livre manifestação de pensamento, fundamentada no princípio utilitarista da maximização do bem-estar geral. Para Mill, a liberdade de expressão é elementar para o progresso individual e social. Ele argumenta que “o único modo pelo qual um ser humano pode ser completamente livre é permitindo-se a liberdade de expressão, mesmo para opiniões que consideramos falsas ou ofensivas” (MILL, 2000, p. 45).

Mill (2000), justifica essa posição radical com base em argumentos pragmáticos e epistêmicos, pois, a supressão de uma opinião, mesmo que seja minoritária ou impopular, impede a sociedade de avançar. Isso porque, uma opinião silenciada pode conter a verdade, e ao suprimi-la, a humanidade perde a chance de corrigir seus equívocos. Mesmo que essa opinião seja falsa, é provável que ela contenha uma parcela de verdade, e a confrontação com o erro é fundamental para que a verdade

completa se revele.

Além disso, mesmo que a opinião dominante seja inteiramente verdadeira, se não for constantemente desafiada, ela se torna um dogma sem vida, um mero preconceito, esvaziado de um entendimento genuíno de seus próprios fundamentos.

Dessa forma, o livre mercado de ideias é o ambiente ideal para a descoberta da verdade e para o desenvolvimento intelectual e moral de uma sociedade. Essa defesa da livre expressão, no entanto, não é absoluta. Há limites, especialmente quando a expressão de uma ideia causa dano direto a outras pessoas, conforme o princípio do dano (*harm principle*).

No plano das declarações de direitos, que traduziram essas ideias filosóficas em normas jurídicas, a Declaração dos Direitos do Homem e do Cidadão de 1789, fruto da Revolução Francesa e marco do constitucionalismo moderno, estabelece em seu artigo 11:

A livre comunicação dos pensamentos e das opiniões é um dos direitos mais preciosos do homem; todo cidadão pode, portanto, falar, escrever, imprimir livremente, respondendo, todavia, pelos abusos dessa liberdade nos casos determinados pela lei” (FRANÇA, 1789, tradução nossa).

Essa formulação pioneira já delinea a tensão inerente ao direito: sua fundamentalidade e a necessidade de limites legais para coibir abusos, um equilíbrio que permanece central nos debates contemporâneos. Similarmente, a Primeira Emenda à Constituição dos Estados Unidos (1791) proíbe o Congresso de legislar restringindo a liberdade de expressão ou de imprensa, estabelecendo um padrão de proteção robusto, embora interpretado de formas diversas ao longo da história norte-americana.

O século XX, marcado por conflitos mundiais e regimes totalitários que suprimiram brutalmente as liberdades, testemunhou a internacionalização da proteção dos direitos humanos. A Declaração Universal dos Direitos Humanos (DUDH), aprovada pela Assembleia Geral da ONU em 1948 como uma resposta direta às atrocidades da Segunda Guerra Mundial, consagrou a liberdade de expressão em seu artigo 19, que reconhece o direito universal de “procurar, receber e difundir informações e ideias por quaisquer meios e independentemente de fronteiras” (ONU, 1948).

Esta disposição, de imenso valor simbólico e político, representou um consenso global sobre a importância da liberdade de expressão para a dignidade humana e a paz mundial. Sua força normativa foi posteriormente reforçada e detalhada em tratados internacionais vinculantes, como o Pacto Internacional dos Direitos Civis e Políticos (PIDCP, 1966), em seu artigo 19, e as convenções regionais, como a Convenção Americana sobre Direitos Humanos (Pacto de San José da Costa Rica, 1969), em seu artigo 13, e a Convenção Europeia dos Direitos Humanos (1950), em seu artigo 10.

Estes tratados não apenas reiteram o direito, mas também estabelecem os parâmetros para suas legítimas restrições, que devem ser previstas em lei, necessárias em uma sociedade democrática e voltadas à proteção de outros direitos ou bens jurídicos relevantes (como a segurança nacional, a ordem pública, a saúde ou a moral públicas, e os direitos e reputação de outrem).

A partir dessa sólida base filosófica e jurídica, consolidada ao

longo de séculos, a liberdade de expressão passou a ser amplamente reconhecida não apenas como um direito individual, mas como um instrumento essencial para o exercício pleno da cidadania, o controle democrático do poder, a formação da opinião pública e a garantia do pluralismo político e cultural. O filósofo do direito Ronald Dworkin, em sua influente obra *Levando os direitos a sério*, reforça essa dimensão estrutural ao afirmar que “sem liberdade de expressão, o cidadão se torna um súdito, impedido de participar criticamente das decisões que regem sua vida” (DWORKIN, 2002, p. 240).

Nesse sentido, o direito à expressão transcende a esfera meramente individual; ele é condição de possibilidade para a própria legitimidade do sistema democrático, pois viabiliza o debate público informado, a crítica às autoridades e a formação de consensos e dissensos que caracterizam uma sociedade aberta. Habermas (2003), por sua vez, enfatiza o papel da comunicação livre e racional na esfera pública como fundamento da ação comunicativa e da própria democracia deliberativa.

No contexto brasileiro, a trajetória da liberdade de expressão foi igualmente marcada por avanços e retrocessos. Embora reconhecida formalmente em textos constitucionais anteriores, começando no período imperial, onde a nossa primeira Constituição de 1824 já previa a liberdade de expressão, porém na prática ela era limitada, pois a censura prévia nos jornais foi instituída no mesmo ano, mostrando que o direito não era plenamente respeitado, sua efetividade ao longo da história foi frequentemente cerceada por regimes autoritários, como durante o Estado Novo (1937-1945) que extinguiu a previsão de liberdade de expressão,

com predominância da censura e, mais recentemente, a Ditadura Militar (1964-1985), período em que a censura prévia e a perseguição a opositores políticos eram práticas correntes, principalmente após a instituição do Ato Institucional nº 5 (AI-5), que suspendeu direitos políticos e intensificou a repressão. (MENDES, [2025]) Foi a Constituição Federal de 1988, promulgada em um contexto de redemocratização e ampla participação social, que conferiu à liberdade de expressão a mais robusta proteção jurídica de nossa história, trataremos dela, de forma mais dedicada, no próximo tópico.

Dessa forma, a análise da evolução histórica e filosófica da liberdade de expressão revela um percurso complexo, que parte de suas origens na antiguidade clássica, consolida-se com o pensamento liberal e iluminista, e se universaliza no direito internacional dos direitos humanos. Essa trajetória demonstra não apenas sua raiz liberal-individualista, focada na autonomia do sujeito, mas também seu papel cada vez mais reconhecido como direito coletivo e estrutural, indispensável ao funcionamento das democracias contemporâneas. A defesa intransigente desse direito exige, portanto, vigilância constante e aprofundamento crítico, especialmente em tempos de crescente polarização política, desinformação deliberada e proliferação do discurso de ódio facilitada pelos meios digitais, desafios que demandam respostas jurídicas e sociais equilibradas e eficazes.

2.2.2 A Liberdade de Expressão no Sistema Constitucional Brasileiro

A Constituição da República Federativa do Brasil de 1988, promulgada em um momento histórico de redemocratização após mais de

duas décadas de regime militar, representou uma profunda ruptura com o autoritarismo e a instauração de um novo pacto normativo fundado na dignidade da pessoa humana (art. 1º, III), nos direitos e garantias fundamentais e no pluralismo político (art. 1º, V). Nesse arcabouço axiológico e normativo, a liberdade de expressão ocupa uma posição de destaque inegável, sendo tratada não apenas como um direito fundamental de primeira geração (ou dimensão), essencial à autonomia individual, mas também como uma garantia institucional indispensável para a própria existência e vitalidade do regime democrático e para o exercício efetivo da cidadania plena. Sua consagração reflete a opção inequívoca do constituinte originário por uma sociedade aberta, participativa e crítica, em contraposição direta à censura e à repressão características do período anterior.

O núcleo normativo da liberdade de expressão na Carta Magna brasileira encontra-se primordialmente no artigo 5º, que arrola os direitos e deveres individuais e coletivos. O inciso IV dispõe que "é livre a manifestação do pensamento, sendo vedado o anonimato", enquanto o inciso IX assegura que "é livre a expressão da atividade intelectual, artística, científica e de comunicação, independentemente de censura ou licença". Complementarmente, o artigo 220 estabelece que "a manifestação do pensamento, a criação, a expressão e a informação, sob qualquer forma, processo ou veículo não sofrerão qualquer restrição, observado o disposto nesta Constituição", vedando ainda, em seu § 2º, "toda e qualquer censura de natureza política, ideológica e artística". Essas disposições, interpretadas sistematicamente, conferem à liberdade de

expressão um duplo caráter, como bem aponta a doutrina majoritária: uma dimensão subjetiva, enquanto direito fundamental oponível ao Estado e, em certas circunstâncias, a particulares (eficácia horizontal), que protege a esfera de autonomia de cada cidadão para formar e externar suas opiniões e ideias; e uma dimensão objetiva, enquanto elemento estruturante e garantia institucional para o funcionamento saudável da democracia, viabilizando o debate público, a formação da opinião pública, o controle social dos atos do poder e o pluralismo de ideias (MENDES; BRANCO, 2020). Conforme destaca com acuidade o ministro do STF, Luís Roberto Barroso:

A liberdade de expressão ocupa posição preferencial (*preferred position*) na ordem constitucional brasileira, sendo condição para o exercício de outros direitos fundamentais, como o direito de voto, de reunião, de associação e de participação política em geral (BARROSO, 2017, p. 47).

Essa teoria da posição preferencial, originária da jurisprudência norte-americana e recepcionada pela doutrina e jurisprudência brasileiras, não significa que a liberdade de expressão seja um direito absoluto, mas sim que ela possui um peso abstrato maior em eventuais ponderações com outros direitos ou interesses constitucionalmente protegidos. Isso implica um ônus argumentativo mais elevado para justificar qualquer restrição a essa liberdade, que deve ser sempre excepcional, estritamente necessária e proporcional. Silva (2010) aprofunda essa análise, argumentando que a preferência se justifica pelo papel instrumental da liberdade de expressão para a democracia e para a busca da verdade, exigindo um escrutínio judicial mais rigoroso sobre as leis ou atos que a restrinjam.

Todavia, a jurisprudência que vem sendo construída no Supremo

Tribunal Federal, deixa claro que a liberdade de expressão, apesar de sua posição preferencial, não é um direito absoluto ou ilimitado. Ela encontra limites imanentes em outros direitos e valores constitucionais igualmente relevantes, como a dignidade da pessoa humana (fundamento da República, art. 1º, III), os direitos da personalidade (honra, imagem, privacidade, intimidade – art. 5º, X), a proteção da infância e da juventude, a vedação ao racismo (art. 5º, XLII) e a própria defesa do Estado Democrático de Direito contra discursos que visem sua ruptura.

Os direitos fundamentais não são absolutos, devendo ser interpretados e aplicados com base no princípio da concordância prática ou harmonização, que busca otimizar a realização de todos os direitos em conflito em um caso concreto, evitando o sacrifício total de um em detrimento do outro” (SARLET; MARINONI; MITIDIERO, 2021, p. 205).

A ponderação de interesses, guiada pelo princípio da proporcionalidade (necessidade, adequação e proporcionalidade em sentido estrito), torna-se, assim, a ferramenta hermenêutica central para solucionar colisões entre a liberdade de expressão e outros bens jurídicos constitucionais.

Nesse sentido, a liberdade de expressão no sistema constitucional brasileiro assume uma configuração complexa e dinâmica, que busca garantir um espaço amplo para a circulação de ideias e o debate público, protegendo-a vigorosamente contra interferências indevidas do Estado (especialmente a censura prévia), mas que, ao mesmo tempo, impõe deveres e responsabilidades aos indivíduos e às plataformas digitais.

A responsabilização *a posteriori* por abusos no exercício desse direito – seja por danos morais, crimes contra a honra, racismo, incitação à violência ou outros ilícitos – é parte integrante do regime constitucional

da liberdade de expressão, conforme previsto na própria Constituição (art. 5º, V e X; art. 220, § 1º). Trata-se, portanto, de uma liberdade ampla, preferencial, mas não absoluta; uma liberdade que deve ser exercida com responsabilidade, em respeito aos demais direitos fundamentais e aos pilares da ordem democrática, não legitimando a ofensa gratuita, o preconceito, a desinformação dolosa ou a incitação ao ódio e à violência. O desafio contemporâneo reside em encontrar o equilíbrio adequado entre a máxima proteção da liberdade e a coibição eficaz de seus abusos, especialmente no ambiente digital, tema que será aprofundado nos capítulos subsequentes.

2.2.3 Liberdade de Informação, Crítica, Imprensa e Opinião

A liberdade de expressão, consagrada como direito fundamental multifacetado no ordenamento jurídico brasileiro e internacional, não se apresenta como um bloco monolítico. Pelo contrário, ela se desdobra em diversas manifestações específicas, interdependentes e dotadas de relevância jurídica autônoma, que garantem a sua plena efetividade no contexto democrático. Entre essas dimensões, destacam-se a liberdade de informação, a liberdade de crítica, a liberdade de imprensa e a liberdade de opinião. Compreender as nuances e os contornos de cada uma dessas liberdades é elementar para a correta aplicação do direito e para a salvaguarda do pluralismo, da transparência e da cidadania ativa que caracterizam o Estado Democrático de Direito.

A liberdade de informação, em sua acepção mais ampla, compreende um feixe de direitos interligados: o direito de buscar

ativamente informações (direito de acesso), o direito de produzir e elaborar informações, o direito de difundir informações e ideias (direito de informar) e o direito de receber informações (direito de ser informado). A Constituição Federal de 1988 assegura essa liberdade de forma robusta, tanto em sua dimensão individual quanto institucional. O artigo 5º, inciso XIV, garante a todos o acesso à informação, resguardado o sigilo da fonte quando necessário ao exercício profissional. De maneira ainda mais abrangente, o artigo 220 estabelece que "a manifestação do pensamento, a criação, a expressão e a informação, sob qualquer forma, processo ou veículo não sofrerão qualquer restrição, observado o disposto nesta Constituição". Essa garantia abarca não apenas a liberdade de expressar o que se pensa, mas também o direito fundamental de ter acesso aos fatos, dados e conhecimentos necessários para a formação de uma opinião consciente e para a participação informada nos assuntos públicos. A função eminentemente informativa da comunicação social é, portanto, essencial para a formação da opinião pública livre, para o controle democrático das instituições e para a própria deliberação coletiva. Como bem aponta Barroso, "a liberdade de informação jornalística é um dos pilares da *accountability* democrática e da transparência na gestão pública, permitindo que os cidadãos fiscalizem os atos de seus representantes" (BARROSO, 2017, p. 122). A Lei de Acesso à Informação (Lei nº 12.527/2011) veio a regulamentar o direito de acesso à informação pública, reforçando essa dimensão da cidadania.

A liberdade de crítica, por sua vez, representa a faculdade de formular juízos de valor, avaliações e apreciações sobre pessoas, ideias,

obras, instituições ou atos, especialmente aqueles emanados de figuras públicas e do governo (BARROSO, 2017). A crítica, mesmo quando ácida, contundente ou incômoda para seus destinatários, constitui uma das formas mais legítimas e necessárias de exercício da liberdade de expressão em uma democracia. Ela é o motor do debate público, permitindo o escrutínio das ações estatais e a contestação de ideias dominantes. Paulo Bonavides (2004), com sua reconhecida autoridade, destaca que “a crítica, sobretudo a crítica política, pertence ao núcleo duro da liberdade de expressão, sendo sua restrição um sintoma inequívoco de regressão autoritária e de asfixia democrática” (BONAVIDES, 2004, p. 386). Obviamente, a liberdade de crítica não é ilimitada, encontrando barreiras nos direitos da personalidade, como a honra e a imagem. Contudo, a jurisprudência, tanto nacional quanto internacional, tende a conferir uma proteção mais ampla à crítica dirigida a agentes públicos ou a assuntos de interesse geral, reconhecendo uma maior tolerância a manifestações severas nesse âmbito, em nome do interesse público na fiscalização do poder (MENDES; BRANCO, 2020).

A liberdade de imprensa, frequentemente tratada como sinônimo de liberdade de informação, possui contornos próprios e representa uma das manifestações institucionais mais relevantes da liberdade de expressão. Ela assegura aos meios de comunicação social (jornais, revistas, rádio, televisão, internet) a prerrogativa de investigar, apurar, editar e divulgar notícias, fatos e opiniões, sem sofrer censura prévia ou interferências indevidas por parte do Estado. A imprensa livre desempenha um papel importante como "cão de guarda" da democracia, fiscalizando os

poderes constituídos, denunciando abusos e garantindo a pluralidade de vozes no espaço público. A decisão do STF na Arguição de Descumprimento de Preceito Fundamental (ADPF) 130/DF foi um marco na reafirmação dessa liberdade no Brasil pós-1988, ao banir a Lei de Imprensa do regime militar e vedar qualquer forma de censura. No entanto, essa ampla liberdade conferida à imprensa não é absoluta e não a isenta de responsabilidades. No julgamento do Recurso Extraordinário n.º 511.961/SP, o STF, embora reafirmando a liberdade de informação jornalística.

A liberdade de informação jornalística, embora ampla, não se sobrepõe, de modo absoluto, aos direitos da personalidade (honra, imagem, privacidade, intimidade), devendo ser exercida com diligência, com base na veracidade dos fatos e na pertinência da informação divulgada (BRASIL, STF, RE 511961/SP, Rel. Min. Carlos Britto, julgado em 17 jun. 2009).

Essa decisão evidencia a aplicação do princípio da concordância prática ou harmonização: a imprensa tem o direito e o dever de informar, mas deve fazê-lo de forma responsável, checando a veracidade dos fatos (dever de diligência) e evitando a divulgação de informações que violem injustificadamente a esfera íntima dos indivíduos. A responsabilidade civil (indenização por danos morais e materiais) e, em casos mais graves, a responsabilidade penal (crimes contra a honra) são os mecanismos previstos para coibir os abusos.

Finalmente, a liberdade de opinião protege as manifestações de convicções, crenças e juízos subjetivos, sejam eles de natureza filosófica, religiosa, política, moral ou estética. Diferentemente das informações factuais, que podem ser aferidas quanto à sua veracidade ou falsidade, as

opiniões expressam pontos de vista pessoais e, por essa razão, gozam, em princípio, de um grau ainda mais elevado de proteção jurídica, sendo mais resistentes a restrições. Robert Alexy (2015), em sua influente teoria dos direitos fundamentais.

Juízos de valor não são suscetíveis de demonstração empírica verdadeira ou falsa, e por isso sua limitação exige um critério mais estrito de justificação, baseado não na verdade, mas na ponderação com outros direitos ou valores. (ALEXY, 2015, p. 448).

Isso não significa, contudo, que toda e qualquer opinião esteja imune a limites. Mesmo as opiniões podem ultrapassar as fronteiras da licitude quando se revestem de caráter manifestamente discriminatório, quando incitam ao ódio, à violência ou à subversão da ordem democrática, ou quando configuram ofensa grave e gratuita à dignidade de outrem. A linha divisória entre a opinião crítica, ainda que dura, e o discurso de ódio ou a injúria discriminatória é tênue e desafiadora, exigindo análise cuidadosa do contexto, da intenção do emissor e do impacto sobre o receptor e a sociedade.

2.2.4 Limites Constitucionais E Cláusulas De Vedação À Liberdade De Expressão

Embora a liberdade de expressão seja alçada à condição de direito fundamental cardeal e pressuposto da própria democracia em todos os sistemas constitucionais contemporâneos, sua natureza não é, e nem poderia ser, absoluta. Essa compreensão não representa uma fragilidade, mas sim uma decorrência lógica e necessária da própria estrutura axiológica do Estado Constitucional de Direito, que se fundamenta na

coexistência harmônica de uma pluralidade de direitos e valores, todos eles igualmente dignos de proteção. O exercício irrestrito de um único direito individual, levado às últimas consequências, poderia aniquilar ou comprometer severamente a existência de outros direitos igualmente fundamentais, gerando uma situação de desequilíbrio e injustiça incompatível com o ideal de uma sociedade livre e igualitária. Nesse contexto, o princípio da unidade da Constituição e, mais especificamente, o princípio da concordância prática ou da harmonização dos direitos fundamentais, emergem como vetores hermenêuticos indispensáveis. Eles orientam o intérprete e o aplicador do direito na busca por soluções que otimizem a realização de todos os direitos em colisão, estabelecendo limitações legítimas à liberdade de expressão sempre que seu exercício ameace violar a dignidade da pessoa humana, a igualdade, a segurança coletiva ou outros bens jurídicos constitucionalmente tutelados (CANOTILHO, 2003; ÁVILA, 2011).

No ordenamento jurídico brasileiro, os limites à liberdade de expressão encontram-se delineados tanto de forma expressa quanto implícita no texto da Constituição Federal de 1988, refletindo as escolhas valorativas do poder constituinte originário. De maneira explícita, o artigo 5º, inciso IV, ao mesmo tempo que garante a livre manifestação do pensamento, veda o anonimato. Essa vedação não é uma restrição à liberdade em si, mas uma condição para seu exercício responsável, assegurando a possibilidade de identificação e responsabilização jurídica do autor por eventuais abusos ou ilícitos cometidos. O inciso V do mesmo artigo garante o direito de resposta, proporcional ao agravo, além da

indenização por dano material, moral ou à imagem, estabelecendo um mecanismo de reparação *a posteriori*. O inciso X protege expressamente a intimidade, a vida privada, a honra e a imagem das pessoas, assegurando o direito à indenização pelo dano material ou moral decorrente de sua violação, configurando limites claros ao que pode ser dito ou divulgado sobre terceiros. De forma ainda mais contundente, o inciso XLII estabelece que “a prática do racismo constitui crime inafiançável e imprescritível, sujeito à pena de reclusão, nos termos da lei”, deixando inequívoco que manifestações de cunho racista não encontram guarida sob o manto protetor da liberdade de expressão. Outros limites explícitos na constituição podem ser encontrados na proteção à infância e à juventude (art. 227), que justifica restrições a conteúdos impróprios, e nas vedações à propaganda de guerra, de processos violentos para alterar a ordem política e social, e de preconceitos de raça ou classe (art. 220, § 1º, c/c art. 5º, XVI e XVII, interpretados sistematicamente).

Adicionalmente, existem limites implícitos que decorrem dos princípios e valores estruturantes da ordem constitucional. O artigo 220, *caput*, ao afirmar que a manifestação do pensamento e a informação não sofrerão qualquer restrição, ressalva expressamente: "observado o disposto nesta Constituição". Isso significa que a própria liberdade de expressão encontra limites nos demais preceitos constitucionais. O mais fundamental desses limites é a dignidade da pessoa humana (art. 1º, III), valor-fonte de todo o ordenamento jurídico, que impede que a liberdade de expressão seja utilizada como instrumento para degradar, humilhar ou negar a humanidade de indivíduos ou grupos. O pluralismo político (art. 1º, V) e a

construção de uma sociedade livre, justa e solidária (art. 3º, I), bem como a promoção do bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação (art. 3º, IV), também funcionam como balizas interpretativas, impedindo que a liberdade de expressão seja invocada para defender discursos antidemocráticos, totalitários ou que fomentem a exclusão e a intolerância. Como ensina Ingo Wolfgang Sarlet (2021), com precisão, “nenhum direito fundamental pode ser interpretado ou aplicado de forma isolada, desgarrado do sistema constitucional em que se insere. Pelo contrário, deve ser compreendido no contexto do conjunto de normas constitucionais e dos princípios estruturantes do Estado, buscando-se sempre a máxima efetividade de todos os valores em jogo” (SARLET; MARINONI; MITIDIERO, 2021, p. 210).

Nesse sentido, a liberdade de expressão deve ser constantemente compatibilizada com os princípios da dignidade humana, da igualdade e da proibição de qualquer forma de discriminação, sendo não apenas legítimo, mas um dever de o Estado intervir para vedar abusos e proteger grupos historicamente vulnerabilizados contra discursos que perpetuem sua marginalização.

Essa compreensão dos limites da liberdade de expressão no Brasil está em ampla consonância com os parâmetros estabelecidos no direito internacional dos direitos humanos. A Convenção Americana sobre Direitos Humanos (Pacto de San José da Costa Rica), incorporada ao ordenamento brasileiro com status supralegal (ou, para alguns, constitucional), prevê em seu art. 13, § 5º, que a lei pode proibir "toda

propaganda a favor da guerra, bem como toda apologia ao ódio nacional, racial ou religioso que constitua incitação à discriminação, à hostilidade, ao crime ou à violência". A Corte Interamericana de Direitos Humanos (Corte IDH), intérprete autêntica da Convenção, tem reiterado essa posição.

É fundamental destacar, contudo, que a existência de limites não pode ser um pretexto para a imposição de restrições arbitrárias ou desproporcionais à liberdade de expressão, nem para a perseguição de opositores ou o silenciamento de vozes críticas ao poder estabelecido. A aplicação de qualquer limitação deve ser sempre excepcional e rigorosamente justificada à luz do princípio da proporcionalidade, em sua tríplice dimensão: adequação (a medida restritiva deve ser apta a atingir o fim legítimo almejado), necessidade (a medida deve ser a menos gravosa possível entre as igualmente eficazes) e proporcionalidade em sentido estrito (as vantagens da restrição devem superar suas desvantagens, ponderando-se os bens jurídicos em conflito). Como sublinha Robert Alexy (2015), “toda restrição a direitos fundamentais deve ser submetida a um exame de proporcionalidade, que funciona como um critério racional para a solução de colisões entre princípios” (ALEXY, 2015, p. 519). Assim, o Estado só deve intervir para restringir a expressão quando houver um dano concreto ou um risco claro e iminente (*clear and present danger*) de violação a outros direitos fundamentais, e nunca com base em meras suposições, abstrações, paternalismos ou moralismos políticos. A vagueza ou a amplitude excessiva das leis restritivas (*overbreadth*) também são incompatíveis com a liberdade de expressão.

O debate contemporâneo em torno dos discursos de ódio, da desinformação e da regulação das plataformas digitais reacende com intensidade essa tensão inerente entre liberdade e limitação. A velocidade e o alcance da disseminação de conteúdos nocivos na internet, potencializados por algoritmos, criam desafios inéditos. A ausência de uma legislação específica e abrangente sobre discurso de ódio no Brasil (embora existam tipos penais como racismo, injúria racial, incitação ao crime, etc.) dificulta a aplicação uniforme e previsível das cláusulas de vedação, gerando um vácuo normativo que tem sido preenchido, muitas vezes de forma casuística, por decisões judiciais e pelas próprias políticas de moderação de conteúdo das plataformas digitais privadas. Isso, por sua vez, levanta novas e complexas questões sobre segurança jurídica, devido processo legal, o poder privado de censura exercido pelas *big techs* (as empresas gigantes da tecnologia), a transparência dos algoritmos e a necessidade (ou não) de uma regulação estatal mais robusta, como proposto no PL 2630/2020 (Lei das Fake News), objeto de intenso debate público.

Portanto, os limites constitucionais e as cláusulas de vedação à liberdade de expressão no Brasil não devem ser vistos como uma negação ou diminuição de sua importância, mas sim como uma condição essencial para sua própria legitimidade e sustentabilidade em uma sociedade democrática e pluralista. Eles são a expressão do compromisso fundamental do Estado com a proteção da dignidade de todas as pessoas, com a igualdade substancial e com a prevenção de discursos que possam corroer os alicerces da convivência pacífica e da própria democracia. A

definição da fronteira delicada entre o discurso protegido, ainda que chocante ou ofensivo, e o discurso ilícito, que incita ao ódio ou à violência, deve ser sempre traçada com base em critérios jurídicos objetivos, claros e previsíveis, por meio de leis precisas e de uma interpretação judicial prudente e contextualizada, garantindo-se, assim, tanto a máxima amplitude possível para a liberdade quanto a responsabilização efetiva por seus abusos.

2.3 DISCURSO DE ÓDIO: CONCEITO, TIPIIFICAÇÃO E CONTROVÉRSIAS

2.3.1 Conceito Jurídico e Filosófico de Discurso de Ódio

A expressão “discurso de ódio” (ou *hate speech*, em sua formulação anglo-saxônica original, que ganhou ampla circulação internacional) ocupa uma posição de inegável centralidade nos debates jurídicos, políticos e sociais contemporâneos sobre os limites éticos e legais da liberdade de expressão e a necessária proteção dos direitos humanos em sociedades plurais. Trata-se, contudo, de um conceito polissêmico e contestado que, apesar de sua vasta utilização em discursos acadêmicos, midiáticos e em documentos normativos, ainda carece de uma definição jurídica universalmente aceita e unívoca no direito internacional. Essa ausência de consenso reflete a profunda tensão entre a proteção da liberdade de manifestação, mesmo de ideias que possam chocar ou ofender, e a necessidade de coibir expressões que incitem à violência, à discriminação e que neguem a igual dignidade de grupos vulneráveis (WALDRON, 2012).

Em termos gerais, e buscando uma síntese operacional, o discurso

de ódio pode ser compreendido como qualquer forma de comunicação – seja ela verbal, escrita, visual, simbólica ou comportamental – que ataque ou use linguagem pejorativa ou discriminatória com referência a uma pessoa ou grupo com base em atributos como religião, etnia, nacionalidade, raça, cor, descendência, gênero, orientação sexual, deficiência ou outro fator identitário. Seu propósito ou efeito é, frequentemente, deslegitimar, inferiorizar, estigmatizar, hostilizar, intimidar ou incitar à discriminação, à hostilidade e à violência contra esses indivíduos ou grupos.

A Estratégia e Plano de Ação das Nações Unidas sobre o Discurso de Ódio, embora reconheça a falta de uma definição legal internacional, adota uma definição de trabalho similar, descrevendo-o como "qualquer tipo de comunicação [...] que ataque ou use linguagem pejorativa ou discriminatória com referência a uma pessoa ou grupo [...]" (UNITED NATIONS, 2019).

No plano filosófico, a controvérsia em torno do discurso de ódio está intrinsecamente vinculada à tensão clássica do pensamento liberal entre a autonomia individual e a liberdade de expressão, de um lado, e a busca por justiça social, igualdade e proteção contra danos, de outro. Autores como Jeremy Waldron, em sua influente obra *The Harm in Hate Speech*, argumentam vigorosamente que o discurso de ódio não pode ser tratado como uma mera manifestação de opinião divergente ou ofensiva, a ser tolerada em nome do livre mercado de ideias.

Para Waldron (2012), o dano causado pelo discurso de ódio transcende a ofensa individual; ele ataca a "dignidade pública" das

peçoas, minando a "garantia social básica de que serão tratadas como membros da sociedade em situação de igualdade" (WALDRON, 2012, p. 5, tradução nossa). Ao espalhar mensagens que retratam certos grupos como inferiores, perigosos ou indesejáveis, o discurso de ódio polui o ambiente social, corrói os laços de confiança cívica e nega aos membros desses grupos a segurança fundamental de poderem participar da vida pública sem medo de assédio, humilhação ou violência. Nesse sentido, o discurso de ódio opera como um mecanismo insidioso de violência simbólica e estrutural, que não apenas fere sentimentos, mas ativamente constrói e reforça hierarquias sociais opressivas, naturalizando a exclusão e dificultando a realização da igualdade substantiva.

Em contraposição a essa visão, pensadores alinhados a uma tradição liberal mais robusta na defesa da liberdade de expressão, como Ronald Dworkin (2002), embora reconheçam os potenciais danos sociais do discurso odioso, expressam ceticismo quanto à legitimidade e à eficácia de sua proibição estatal. Dworkin argumenta que a liberdade de expressão possui um valor intrínseco ligado à autonomia moral dos indivíduos e um valor instrumental único para a legitimidade do processo democrático. Para ele, “a liberdade de expressão não é apenas um direito individual, mas uma salvaguarda da legitimidade política, pois permite que os cidadãos participem da formação da vontade coletiva” (DWORKIN, 2002, p. 245).

Proibir certas ideias, mesmo as mais repugnantes, poderia ser interpretado como uma forma de paternalismo estatal, tratando os cidadãos como incapazes de discernir e rejeitar concepções errôneas, além de criar um precedente perigoso para a supressão de discursos dissidentes ou

impopulares no futuro. Dworkin (2002) sugere que a resposta mais adequada ao discurso de ódio não seria a censura, mas sim o contradiscurso (*counterspeech*), o debate robusto e a educação cívica. Contudo, é importante notar que mesmo Dworkin e outros defensores de uma ampla proteção à liberdade de expressão admitem que essa liberdade não é absoluta, reconhecendo a possibilidade de restrições em casos extremos, como a incitação direta e iminente à violência (seguindo a linha do teste de *clear and present danger* da Suprema Corte dos EUA), embora discordem sobre onde exatamente traçar essa linha divisória.

No campo do direito internacional dos direitos humanos, embora não haja uma definição única de "discurso de ódio", diversos tratados e organismos internacionais estabelecem parâmetros importantes para sua identificação e restrição. O Pacto Internacional sobre Direitos Civis e Políticos (PIDCP), em seu artigo 20, parágrafo 2º, determina que "será proibida por lei toda apologia do ódio nacional, racial ou religioso que constitua incitamento à discriminação, à hostilidade ou à violência".

Essa disposição impõe aos Estados Partes uma obrigação positiva de legislar contra as formas mais graves de discurso de ódio, que atingem o limiar da incitação. O Comitê de Direitos Humanos da ONU, em seu Comentário Geral nº 34 sobre o artigo 19 (liberdade de expressão), esclarece que as restrições previstas no artigo 20 são compatíveis com a liberdade de expressão, mas devem ser interpretadas restritivamente e aplicadas com cautela, exigindo prova da intenção de incitar e da probabilidade de ocorrência do resultado danoso.

No âmbito regional europeu, a Corte Europeia de Direitos

Humanos (CEDH), interpretando o artigo 10 da Convenção Europeia (que garante a liberdade de expressão, mas prevê restrições), tem desenvolvido uma jurisprudência consistente no sentido de que expressões que incitem à violência, ao ódio, à discriminação ou que neguem crimes contra a humanidade (como o Holocausto) não se beneficiam da proteção da Convenção, podendo inclusive ser consideradas um abuso de direito (artigo 17). No caso *Erbakan v. Turkey* (2006), a Corte afirmou que “a tolerância e o respeito à igualdade de todos os seres humanos constituem os fundamentos de uma sociedade democrática e pluralista” e, por isso, manifestações que propaguem, incitem ou justifiquem o ódio baseado na intolerância são incompatíveis com esses valores e não merecem proteção (TEDH, 2006). O Conselho da Europa, por meio de sua Comissão Europeia contra o Racismo e a Intolerância (ECRI), também emitiu Recomendações de Política Geral (como a nº 15, sobre combate ao discurso de ódio) que oferecem definições e diretrizes aos Estados membros, definindo discurso de ódio como o fomento, promoção ou incitação ao ódio, humilhação ou menosprezo de uma pessoa ou grupo, assim como o assédio, insulto, estereotipagem negativa, estigmatização ou ameaça, em razão de características como raça, cor, ascendência, origem nacional ou étnica, idade, deficiência, língua, religião ou crença, sexo, gênero, identidade de gênero, orientação sexual, entre outras.

Na esfera interamericana, a Corte Interamericana de Direitos Humanos (Corte IDH) segue uma linha interpretativa semelhante. Seu parecer consultivo OC-5/85, sobre a colegiação obrigatória de jornalistas, já afirmava que a liberdade de expressão, garantida no artigo 13 da

Convenção Americana, não protege manifestações que:

Constituam propaganda a favor da guerra e apologia ao ódio nacional, racial ou religioso que constituam incitações à violência ou qualquer outra ação ilegal semelhante contra qualquer pessoa ou grupo de pessoas, por qualquer motivo, inclusive os de raça, cor, religião, idioma ou origem nacional (CORTE IDH, 1985, par. 66).

Portanto, existe um consenso robusto nos principais sistemas internacionais e regionais de direitos humanos de que a liberdade de expressão, embora ampla, encontra limites claros e necessários nas manifestações que negam a dignidade intrínseca e a igualdade de grupos historicamente discriminados, especialmente quando há incitação à violência ou hostilidade.

No Brasil, a situação é peculiar. Embora o termo "discurso de ódio" não esteja expressamente positivado em uma lei específica que o defina de forma abrangente, ele tem sido cada vez mais incorporado pela jurisprudência dos tribunais superiores, pela doutrina especializada e pela atuação do Ministério Público para caracterizar manifestações que extrapolam os limites do dissenso legítimo, da crítica ácida ou da opinião controversa, configurando condutas discriminatórias vedadas pela Constituição e por leis infraconstitucionais. A Constituição Federal de 1988 estabelece um sólido fundamento para a repressão ao discurso de ódio, ao elencar como objetivos fundamentais da República a construção de uma sociedade livre, justa e solidária (art. 3º, I) e a promoção do bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação (art. 3º, IV). Além disso, como já mencionado, o artigo 5º, inciso XLII, classifica o racismo como crime inafiançável e imprescritível, e o inciso XLIV pune a ação de grupos

armados contra a ordem constitucional e o Estado Democrático. Essas normas constitucionais, interpretadas à luz do princípio da dignidade da pessoa humana (art. 1º, III), fornecem a base axiológica para a vedação de discursos que atentem contra a igualdade e a própria estrutura democrática.

No plano infraconstitucional, a Lei nº 7.716/1989 (Lei Caó) define os crimes resultantes de preconceito de raça ou de cor, punindo diversas condutas discriminatórias, incluindo a prática, indução ou incitação à discriminação ou preconceito (art. 20). O Supremo Tribunal Federal, no julgamento da Ação Direta de Inconstitucionalidade por Omissão (ADO) nº 26/DF e do Mandado de Injunção nº 4.733/DF, em 2019, deu um passo significativo ao equiparar a homofobia e a transfobia ao crime de racismo, previsto na sobredita lei, até que o Congresso Nacional legisle sobre a matéria.

No voto condutor, o Ministro Celso de Mello afirmou que “a homofobia e a transfobia, assim como o racismo, traduzem práticas odiosas, nefastas e repugnantes, que vulneram o postulado da dignidade humana e que são incompatíveis com os valores ético-jurídicos que informam e estruturam a própria Constituição da República” (BRASIL, STF, ADO 26/DF, 2019). Essa decisão ampliou o escopo de proteção contra o discurso de ódio no Brasil, reconhecendo a vulnerabilidade específica da população LGBTQIA+.

Outras legislações, como o Estatuto da Criança e do Adolescente (1990), o Estatuto do Idoso (2003) e a Lei Maria da Penha (2006), também contêm disposições que podem ser aplicadas para coibir discursos de ódio direcionados a esses grupos. O Marco Civil da Internet (Lei nº

12.965/2014), embora não trate especificamente de discurso de ódio, estabelece princípios como a proteção da privacidade e dos dados pessoais, e a responsabilização dos agentes de acordo com suas atividades, que são relevantes para o debate sobre moderação de conteúdo online.

Dessa forma, embora o Brasil ainda careça de um marco legal específico e sistematizado que defina claramente o conceito de discurso de ódio e estabeleça procedimentos para sua identificação e repressão, é inegável que o ordenamento jurídico nacional já oferece fundamentos constitucionais, legais e jurisprudenciais robustos para reprimir manifestações que violam a dignidade humana por meio da propagação do ódio, da intolerância e da discriminação contra minorias e grupos vulneráveis.

O desafio reside em aprimorar a legislação existente, conferir maior segurança jurídica na aplicação do conceito pelos tribunais e pelas plataformas digitais, e desenvolver estratégias eficazes que combinem medidas repressivas proporcionais com ações preventivas e educativas.

2.3.2 Diferenças entre Discurso Ofensivo, de Ódio e Discriminatório

A precisa delimitação conceitual entre as categorias de discurso ofensivo, discurso de ódio e discurso discriminatório revela-se como uma tarefa hermenêutica de fundamental importância para a aplicação justa e legítima dos limites à liberdade de expressão nas sociedades democráticas contemporâneas. Embora essas três formas de manifestação possam, em determinadas situações concretas, sobrepor-se ou coexistir – uma expressão pode ser simultaneamente ofensiva, discriminatória e incitadora

de ódio –, é importante reconhecer suas distinções analíticas. Nem toda expressão que causa ofensa ou desconforto pode ser classificada como discurso de ódio, tampouco toda manifestação que reflete ou reforça uma discriminação estrutural se enquadra automaticamente como um ilícito passível de sanção jurídica severa. Essa diferenciação cuidadosa é essencial não apenas para proteger os indivíduos e grupos vulneráveis contra os danos específicos causados pelo ódio, mas também para evitar que o Estado, sob o pretexto de combater a intolerância, cometa abusos na repressão à expressão, cerceando o dissenso legítimo e comprometendo a vitalidade do debate público pluralista, elemento indispensável à democracia.

O discurso meramente ofensivo, em primeiro lugar, consiste em manifestações que, embora possam causar indignação, repulsa, desconforto ou mágoa a determinados indivíduos ou grupos, não visam, primordial e diretamente, à exclusão social sistemática, à incitação à violência ou à negação da dignidade intrínseca de sujeitos com base em suas identidades coletivas. A jurisprudência internacional de direitos humanos, notadamente a da Corte Europeia de Direitos Humanos (CEDH), tem consistentemente reconhecido que a liberdade de expressão, nos termos do artigo 10 da Convenção Europeia, abrange não apenas as informações ou ideias recebidas favoravelmente ou consideradas inofensivas ou indiferentes, mas também aquelas “que ofendem, chocam ou perturbam o Estado ou qualquer setor da população”.

Essa proteção estende-se a ideias que desafiam o *status quo*, questionam crenças estabelecidas ou utilizam linguagem contundente,

satírica ou mesmo vulgar. Como afirmado no célebre caso *Handyside v. United Kingdom* (1976, par. 49), “tal são as demandas do pluralismo, da tolerância e da abertura de espírito, sem as quais não há ‘sociedade democrática’”. Nessa perspectiva, o dissenso político, a crítica artística, o humor ácido, o sarcasmo e até a linguagem considerada de mau gosto, embora possam ser subjetivamente ofensivos para alguns, são compreendidos como parte integrante e necessária do discurso político, social e cultural em uma sociedade aberta e pluralista. A ofensa, por si só, não é critério suficiente para justificar a restrição da liberdade de expressão, sob pena de se instaurar uma tirania da sensibilidade ou do politicamente correto que asfixiaria o debate crítico.

O discurso discriminatório, por sua vez, situa-se em um patamar distinto. Caracteriza-se por manifestações, práticas ou representações que promovem, reforçam, justificam ou perpetuam tratamentos desiguais e hierarquias sociais injustas, frequentemente baseadas em categorias suspeitas como raça, etnia, gênero, religião, orientação sexual, identidade de gênero, deficiência, origem nacional ou condição socioeconômica (MOREIRA, 2020).

A discriminação discursiva pode ocorrer de forma direta e explícita (por exemplo, a recusa de um serviço com base na cor da pele) ou, mais frequentemente, de forma indireta, sutil e até mesmo inconsciente. Ela pode manifestar-se não apenas através de linguagem agressiva, mas também pela omissão, pela ausência sistemática de representatividade de certos grupos na mídia, pela reiteração de estereótipos negativos que associam grupos a características indesejáveis (criminalidade, preguiça,

incapacidade), ou pela adoção de critérios aparentemente neutros que, na prática, produzem um impacto desproporcionalmente adverso sobre grupos específicos. Como observa Pierre Bourdieu (1989) em sua análise da violência simbólica.

É uma violência que se exerce com o consentimento tácito, e muitas vezes inconsciente, daqueles que a sofrem e também daqueles que a exercem, na medida em que uns e outros não percebem que a exercem ou que a sofrem” (BOURDIEU, 1989, p. 18).

Assim, discursos discriminatórios nem sempre são evidentes ou intencionalmente hostis, mas podem operar de maneira insidiosa e persistente na manutenção de estruturas de poder e exclusão. O combate ao discurso discriminatório envolve não apenas a repressão a atos explícitos, mas também a implementação de políticas públicas, ações afirmativas e estratégias educativas que visem desconstruir preconceitos e promover a igualdade material.

Já o discurso de ódio (*hate speech*) representa uma forma qualificada e particularmente grave de manifestação discriminatória. Sua especificidade reside na intenção (explícita ou inferida do contexto) ou no efeito provável de incitar, promover, justificar ou encorajar o ódio, a hostilidade, a discriminação ou a violência contra indivíduos ou grupos identificáveis com base em suas características protegidas. Ele possui um conteúdo não apenas excludente ou depreciativo, mas também persecutório, inflamatório e desumanizador, mobilizando emoções intensas de repulsa, medo, desprezo ou aversão. O discurso de ódio ataca a própria condição de igualdade e pertença dos membros do grupo-alvo à comunidade política, tratando-os como "outros" indesejáveis, inferiores ou

perigosos. A Corte Interamericana de Direitos Humanos, como visto, diferencia claramente o discurso de ódio das demais formas de manifestação ao afirmar que “expressões que constituam incitação direta e pública à violência ou qualquer outra ação ilegal similar contra qualquer pessoa ou grupo de pessoas por motivos de raça, cor, religião, idioma ou origem nacional [...] ou que façam apologia ao ódio nacional, racial ou religioso que constitua incitação à discriminação, à hostilidade, ao crime ou à violência” não são protegidas pela liberdade de expressão (CIDH, OC-5/85, 1985; CIDH). O elemento chave aqui é a *incitação* – a conclamação, ainda que implícita, a uma ação ou atitude hostil por parte de terceiros contra o grupo visado.

A principal diferença, portanto, entre esses três tipos de discurso reside no grau de intensidade da lesão ou do risco de lesão aos direitos fundamentais alheios e na finalidade ou função pragmática subjacente à manifestação. Enquanto o discurso meramente ofensivo, embora possa ferir suscetibilidades, geralmente se insere no âmbito do debate de ideias (ainda que de forma rude) e goza de ampla proteção constitucional, o discurso discriminatório, ao perpetuar desigualdades, já justifica a intervenção estatal por meio de políticas antidiscriminatórias e, em certos casos, sanções civis ou administrativas, mas nem sempre exige a repressão penal. O discurso de ódio, por sua vez, ao incitar ativamente a hostilidade e a violência contra grupos vulneráveis, representa uma ameaça mais direta e grave à dignidade humana, à igualdade e à própria coesão social, podendo justificar, como último recurso, sanções jurídicas mais severas, inclusive de natureza penal. Contudo, essa repressão penal deve ser

reservada aos casos mais graves e inequívocos, para evitar a banalização do direito penal e a criminalização indevida da expressão.

O julgamento da ADPF 187/DF pelo Supremo Tribunal Federal, que reconheceu a constitucionalidade das manifestações públicas em defesa da descriminalização das drogas (a chamada "Marcha da Maconha"), ilustra bem a proteção conferida ao discurso impopular ou mesmo ofensivo para alguns setores, desde que não descambe para a apologia direta ao crime ou incitação ao consumo. No voto do Ministro Ayres Britto, ficou claro que “não há liberdade de expressão só para as ideias simpáticas, ou que gozam de consenso social. A democracia pressupõe o dissenso, a divergência, a possibilidade de se defenderem posições minoritárias ou contramajoritárias” (BRASIL, STF, ADPF 187/DF, julgado em 15 jun. 2011). Esse entendimento reforça que a crítica a leis existentes ou a defesa de mudanças legislativas, mesmo sobre temas controversos, faz parte do jogo democrático e não pode ser confundida, *a priori*, com discurso de ódio ou incitação ao ilícito. O Estado deve atuar com máxima moderação (*self-restraint*) para não inviabilizar o debate público legítimo sobre políticas públicas.

Ademais, a análise da natureza de um discurso (ofensivo, discriminatório ou de ódio) não pode prescindir de uma avaliação cuidadosa do *contexto* comunicacional em que ele se insere. Fatores como o local da manifestação (um debate acadêmico, uma rede social, um comício político, uma conversa privada), o público-alvo, o momento histórico e social, a relação de poder entre o emissor e o receptor, e a posição de autoridade ou influência do emissor são cruciais para

determinar o significado e o impacto potencial da mensagem. Uma piada considerada inofensiva em um círculo íntimo pode adquirir contornos discriminatórios ou de ódio se proferida por uma figura pública em um meio de comunicação de massa.

É apenas o conteúdo semântico ou literal da mensagem que importa para o direito, mas sua função pragmática no contexto social em que é proferida, os atos de fala que ela realiza” (MAFEI, 2016, p. 74).

Essa análise contextual é particularmente desafiadora no ambiente digital, onde as mensagens podem ser facilmente descontextualizadas, viralizadas e amplificadas por algoritmos, potencializando seus efeitos nocivos.

Por todas essas razões, a distinção entre discurso ofensivo, discurso discriminatório e discurso de ódio, embora conceitualmente necessária, exige na prática uma aplicação criteriosa e sensível ao contexto, guiada por parâmetros técnicos e jurídicos sólidos, baseados na hermenêutica constitucional e nos padrões internacionais de direitos humanos. Essa diferenciação é vital não apenas para assegurar a proteção adequada aos grupos vulnerabilizados contra as diversas formas de violência discursiva, mas também para preservar o núcleo essencial da liberdade de expressão como pilar da democracia, evitando que o legítimo combate ao discurso de ódio seja desvirtuado e instrumentalizado para fins de censura política, perseguição ideológica ou silenciamento do dissenso.

2.3.3 A dignidade da pessoa humana como limite à manifestação do pensamento

No complexo edifício do Estado Constitucional de Direito, caracterizado pela interdependência e, por vezes, pela tensão entre múltiplos princípios e direitos fundamentais, a liberdade de manifestação do pensamento encontra sua principal e mais robusta fronteira jurídico-axiológica no princípio da dignidade da pessoa humana. Previsto expressamente como um dos fundamentos da República Federativa do Brasil no artigo 1º, inciso III, da Constituição de 1988, este princípio transcende a condição de mero postulado ético-filosófico abstrato, assumindo a natureza de norma jurídica fundamental, dotada de densidade axiológica própria e de uma eficácia irradiadora (*Ausstrahlungswirkung*) que permeia e informa a interpretação e aplicação de todo o ordenamento jurídico, em especial o catálogo de direitos fundamentais.

A dignidade da pessoa humana, em sua acepção consolidada pela dogmática constitucional brasileira e internacional, remonta a raízes filosóficas profundas, notadamente já citada ética kantiana, que postula que o ser humano deve ser sempre tratado como um fim em si mesmo, jamais como mero meio ou instrumento para a consecução de outros fins (KANT, 1986). Transportada para o plano jurídico-constitucional, a dignidade configura-se como o valor-fonte do sistema, o núcleo normativo intangível que confere unidade e sentido a todos os direitos e garantias. Ela exige do Estado e da sociedade não apenas uma postura de abstenção, de não praticar condutas que violem a integridade física, psíquica e moral dos indivíduos (dimensão negativa ou defensiva), mas também uma atuação

positiva, de promoção ativa das condições materiais e existenciais necessárias para o pleno desenvolvimento da personalidade humana em sua individualidade e em sua dimensão social e comunitária (dimensão positiva ou prestacional). Como observa com precisão Ingo Wolfgang Sarlet (2021):

A dignidade da pessoa humana possui um caráter normativo vinculante e opera simultaneamente como critério de validade material para os atos normativos e as políticas públicas, como limite intransponível ao exercício de quaisquer liberdades individuais ou coletivas, e como diretriz interpretativa para todo o ordenamento jurídico” (SARLET, 2021, p. 215).

A dignidade, portanto, não é apenas *um* direito fundamental, mas o próprio fundamento e o fim último de todos os direitos fundamentais. No campo específico da colisão entre a liberdade de expressão e outros bens jurídicos, a dignidade da pessoa humana emerge como o limite material por excelência. Ela impõe restrições legítimas quando a manifestação do pensamento deixa de ser um exercício crítico, informativo ou artístico, inserido no fluxo do debate democrático, e passa a configurar uma prática de exclusão, humilhação, estigmatização, desumanização ou negação da alteridade. O direito de expressar livremente as próprias ideias encontra, assim, limites imanescentes que não decorrem de uma censura externa ou de uma tutela paternalista imposta pelo Estado, mas da própria necessidade intrínseca de preservar a ordem constitucional democrática, que se funda no reconhecimento da igual dignidade de todos os seres humanos e na vedação peremptória de qualquer forma de discriminação atentatória a essa dignidade (art. 3º, IV, CF/88). Como afirma Daniel Sarmento, a dignidade funciona como um "superprincípio" que orienta a ponderação em casos de

conflito, conferindo um peso especial à proteção de grupos vulneráveis contra discursos que historicamente serviram para justificar sua opressão (SARMENTO, 2016).

A dogmática contemporânea dos direitos fundamentais, especialmente a partir da influente teoria da ponderação desenvolvida por Robert Alexy (2015), fornece os instrumentos metodológicos adequados para compreender e operacionalizar essa limitação. Os direitos fundamentais, em sua maioria, possuem a estrutura de princípios – normas que ordenam que algo seja realizado na maior medida possível, dentro das possibilidades fáticas e jurídicas existentes. Quando dois princípios colidem em um caso concreto (como a liberdade de expressão e a dignidade/honra de outrem), a solução não se dá pela invalidação de um deles, mas pela ponderação, que busca estabelecer uma relação de precedência condicionada entre eles, com base nos critérios de adequação (a restrição a um princípio deve ser apta a promover o outro), necessidade (a restrição deve ser a menos gravosa possível entre as igualmente eficazes) e proporcionalidade em sentido estrito (a importância de satisfazer o princípio que prevalece deve ser maior que a intensidade da restrição imposta ao princípio que cede). Aplicando-se esse raciocínio, pode-se afirmar que a liberdade de expressão, embora dotada de uma posição preferencial (*prima facie*) na ordem constitucional devido ao seu papel instrumental para a democracia, pode e deve ser restringida sempre que essa restrição se mostre adequada, necessária e proporcional em sentido estrito para proteger de forma eficaz a dignidade de pessoas ou grupos vulnerabilizados contra ataques graves (ALEXY, 2015).

Se faz mister sublinhar, contudo, que a invocação da dignidade da pessoa humana como limite à liberdade de expressão não pode ser banalizada ou utilizada de forma arbitrária para silenciar críticas legítimas ou opiniões meramente impopulares. A proteção jurídica da dignidade não se confunde com um apelo subjetivista à tutela contra sentimentos de ofensa, desconforto ou indignação moral. A dignidade, enquanto limite objetivo à liberdade de expressão, não se presta a proteger a suscetibilidade individual exacerbada ou a impor um código de conduta puramente moralista ao debate público. Ela visa, antes, à salvaguarda de condições mínimas e objetivas de respeito mútuo, reconhecimento e inclusão no espaço público, protegendo os indivíduos de ataques que neguem sua igual humanidade, sua autonomia ou sua integridade moral de forma grave e sistemática. Como bem adverte Luís Roberto Barroso, é preciso ter cautela.

Não confundir discurso meramente ofensivo, chocante ou imoral com discurso intrinsecamente inconstitucional. Só este último, que configura um grave atentado à dignidade humana ou uma incitação clara à violência ou discriminação, pode ser legitimamente reprimido pelo Direito” (BARROSO, 2017, p. 129).

Nesse sentido, a aplicação concreta do princípio da dignidade como fator limitador da liberdade de expressão demanda uma fundamentação jurídica rigorosa e contextualizada. É preciso demonstrar, em cada caso, que a manifestação em questão não apenas causa desconforto ou ofensa subjetiva, mas atinge o núcleo essencial da identidade, da honra objetiva ou da integridade moral do indivíduo ou grupo atingido, de forma sistemática, intencional (ou com grave negligência quanto ao seu potencial danoso) e com potencial lesivo concreto no contexto social em que se

insere. A já citada decisão do STF na ADO 26/DF, relativa à criminalização da homofobia e da transfobia, ilustra essa preocupação ao fundamentar a necessidade da intervenção penal não em meras ofensas, mas no histórico de violência e discriminação estrutural sofrido pela população LGBTQIA+ no Brasil, reconhecendo que discursos de ódio contra esse grupo contribuem para a perpetuação dessa violência. O Ministro Celso de Mello, em seu voto, enfatizou que:

A proteção constitucional da liberdade de expressão não pode converter-se em criminoso salvo-conduto para a prática de atos de ódio, intolerância e discriminação que atentem contra a dignidade e a integridade física e psíquica de pessoas LGBTQIA+ (BRASIL, STF, ADO 26/DF, 2019).

A dignidade da pessoa humana consolida-se, no constitucionalismo democrático contemporâneo e, em particular, no ordenamento brasileiro pós-1988, como o limite material substancial por excelência à manifestação do pensamento. Ela atua como uma cláusula de contenção jurídica e ética indispensável, especialmente nos casos em que a expressão resvala para o discurso de ódio, a incitação à violência, a negação da igualdade ou a promoção da exclusão de grupos vulneráveis. Essa limitação não representa uma antinomia ou uma contradição à importância da liberdade de expressão, mas sim uma condição de sua própria legitimidade e sustentabilidade em uma sociedade pluralista. O objetivo não é inibir o dissenso, a crítica ou a diversidade de opiniões, mas sim preservar o próprio ambiente de respeito recíproco e reconhecimento da igual humanidade sobre o qual se assenta a possibilidade de um discurso público verdadeiramente democrático. A complexa responsabilidade hermenêutica que recai sobre os intérpretes e aplicadores do direito é,

portanto, a de traçar, caso a caso e com base em critérios objetivos e ponderados, a delicada fronteira entre a liberdade legítima e o abuso intolerável, entre a expressão protegida e a opressão discursiva, entre a crítica necessária e a humilhação vedada.

2.3.4 Tipificação Penal Brasileira

A resposta do sistema penal ao fenômeno complexo e multifacetado do discurso de ódio no ordenamento jurídico brasileiro apresenta-se, de maneira inequívoca, como fragmentada, insuficiente e marcada por significativas lacunas. Essa situação reflete, em grande medida, a ausência de um marco normativo específico e abrangente que defina com clareza e precisão o que se entende por discurso de ódio no contexto nacional e quais devem ser suas consequências jurídicas, especialmente na esfera criminal, considerada a *ultima ratio* da intervenção estatal. Embora a Constituição Federal de 1988 estabeleça bases sólidas para a proteção da dignidade humana (art. 1º, III) e da igualdade (art. 5º, *caput*), e vede expressamente manifestações racistas (art. 5º, XLII), a legislação penal infraconstitucional ainda opera com tipos penais dispersos e, em muitos casos, inadequados para capturar a especificidade e a gravidade dos discursos que incitam à violência, à discriminação ou à hostilidade contra grupos historicamente marginalizados e vulnerabilizados na sociedade brasileira.

A principal norma penal historicamente mobilizada para o enfrentamento de manifestações discriminatórias é a Lei n.º 7.716/1989, conhecida como Lei Caó, que define os crimes resultantes de preconceito

de raça ou de cor. Originalmente concebida para combater o racismo em suas diversas manifestações, a abrangência desta lei foi objeto de uma reinterpretação extensiva e histórica por parte do Supremo Tribunal Federal (STF) no julgamento conjunto da Ação Direta de Inconstitucionalidade por Omissão (ADO) n.º 26/DF e do Mandado de Injunção (MI) n.º 4.733/DF, em 2019. Nessa ocasião, a Corte reconheceu a mora inconstitucional do Poder Legislativo em criminalizar especificamente a homofobia e a transfobia e, por maioria expressiva, determinou que as condutas homofóbicas e transfóbicas se enquadrassem, por analogia, nos diversos tipos penais definidos na Lei nº 7.716/89, até que o Congresso Nacional edite legislação específica sobre a matéria. O voto condutor do Ministro Celso de Mello ressaltou que:

A omissão legislativa em criminalizar tais condutas representa uma grave e inaceitável falha do Estado brasileiro em cumprir seu dever constitucional de proteger a dignidade e a segurança de pessoas LGBTQIA+ contra atos de violência e discriminação” (STF, 2019).

Esse julgamento representou um importante avanço na proteção de grupos vulneráveis, mas também suscitou debates sobre os limites da interpretação judicial em matéria penal, dada a estrita vinculação ao princípio da legalidade (*nullum crimen, nulla poena sine praevia lege poenali*).

Apesar desse importante precedente jurisprudencial, a ausência de uma tipificação penal autônoma e bem delineada para o discurso de ódio como fenômeno comunicativo específico continua a gerar considerável insegurança jurídica e dificuldades operacionais significativas na repressão eficaz dessas práticas. A própria redação dos tipos penais

previstos na Lei n.º 7.716/1989, como o artigo 20, que pune a prática, indução ou incitação à discriminação ou preconceito, exige frequentemente a demonstração de elementos subjetivos específicos do agente (o dolo específico de discriminar ou incitar) e a caracterização explícita de uma motivação baseada em raça, cor, etnia, religião ou procedência nacional (agora estendida à orientação sexual e identidade de gênero pela decisão do STF). Essa exigência probatória rigorosa, aliada à dificuldade de comprovar a intenção discriminatória em discursos mais velados, irônicos, simbólicos ou que se valem de ambiguidades, limita consideravelmente a eficácia da lei diante das formas contemporâneas e muitas vezes sutis de propagação do ódio, especialmente no ambiente digital. A necessidade de provar o dolo específico pode levar à impunidade de discursos objetivamente danosos, mas cuja intenção maliciosa é difícil de comprovar cabalmente (MEYER-PFLUG, 2009).

Adicionalmente, outros dispositivos do Código Penal que poderiam, em tese, ser aplicados – como os crimes contra a honra (calúnia, difamação e injúria, previstos nos arts. 138 a 140), a incitação ao crime (art. 286) e a apologia de crime ou criminoso (art. 287) – mostram-se frequentemente inadequados ou insuficientes para lidar com a complexidade e a escala do discurso de ódio na era digital. Os crimes contra a honra, por exemplo, são tradicionalmente concebidos para proteger a reputação de indivíduos específicos, sendo de difícil aplicação quando o discurso de ódio se dirige a grupos inteiros de forma genérica e impessoal, atacando sua dignidade coletiva mais do que a honra individual de seus membros. A injúria racial (art. 140, § 3º, CP), embora mais

próxima, ainda exige a ofensa direta à dignidade ou decoro de pessoa determinada, utilizando elementos referentes à raça, cor, etnia, religião, origem ou condição de pessoa idosa ou com deficiência. Já os crimes de incitação e apologia ao crime pressupõem uma relação mais direta com a prática de um delito específico, o que nem sempre ocorre no discurso de ódio, que pode incitar à hostilidade e à discriminação de forma mais difusa, sem necessariamente conclamar a um crime determinado. Essa inadequação tipológica deixa muitas manifestações odiosas em uma zona cinzenta de impunidade penal (BITENCOURT, 2021).

A ausência de uma legislação penal específica e atualizada também impacta negativamente a responsabilização das plataformas digitais (redes sociais, aplicativos de mensagens, motores de busca), que desempenham um papel primordial como intermediárias, amplificadoras e, por vezes, curadoras desses discursos. O Marco Civil da Internet (Lei n.º 12.965/2014), embora estabeleça princípios importantes para o uso da rede no Brasil, adota um regime de responsabilidade civil subjetiva para os provedores de aplicações de internet (como redes sociais) por danos decorrentes de conteúdo gerado por terceiros. Salvo em casos específicos (como violação de direitos autorais ou divulgação não consentida de imagens íntimas), a plataforma só pode ser responsabilizada civilmente se, após ordem judicial específica, não tomar as providências para tornar indisponível o conteúdo apontado como infringente (art. 19). Esse modelo, conhecido como *notice and judicial takedown*, embora vise proteger a liberdade de expressão e evitar a censura privada em larga escala, tem sido criticado por sua lentidão e por potencialmente contribuir para a

impunidade de discursos manifestamente ilícitos, como o de ódio, que podem se alastrar rapidamente e causar danos irreparáveis antes que uma ordem judicial seja obtida e cumprida.

A arquitetura normativa da internet brasileira, focada na responsabilidade *ex post* e mediante ordem judicial, ainda não foi suficientemente adaptada para responder aos desafios específicos do discurso odioso algorítmico, que opera por meio de viralização automatizada, microdirecionamento e exploração de vieses cognitivos (MAFEI, 2016, p. 142).

Em contraste com a abordagem brasileira, diversos outros ordenamentos jurídicos têm adotado posturas mais proativas e específicas para lidar com o discurso de ódio online e a responsabilidade das plataformas. Na Alemanha, a Lei de Aplicação da Rede (Netzwerkdurchsetzungsgesetz - NetzDG), em vigor desde 2017, impõe obrigações rigorosas às grandes redes sociais (com mais de 2 milhões de usuários registrados na Alemanha) para remover ou bloquear o acesso a conteúdo "manifestamente ilegal" (incluindo incitação ao ódio, negação do Holocausto, etc.) no prazo de 24 horas após o recebimento de uma notificação, sob pena de multas que podem chegar a 50 milhões de euros. Conteúdos não manifestamente ilegais devem ser avaliados e removidos em até 7 dias. A lei também exige que as plataformas mantenham procedimentos eficazes de gestão de reclamações e publiquem relatórios de transparência detalhados. Embora criticada por alguns por potencialmente incentivar a remoção excessiva de conteúdo (*overblocking*) e delegar poderes de censura a empresas privadas, a NetzDG representa um marco na tentativa de responsabilizar as plataformas pela disseminação de conteúdo ilícito. No Reino Unido, o ambicioso Online Safety Act, aprovado em 2023 após longo debate,

estabelece um "dever de cuidado" (*duty of care*) para as plataformas online em relação à segurança de seus usuários, especialmente crianças.

A lei cria um órgão regulador independente (Ofcom) com amplos poderes para fiscalizar o cumprimento desse dever, exigindo que as plataformas implementem sistemas e processos robustos para identificar, avaliar e mitigar os riscos associados a conteúdos ilegais (como terrorismo, abuso infantil e discurso de ódio criminalizado) e também a conteúdos considerados "nocivos" para crianças (como pornografia, automutilação, suicídio, distúrbios alimentares). O não cumprimento pode acarretar multas de até 10% do faturamento global anual da empresa. Essas experiências estrangeiras, embora com modelos e intensidades distintas, apontam para uma tendência global de maior regulação e responsabilização das plataformas digitais no combate a conteúdos prejudiciais, incluindo o discurso de ódio.

No Brasil, diversas propostas legislativas têm sido apresentadas nos últimos anos com o objetivo de preencher a lacuna existente na tipificação do discurso de ódio e na regulação das plataformas. O Projeto de Lei (PL) 2.370/2019, por exemplo, propõe a alteração da Lei nº 7.716/89 para incluir um tipo penal específico de "incitação ao ódio", definido como "incitar, publicamente, por qualquer meio, inclusive pela rede mundial de computadores, o ódio ou a aversão contra grupo de pessoas por motivo de raça, cor, etnia, religião, procedência nacional, orientação sexual ou identidade de gênero". Já o PL 2630/2020, conhecido como "PL das Fake News" ou Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet, embora não tipifique

diretamente o discurso de ódio, estabelece uma série de obrigações para provedores de redes sociais e de mensagens privadas, incluindo deveres de transparência sobre publicidade e impulsionamento de conteúdo, moderação de conteúdo, identificação de contas automatizadas e representação legal no Brasil.

O projeto também prevê a criação de um órgão de autorregulação e estabelece sanções administrativas para o descumprimento das obrigações. Ambas as propostas, no entanto, enfrentaram (e ainda enfrentam) forte resistência e intenso debate público, com setores expressando preocupações legítimas sobre os riscos de censura, a vagueza de conceitos, o impacto sobre a liberdade de expressão e a inovação, e a potencial criação de mecanismos de vigilância estatal ou privada. Essa dificuldade em alcançar um consenso legislativo revela a complexidade de se encontrar o ponto de equilíbrio adequado entre a necessidade de proteger grupos vulneráveis contra discursos nocivos e a salvaguarda do espaço democrático de crítica, debate e dissenso.

Portanto, é possível afirmar com segurança que o ordenamento jurídico brasileiro, apesar de dispor de alguns instrumentos jurídicos parciais e de contar com uma jurisprudência atuante do STF em certos aspectos, ainda carece de uma tipificação penal clara, sistemática, proporcional e eficaz para o discurso de ódio, especialmente em sua manifestação online. Essa ausência de normatização específica e adequada compromete não apenas a efetividade da repressão penal, dificultando investigações, processos e condenações, mas também enfraquece a função simbólica e preventiva da lei penal na reafirmação dos valores

fundamentais da Constituição, notadamente a dignidade da pessoa humana, a igualdade material e o pluralismo democrático. A superação dessa lacuna exige um debate legislativo qualificado, informado pela doutrina, pela jurisprudência comparada e pelos padrões internacionais de direitos humanos, que seja capaz de construir soluções normativas que conciliem, de forma equilibrada e proporcional, a proteção contra o ódio e a garantia da liberdade.

2.3.5 Parâmetros do Sistema Interamericano e Europeu de Direitos Humanos

A compreensão aprofundada do fenômeno do discurso de ódio e da complexa tarefa de delinear seus limites jurídicos no ordenamento brasileiro requer, indispensavelmente, um diálogo constante e atento com os desenvolvimentos normativos e jurisprudenciais dos sistemas regionais de proteção aos direitos humanos, com destaque para o Sistema Interamericano e o Sistema Europeu. Estes sistemas, por meio de suas convenções fundadoras e da atuação de suas respectivas cortes – a Corte Interamericana de Direitos Humanos (Corte IDH) e o Tribunal Europeu de Direitos Humanos (TEDH) –, oferecem um corpo robusto de jurisprudência consolidada, princípios orientadores e práticas normativas avançadas que não apenas refletem um consenso regional sobre os padrões mínimos de proteção, mas também contribuem significativamente para a construção de parâmetros interpretativos e para a evolução do direito interno dos Estados-membros, incluindo o Brasil.

Na Convenção Americana sobre Direitos Humanos (CADH), a liberdade de pensamento e expressão encontra-se consagrada no artigo 13.

Este artigo, em seu *caput*, garante a toda pessoa o direito à liberdade de pensamento e de expressão, compreendendo a liberdade de buscar, receber e difundir informações e ideias de toda natureza, sem consideração de fronteiras, verbalmente ou por escrito, ou em forma impressa ou artística, ou por qualquer outro processo de sua escolha. Contudo, a própria Convenção estabelece, de forma explícita e inequívoca, que essa liberdade não é absoluta. O parágrafo 5º do artigo 13 representa uma das cláusulas mais contundentes contra o discurso de ódio no direito internacional regional, ao determinar que “a lei deve proibir toda propaganda a favor da guerra, bem como toda apologia ao ódio nacional, racial ou religioso que constitua incitamento à discriminação, à hostilidade, ao crime ou à violência” (OEA, 1969). Essa disposição não apenas permite, mas *obriga* os Estados Partes a adotarem medidas legislativas para coibir as formas mais graves de discurso de ódio, aquelas que atingem o limiar da incitação.

A Corte Interamericana de Direitos Humanos, como intérprete última da Convenção Americana, tem consistentemente reafirmado e densificado essa limitação em sua jurisprudência. No caso *Kimel vs. Argentina* (2008), embora tratando de crime contra a honra, a Corte reiterou que a liberdade de expressão pode ser sujeita a responsabilidades ulteriores, que devem ser necessárias e proporcionais para assegurar o respeito aos direitos ou à reputação dos demais.

No caso *Usón Ramírez vs. Venezuela* (2009), a Corte analisou a condenação de um militar por ofensas às Forças Armadas e, embora tenha concluído pela violação da liberdade de expressão no caso concreto, reafirmou que discursos que incitem à violência ou que neguem direitos

fundamentais não estão protegidos. Mais recentemente, no caso *Álvarez Ramos vs. Venezuela* (2019), a Corte foi explícita ao analisar declarações de altas autoridades estatais contra um canal de televisão, destacando que a liberdade de expressão deve ser exercida em conformidade com os valores democráticos e o respeito aos direitos humanos, o que exclui manifestações que incitem à violência, à discriminação ou que desumanizem pessoas ou grupos com base em suas identidades. (CIDH, 2019).

A Relatoria Especial para a Liberdade de Expressão da Comissão Interamericana de Direitos Humanos também tem produzido importantes relatórios temáticos sobre discurso de ódio, oferecendo parâmetros adicionais para sua identificação e combate.

O Sistema Europeu de Direitos Humanos, estruturado em torno da Convenção Europeia de Direitos Humanos (CEDH, 1950), adota uma abordagem similar, embora com nuances próprias. O artigo 10 da CEDH garante a liberdade de expressão (parágrafo 1º), mas prevê, em seu parágrafo 2º, a possibilidade de restrições, desde que previstas em lei, necessárias em uma sociedade democrática e voltadas à proteção de interesses legítimos, como a segurança nacional, a integridade territorial, a segurança pública, a defesa da ordem e prevenção do crime, a proteção da saúde ou da moral, a proteção da reputação ou dos direitos de outrem, entre outros.

A jurisprudência do Tribunal Europeu de Direitos Humanos (TEDH), sediado em Estrasburgo, tem sido particularmente rica e influente na delimitação dos contornos da liberdade de expressão frente ao discurso

de ódio. O TEDH tem sido firme ao excluir da proteção do artigo 10 os discursos que promovem ou incitam ao ódio racial, étnico, religioso, xenófobo ou homofóbico, ou que negam crimes contra a humanidade. No emblemático caso *Garaudy v. France* (2003), o Tribunal declarou inadmissível a queixa de um autor condenado por negar o Holocausto em seus escritos, afirmando que a negação ou a minimização de crimes contra a humanidade, especialmente quando visa reabilitar ideologias totalitárias ou incitar ao ódio contra grupos que foram vítimas desses crimes, constitui um abuso do direito à liberdade de expressão (nos termos do artigo 17 da CEDH) e, portanto, não merece proteção (TEDH, 2003). Em *Jersild v. Denmark* (1994), contudo, o Tribunal protegeu um jornalista que exibiu entrevista com membros de um grupo racista, por entender que o contexto da reportagem era informativo e crítico, não endossando as opiniões odiosas. Isso demonstra a importância da análise contextual na jurisprudência europeia.

Em sua análise, o TEDH frequentemente emprega a doutrina da “margem de apreciação” (*margin of appreciation*), que reconhece aos Estados parte uma certa latitude ou discricionariedade para avaliar a necessidade e a proporcionalidade das restrições à liberdade de expressão, levando em conta suas particularidades históricas, culturais e sociais.

Essa margem, contudo, não é ilimitada e está sujeita ao controle do próprio Tribunal, sendo mais estreita em matéria de discurso político e mais ampla em áreas como a proteção da moralidade. A aplicação dessa doutrina ao discurso de ódio é complexa, pois, ao mesmo tempo em que se reconhece a necessidade de os Estados combaterem o ódio, também se

busca evitar restrições excessivas que silenciem o debate. O Tribunal Europeu enfatiza que qualquer restrição deve ser prevista em lei clara e acessível, perseguir um fim legítimo e ser “necessária em uma sociedade democrática”, o que implica um teste rigoroso de proporcionalidade entre a medida restritiva e o objetivo almejado (VAN DIJK et al., 2018).

A principal contribuição desses sistemas regionais para o debate brasileiro reside, portanto, em oferecer um parâmetro normativo e jurisprudencial externo, objetivamente comprometido com os mais altos padrões internacionais de direitos humanos, que condiciona inequivocamente o exercício da liberdade de expressão aos valores fundamentais da dignidade humana, da igualdade substancial e da não discriminação. Eles demonstram um consenso regional robusto de que discursos que incitam ao ódio, à violência ou à hostilidade contra grupos vulneráveis não encontram guarida no direito à liberdade de expressão e podem (e em alguns casos, devem) ser restringidos ou punidos pelo Estado. O Brasil, como signatário da Convenção Americana desde 1992 e reconhecedor da jurisdição contenciosa da Corte Interamericana desde 1998, encontra-se juridicamente vinculado a essas diretrizes. Isso impõe aos operadores do direito nacional – legisladores, administradores e, especialmente, juízes – a obrigação de interpretar e aplicar o direito interno, inclusive a Constituição, de forma harmônica com os *padrões* estabelecidos pela Convenção Americana e pela jurisprudência da Corte IDH. Esse dever decorre do princípio do controle de convencionalidade, que exige que as normas e práticas internas sejam compatíveis com as obrigações internacionais assumidas pelo Estado em matéria de direitos

humanos.

A jurisprudência da Corte IDH, portanto, não é mera fonte de inspiração, mas sim um critério vinculante para a aferição da validade e da legitimidade das ações estatais, inclusive no que tange à regulação da liberdade de expressão (MAZZUOLI, 2021). Ignorar esses parâmetros não apenas enfraquece a proteção dos direitos humanos no plano interno, mas também expõe o Brasil à responsabilidade internacional por omissão ou ação incompatível com suas obrigações convencionais.

O diálogo com os Sistemas Interamericano e Europeu de Direitos Humanos é essencial para enriquecer a compreensão e a aplicação dos limites da liberdade de expressão no Brasil. Suas convenções e, sobretudo, a rica jurisprudência de suas Cortes oferecem balizas interpretativas importantes, que reforçam a centralidade da dignidade humana e da igualdade como limites materiais ao discurso, ao mesmo tempo em que alertam para a necessidade de critérios claros, previsibilidade e proporcionalidade na imposição de quaisquer restrições. A incorporação efetiva desses parâmetros internacionais no direito e na prática jurídica brasileira é um passo fundamental para consolidar uma cultura de respeito aos direitos humanos e para enfrentar, de forma legítima e eficaz, os desafios contemporâneos impostos pelo discurso de ódio.

2.4 REDES SOCIAIS, PLATAFORMAS DIGITAIS E RESPONSABILIZAÇÃO

2.4.1 Redes Sociais como Arenas Públicas Contemporâneas: Evolução, Transformações e Desafios

A compreensão das redes sociais digitais como as arenas públicas predominantes na contemporaneidade exige um mergulho analítico que transcenda a mera constatação de sua popularidade, demandando um resgate histórico-conceitual da noção de esfera pública para evidenciar tanto suas continuidades quanto, e principalmente, suas profundas transformações estruturais no ambiente digital. Desde a Antiguidade clássica, os espaços designados como públicos foram concebidos como locais de visibilidade, interação social, confronto de ideias e formação de reconhecimento coletivo, ainda que, historicamente, tenham sido marcados por profundas exclusões e restrições de acesso a segmentos privilegiados da população.

Na *polis* grega, a *ágora* funcionava como o epicentro da vida cívica e política, o local onde os cidadãos – uma categoria restrita a homens livres, proprietários e nascidos na cidade – debatiam os rumos da comunidade, exerciam a retórica e buscavam a persuasão como ferramentas essenciais do discurso público. Contudo, esse idealizado espaço de deliberação era intrinsecamente excludente, negando voz e participação a mulheres, escravos, estrangeiros e não proprietários, refletindo as hierarquias sociais e de poder da época (ARENDDT, 1997).

Durante a longa Idade Média europeia, a centralidade do espaço público, tal como concebido na Antiguidade, foi significativamente deslocada e reconfigurada, passando a orbitar em torno de instituições eclesiásticas e das cortes monárquicas. O debate público tornou-se

limitado, fragmentado e profundamente subordinado à autoridade religiosa e ao poder temporal dos soberanos. A disseminação de ideias consideradas divergentes ou heterodoxas era severamente reprimida; discursos qualificados como heréticos, blasfemos ou sediciosos eram combatidos com rigor pelo direito canônico e pelos tribunais inquisitoriais, que não hesitavam em recorrer a punições extremas como a tortura, a pena capital ou o banimento. Nesse contexto, o que hoje poderíamos identificar como "discursos de ódio" era frequentemente compreendido e tratado à luz da doutrina da verdade revelada e da necessidade de preservar a unidade da fé e a ordem social estabelecida, onde a heresia equivalia a uma ameaça existencial à coletividade e à salvação das almas.

A consolidação da modernidade, impulsionada por eventos como a Reforma Protestante, o Iluminismo e as revoluções burguesas, especialmente a partir dos séculos XVII e XVIII, testemunhou o surgimento e a ascensão da esfera pública burguesa, categoria central na obra seminal de Jürgen Habermas. A expansão do letramento, o desenvolvimento da imprensa, a proliferação de cafés, salões literários e outras associações voluntárias criaram as condições para um novo tipo de debate crítico entre particulares – majoritariamente homens burgueses e letrados – que se reuniam para discutir assuntos de interesse comum, formando uma opinião pública capaz de interpelar e fiscalizar o poder estatal. Habermas (2003) descreve esse novo espaço como caracterizado por um discurso racional-crítico entre particulares que, fazendo uso público da razão, buscavam o bem comum por meio da argumentação e do consenso (HABERMAS, 2003, p. 62).

No entanto, como o próprio Habermas reconheceu posteriormente e como suas críticas feministas e pós-coloniais apontaram (FRASER, 1990; CALHOUN, 1992), esse modelo idealizado da esfera pública burguesa era também profundamente elitista e excludente, marginalizando mulheres, trabalhadores, minorias étnicas e outros grupos subalternos, cujas vozes e preocupações raramente encontravam espaço nesse fórum pretensamente universal. Além disso, mesmo nesse período de florescimento da razão iluminista, os discursos de ódio, agora secularizados, continuaram a circular e a encontrar ressonância, manifestando-se em panfletos racistas, caricaturas antisemitas, teorias pseudocientíficas e discursos eugenistas, que muitas vezes encontraram amparo ou conivência em regimes jurídicos discriminatórios, como na Alemanha nazista, nas leis de segregação racial dos Estados Unidos ou no regime do *apartheid* na África do Sul.

A transição para o século XXI, marcada pela revolução digital e pela popularização massiva da internet, introduziu uma mudança paradigmática de magnitude comparável à invenção da imprensa: a digitalização da esfera pública. As redes sociais digitais – plataformas como Facebook, X (antigo Twitter), Instagram, YouTube, TikTok, entre outras – emergiram rapidamente como o principal palco da comunicação política, cultural, social e interpessoal em escala global. Essas plataformas reconfiguraram radicalmente a dinâmica da comunicação, permitindo uma aparente descentralização da produção e circulação do discurso, onde qualquer usuário pode, em tese, criar e disseminar conteúdo para uma audiência potencialmente massiva. Contudo, essa transformação trouxe

consigo consequências ambivalentes e problemáticas.

A dissolução dos tradicionais filtros editoriais, a erosão da autoridade epistêmica das instituições tradicionais (como a imprensa profissional e a academia), a aceleração vertiginosa do fluxo informacional e a própria arquitetura algorítmica dessas plataformas criaram um terreno fértil para a explosão da desinformação (fake news), da pós-verdade, da polarização afetiva e da disseminação viral de discursos extremados, incluindo o discurso de ódio. Nesse cenário, como observou ironicamente Umberto Eco (2015):

"as redes sociais deram o direito à palavra a legiões de imbecis que, anteriormente, falavam apenas em um bar depois de uma taça de vinho, sem prejudicar a coletividade. [...] O drama da internet é que ela promoveu o idiota da aldeia a portador da verdade" (ECO, 2015).

O "idiota da aldeia", outrora socialmente contido em seu alcance, agora opera como produtor e difusor de "verdades" alternativas, frequentemente amplificado por algoritmos opacos que priorizam o engajamento (cliques, curtidas, compartilhamentos, comentários) acima da veracidade, da relevância cívica ou da qualidade do debate.

A configuração atual das redes sociais digitais, portanto, altera profundamente os pressupostos clássicos sobre os quais se assentavam as teorias da liberdade de expressão e da esfera pública democrática. Como argumenta Shoshana Zuboff em sua análise crítica do "capitalismo de vigilância", os discursos que circulam nessas plataformas não fluem livremente em um mercado de ideias neutro, mas são constantemente monitorados, coletados, analisados e, crucialmente, moldados por mecanismos algorítmicos sofisticados que visam à predição e modificação

do comportamento humano, com finalidades primariamente comerciais (publicidade direcionada) e, cada vez mais, políticas (microdirecionamento de campanhas, manipulação da opinião pública) (ZUBOFF, 2020, p. 12).

Nesse ecossistema digital, a viralização do discurso de ódio, da desinformação e da polarização não é um mero efeito colateral indesejado, mas muitas vezes um resultado funcional da própria arquitetura das plataformas, que são projetadas para maximizar o tempo de permanência e o engajamento dos usuários, recompensando conteúdos que geram reações emocionais intensas, sejam elas positivas ou negativas. A formação de "bolhas de filtros" e "câmaras de eco" algorítmicas contribui para fragmentar a esfera pública, isolando os usuários em realidades informacionais paralelas e dificultando o diálogo e a compreensão mútua entre grupos com visões de mundo distintas (SUNSTEIN, 2017; PARISER, 2012).

Do ponto de vista jurídico e constitucional, essa nova realidade digital levanta questões cruciais e complexas: devem as redes sociais ser consideradas espaços públicos ou privados? Estão elas sujeitas aos mesmos deveres constitucionais que o Estado, especialmente no que tange à proteção de direitos fundamentais como a liberdade de expressão, a igualdade e a dignidade humana? Qual o grau de responsabilidade que deve ser atribuído a essas plataformas pelos conteúdos ilícitos, como o discurso de ódio, que nelas circulam e são amplificados? A doutrina jurídica e a jurisprudência, tanto no Brasil quanto internacionalmente, têm cada vez mais reconhecido que, apesar de sua titularidade formalmente

privada, as grandes plataformas digitais que dominam o ecossistema comunicacional contemporâneo exercem uma função pública essencial e detêm um poder estrutural significativo sobre o fluxo de informações e o debate público (VAN DIJCK; POELL; DE WAAL, 2018).

Esse reconhecimento justifica a aplicação, ainda que de forma adaptada, de certas obrigações constitucionais a esses atores privados (a chamada eficácia horizontal dos direitos fundamentais), especialmente deveres positivos de proteção contra violações de direitos perpetradas por terceiros em seus ambientes. Como destaca Luis Greco (2019), ao analisar a função pública da comunicação digital.

Os provedores de aplicação que organizam e estruturam o debate público digital, como as grandes redes sociais, exercem um poder social e comunicacional de tal magnitude que justifica sua submissão a deveres públicos de justificação, transparência e moderação de conteúdo, compatíveis com os princípios constitucionais (GRECO, 2019, p. 151).

Dessa forma, o enfrentamento jurídico e social dos discursos de ódio e da desinformação nas redes sociais demanda o reconhecimento inequívoco de que esses ambientes digitais constituem as arenas públicas centrais da contemporaneidade, cujo funcionamento e arquitetura afetam diretamente a qualidade da democracia, a integridade dos direitos humanos, a coesão social e a própria legitimidade do debate público. A atuação do Estado e da sociedade civil não pode se limitar à criminalização puramente reativa de manifestações individuais, mas deve avançar para uma abordagem mais estrutural e preventiva, que inclua a regulação democrática da arquitetura digital (com foco em transparência algorítmica, devido processo na moderação de conteúdo e interoperabilidade), a

definição clara de regimes de responsabilidade para as plataformas (que equilibrem liberdade de expressão e proteção contra danos), o fomento à educação midiática crítica e o fortalecimento de mecanismos de resolução de conflitos online que sejam ágeis, acessíveis e eficazes. Somente através de uma combinação de estratégias regulatórias, tecnológicas, educacionais e cívicas será possível mitigar os riscos e potencializar as promessas democráticas da esfera pública digital.

2.4.2 Algoritmos, Viralização e Responsabilidade das Plataformas: A Arquitetura da Amplificação do Ódio

No complexo ecossistema digital que caracteriza a esfera pública contemporânea, os discursos de ódio não apenas encontram um terreno fértil para circular – eles são, de maneira sistemática e muitas vezes intencional, amplificados por sofisticados mecanismos algorítmicos. Estes algoritmos, desenvolvidos e operados pelas grandes plataformas digitais (como Meta/Facebook, Google/YouTube, ByteDance/TikTok, X/Twitter), são projetados com o objetivo primordial de maximizar o tempo de permanência do usuário, o engajamento (interações como cliques, curtidas, compartilhamentos, comentários) e, conseqüentemente, a receita publicitária que sustenta seus modelos de negócio. Essa realidade intrinsecamente tecnológica e econômica rompe com os paradigmas clássicos que informavam a compreensão da liberdade de expressão e impõe um desafio sem precedentes ao direito constitucional e à regulação democrática: a necessidade de compreender, avaliar criticamente e, eventualmente, intervir em um modelo de comunicação que parece estruturalmente orientado à viralização de conteúdos extremados,

polarizadores e, frequentemente, odiosos.

Os algoritmos de recomendação e curadoria de conteúdo, que determinam o que cada usuário vê em seu *feed* de notícias ou na página inicial, são o coração pulsante dessas plataformas. Eles operam com base em lógicas complexas de aprendizado de máquina (*machine learning*) e predição comportamental, que mapeiam incessantemente os padrões de navegação, as preferências implícitas e explícitas, o histórico de interações, os dados demográficos e psicográficos de cada indivíduo para construir perfis detalhados e entregar conteúdos personalizados que maximizem a probabilidade de engajamento e retenção. Como argumenta Shoshana Zuboff (2020) em sua análise do "capitalismo de vigilância", as plataformas digitais deixaram de ser meros intermediários passivos ou espaços neutros de livre circulação de ideias para se tornarem verdadeiras "arquiteturas de modificação comportamental", cujo objetivo central é a extração de "excedente comportamental" (dados sobre a experiência humana) para transformá-lo em previsões de comportamento e, em última instância, em produtos comercializáveis no mercado de futuros comportamentais. Nesse modelo, "a atenção humana não é apenas o recurso a ser capturado, mas o meio pelo qual o comportamento é moldado e direcionado para fins comerciais e, cada vez mais, políticos" (ZUBOFF, 2020, p. 35).

Neste contexto operacional e econômico, os discursos de ódio – por sua natureza frequentemente sensacionalista, polarizadora, emocionalmente carregada e transgressora – tendem a apresentar um alto potencial de viralização, sendo, em muitos casos, inadvertida ou

deliberadamente favorecidos pelos próprios algoritmos de recomendação. Diversos estudos empíricos e investigações jornalísticas têm corroborado a hipótese de que conteúdos extremistas, violentos, conspiratórios, desinformativos ou discriminatórios tendem a gerar níveis mais elevados de engajamento (reações, comentários, compartilhamentos) do que conteúdos considerados mais neutros, factuais ou ponderados (WARDLE; DERAKHSHAN, 2017; RIBEIRO et al., 2019).

Nesse sentido, vemos decisões tomadas pelas plataformas para permitir o livre trânsito dessas ideias, no intuito claro de aumentar os lucros, é o que transparece da recente decisão da Meta, onde permitirá “acusações de anormalidade mental relacionadas a gênero ou orientação sexual”. Tal medida, alertam especialistas, infere que o discurso de que chamar homossexuais ou transexuais de doentes mentais está liberado nas redes da gigante Meta (que engloba Facebook, Instagram e Whatsapp) (NAHAS, 2025).

Os algoritmos, otimizados para maximizar métricas de engajamento, acabam por identificar esses padrões e, conseqüentemente, promover e distribuir mais amplamente tais conteúdos, criando ciclos viciosos de radicalização e polarização. Essa constatação expõe uma tensão fundamental: o interesse econômico das plataformas em manter os usuários engajados pode entrar em conflito direto com a necessidade de proteger a esfera pública de discursos nocivos e preservar um ambiente comunicacional saudável. A viralização do ódio, nesse sentido, pode não ser apenas um "bug" no sistema, mas uma "feature" lucrativa, uma externalidade negativa intrinsecamente ligada ao modelo de negócio

dominante baseado na economia da atenção (WU, 2017).

Esse modelo de negócio, impulsionado por algoritmos opacos e otimização constante para o engajamento, transforma as plataformas digitais em agentes estruturais e poderosos do discurso público contemporâneo, distanciando-as significativamente da imagem de meras intermediárias passivas ou "quadros de avisos" neutros.

As plataformas atuam como editores invisíveis do debate público contemporâneo, escolhendo o que aparece, para quem aparece, em que ordem e com que destaque, por meio de critérios algorítmicos que escapam ao escrutínio público (MAFEI, 2016, p. 152).

Essa função editorial algorítmica, embora distinta da curadoria humana tradicional da imprensa, confere às plataformas um poder informacional e uma capacidade de influenciar a opinião pública comparáveis, ou até superiores, aos dos grandes grupos de mídia tradicionais. Essa constatação alimenta o intenso debate jurídico e político sobre a necessidade de sua equiparação jurídica em termos de responsabilidade civil, regulatória e, em casos extremos, até penal.

Os defensores de uma maior responsabilização das plataformas argumentam que, uma vez que estas exercem um controle técnico significativo sobre a arquitetura do ambiente comunicacional e lucram diretamente com o conteúdo que circula (inclusive o ilícito), elas devem arcar com uma parcela da responsabilidade pelos danos causados por discursos de ódio, desinformação ou outras formas de conteúdo prejudicial que hospedam, promovem ou monetizam. A analogia com os meios de comunicação tradicionais é frequentemente invocada, ressaltando que ambos intermedeiam e moldam o debate público, com a diferença vital de

que as plataformas o fazem em uma escala sem precedentes, de forma automatizada, personalizada e frequentemente opaca. Para essa corrente, a manutenção de regimes de isenção ampla de responsabilidade, de forma matizada, no artigo 19 do Marco Civil da Internet no Brasil (que condiciona a responsabilidade civil à desobediência de ordem judicial específica), contribui para a impunidade digital, incentiva a negligência na moderação de conteúdo e, em última análise, corrói a qualidade da esfera pública democrática (VENTURINI, 2020; GILLESPIE, 2018). Argumenta-se, ainda, que a assimetria informacional e de poder entre as gigantescas plataformas transnacionais e os usuários individuais, a opacidade dos algoritmos e a dificuldade prática de litigar contra essas empresas justificam a imposição de deveres jurídicos positivos mais robustos, como deveres de cuidado (*duty of care*), obrigações de transparência algorítmica, implementação de mecanismos eficazes de moderação e recurso, e cooperação com autoridades.

Por outro lado, os críticos de uma responsabilização mais ampla das plataformas, incluindo as próprias empresas de tecnologia e defensores de uma visão mais libertária da liberdade de expressão online, alertam para os riscos de uma intervenção estatal ou judicial excessiva, que poderia levar à censura privada em larga escala e ao chamado *chilling effect* (efeito inibidor) sobre a livre manifestação. Argumentam que exigir das plataformas o mesmo nível de controle editorial prévio que se espera de uma redação jornalística seria tecnicamente inviável, dada a imensa quantidade de conteúdo gerado por usuários a cada segundo, e desproporcional, podendo resultar na remoção excessiva de conteúdos

legítimos ou controversos por receio de sanções (*overblocking*). Além disso, destacam que grande parte do conteúdo problemático é gerado por terceiros, e responsabilizar as plataformas por atos alheios poderia desincentivar a própria oferta de serviços de intermediação. Defendem, em geral, a manutenção de regimes de responsabilidade limitada (*safe harbors*), combinados com mecanismos de autorregulação, cooperação voluntária e notificação e retirada (*notice and takedown*) após ordem judicial ou notificação clara de ilegalidade manifesta (LEMOS; MARQUES, 2021).

Contudo, a objeção baseada na neutralidade técnica das plataformas e na impossibilidade de controle editorial parece ignorar, como já mencionado, que a própria arquitetura algorítmica dessas plataformas já realiza, por definição, uma seleção, priorização e amplificação de conteúdos com base em critérios predominantemente econômicos (maximização do engajamento e da receita publicitária). Essa curadoria algorítmica, mesmo que não intencional em relação a um conteúdo específico, produz efeitos normativos e sociais concretos, moldando o que é visível e o que permanece oculto na esfera pública digital.

A invocada neutralidade técnica das plataformas é, em grande medida, um mito jurídico e retórico, pois sua arquitetura é intrinsecamente carregada de valores e produz efeitos diretos sobre o exercício e a proteção dos direitos fundamentais, incluindo a liberdade de expressão e o direito à não discriminação (GRECO, 2019, p. 158).

Portanto, a questão central não parece ser *se* as plataformas devem ou não ser responsabilizadas por sua influência no ecossistema informacional, mas sim *como, em que medida, sob quais condições e com*

quais garantias processuais e substantivas essa responsabilização deve ocorrer, de modo a equilibrar a proteção contra discursos nocivos com a salvaguarda da liberdade de expressão e do pluralismo.

Os algoritmos que governam as redes sociais não são meros instrumentos técnicos neutros, e a viralização do discurso de ódio não é um fenômeno acidental ou inevitável. Trata-se, em grande medida, de uma consequência previsível e, por vezes, funcionalmente lucrativa de um modelo de negócio e de uma arquitetura comunicacional orientados à maximização do engajamento a qualquer custo. A responsabilização das plataformas digitais por seu papel na amplificação desses discursos nocivos, longe de representar uma ameaça intrínseca à liberdade de expressão ou uma forma de censura estatal disfarçada, pode ser compreendida como uma exigência de coerência com os princípios constitucionais da dignidade humana, da igualdade, da proteção de grupos vulneráveis e da própria higidez do debate democrático. O desafio reside em construir modelos regulatórios que sejam eficazes em mitigar os danos, transparentes em seus mecanismos, proporcionais em suas exigências e que garantam o devido processo e a proteção das liberdades fundamentais, reconhecendo a complexidade de um ambiente digital que ainda opera com um significativo déficit democrático e um excesso de opacidade corporativa.

2.4.3 Marco Civil da Internet, LGPD e a Complexa Equação entre Liberdade de Expressão, Proteção de Dados e Responsabilidade na Era Digital

A consolidação da internet como o espaço central e definidor da

comunicação pública, da interação social e da formação da opinião política na sociedade contemporânea impôs ao ordenamento jurídico brasileiro a necessidade premente de criar um marco legal próprio. Este marco deveria ser capaz de disciplinar as complexas relações e responsabilidades de usuários, provedores de conexão e de aplicações, e do próprio Estado diante dos novos e multifacetados desafios impostos pela era digital. No Brasil, esse esforço normativo culminou na promulgação da Lei nº 12.965/2014, amplamente conhecida como Marco Civil da Internet (MCI). Concebido como uma espécie de "Constituição da Internet" brasileira, o MCI buscou estabelecer princípios, garantias, direitos e deveres fundamentais para o uso da rede no país, em um processo legislativo inovador que contou com ampla participação da sociedade civil. De forma complementar e igualmente estruturante, a Lei Geral de Proteção de Dados Pessoais (LGPD – Lei nº 13.709/2018), inspirada no Regulamento Geral sobre a Proteção de Dados (GDPR) europeu, veio regular de maneira abrangente o tratamento de dados pessoais, inclusive no ambiente digital, introduzindo um novo paradigma de direitos e obrigações na sociedade da informação e impactando diretamente a forma como as plataformas digitais operam e se relacionam com seus usuários.

O Marco Civil da Internet representou, sem dúvida, uma inovação normativa significativa ao incorporar explicitamente ao ordenamento jurídico brasileiro princípios cardeais para a governança da rede, como a garantia da liberdade de expressão, a proteção da privacidade e dos dados pessoais, a neutralidade da rede (recentemente flexibilizada em alguns aspectos pela jurisprudência) e um regime específico de responsabilidade

civil para os agentes intermediários. No que tange especificamente à liberdade de expressão, o MCI a consagra como um dos fundamentos do uso da internet no Brasil (art. 3º, I), assegurando "a liberdade de expressão, comunicação e manifestação de pensamento, nos termos da Constituição Federal" (art. 8º). No entanto, essa garantia não é absoluta e deve ser interpretada em harmonia com outros princípios e direitos igualmente fundamentais estabelecidos pela própria lei e pela Constituição, como a proteção da privacidade (art. 3º, II), a proteção dos dados pessoais (art. 3º, III), a preservação da dignidade da pessoa humana, a defesa dos direitos humanos e o desenvolvimento da personalidade (art. 2º, parágrafo único, c/c art. 7º).

Um dos pontos mais debatidos e controversos do MCI reside no regime de responsabilidade civil dos provedores de aplicações de internet (categoria que abrange as redes sociais, plataformas de vídeo, motores de busca, etc.) por danos decorrentes de conteúdo gerado por terceiros. O artigo 19 do MCI adotou a regra conhecida como "notice and judicial takedown" (notificação e remoção judicial), segundo a qual o provedor de aplicações de internet somente poderá ser responsabilizado civilmente por danos decorrentes de conteúdo gerado por terceiros se, após ordem judicial específica que determine a indisponibilização do conteúdo apontado como infringente, não tomar as providências para, no âmbito e nos limites técnicos do seu serviço e dentro do prazo assinalado, tornar indisponível o conteúdo apontado como infringente. Essa regra, que excepciona casos específicos como violação de direitos autorais (art. 19, §1º) e divulgação não consentida de imagens íntimas (art. 21), foi inicialmente celebrada por

muitos setores como um mecanismo essencial para proteger a liberdade de expressão contra a censura privada e para conferir segurança jurídica às plataformas, evitando que elas fossem obrigadas a monitorar proativamente todo o conteúdo postado por milhões de usuários. Contudo, com o passar do tempo e a crescente percepção dos danos causados pela disseminação viral de desinformação, discurso de ódio e outros conteúdos ilícitos, o artigo 19 passou a ser alvo de intensas críticas por sua alegada ineficiência prática e por criar uma espécie de "porto seguro" excessivamente protetivo para as plataformas.

Os críticos do artigo 19 argumentam que a exigência de uma ordem judicial específica para cada conteúdo ilícito a ser removido sob pena de responsabilização civil é um procedimento excessivamente lento, burocrático e inadequado para a dinâmica veloz e viral da internet. Como observa Danilo Doneda 2020, um dos especialistas que participou da elaboração do MCI, mas que posteriormente reviu sua posição sobre o artigo 19.

A exigência de ordem judicial específica para responsabilização, embora bem-intencionada na origem para evitar a censura privada, acabou por criar uma blindagem prática excessiva às plataformas, mesmo quando estas lucram com a circulação de conteúdos manifestamente ilícitos ou quando falham em aplicar suas próprias políticas de moderação (DONEDA, 2020, p. 122).

De fato, em situações de rápida viralização de discursos de ódio, campanhas de desinformação coordenadas ou ameaças graves – que podem atingir milhões de pessoas em questão de horas ou dias –, a necessidade de aguardar uma decisão judicial individualizada torna a resposta estatal frequentemente tardia e ineficaz, permitindo que os danos

se consumam e se tornem, muitas vezes, irreversíveis. Essa crítica se intensifica em contextos de alta sensibilidade, como períodos eleitorais, crises sanitárias ou situações de violência iminente contra grupos vulneráveis.

A constitucionalidade e a interpretação do artigo 19 do MCI estiveram sob intenso escrutínio do Supremo Tribunal Federal (STF), no âmbito de dois Recursos Extraordinários com repercussão geral reconhecida (RE 1.037.396 - Tema 533 e RE 1.057.258 - Tema 987). O julgamento, opunha, de um lado, aqueles que defendem a manutenção da regra da ordem judicial como garantia essencial à liberdade de expressão e à inovação, e, de outro, aqueles que propõem uma reinterpretação ou mesmo a declaração de inconstitucionalidade parcial do dispositivo, para permitir a responsabilização das plataformas em casos de omissão diante de conteúdos manifestamente ilícitos ou quando houver descumprimento de seus próprios termos de serviço. A decisão final do STF foi tomada recentemente, em 26 de junho de 2025, pelo placar de 8x3 os ministros decidiram que o art. 19 do Marco Civil da Internet (Lei 12.965/2014) é parcialmente inconstitucional:

De acordo com a decisão, enquanto o Congresso Nacional não editar nova lei sobre o tema, a plataforma será responsabilizada civilmente pelos danos decorrentes de conteúdos gerados por terceiros em casos de crimes em geral ou atos ilícitos se, após receber um pedido de retirada, deixar de remover o conteúdo. A regra também vale para os casos de contas denunciadas como falsas (BRASIL, 2025).

Paralelamente ao debate sobre o artigo 19, a Lei Geral de Proteção de Dados Pessoais (LGPD) introduziu um novo e importante eixo de análise e potencial responsabilização das plataformas digitais. Ao regular

o tratamento de dados pessoais, inclusive por meios automatizados, a LGPD estabelece um conjunto de princípios e direitos que incidem diretamente sobre as práticas algorítmicas das plataformas, incluindo aquelas relacionadas à moderação, recomendação e direcionamento de conteúdo.

A LGPD exige que todo tratamento de dados pessoais observe princípios como finalidade (propósitos legítimos, específicos, explícitos e informados), adequação (compatibilidade com as finalidades), necessidade (limitação ao mínimo necessário), livre acesso, qualidade dos dados, transparência, segurança, prevenção, não discriminação e responsabilização e prestação de contas (*accountability*) (art. 6º). O princípio da não discriminação, em particular, veda o tratamento de dados para fins discriminatórios ilícitos ou abusivos. Além disso, o artigo 20 da LGPD confere ao titular dos dados o direito de solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluindo decisões destinadas a definir seu perfil pessoal, profissional, de consumo e de crédito, ou os aspectos de sua personalidade. Embora a redação final do artigo 20 tenha sido objeto de vetos e controvérsias quanto à obrigatoriedade da revisão humana, ele representa um reconhecimento legal da necessidade de escrutínio e contestação das decisões algorítmicas.

A relação entre a LGPD, os algoritmos e o discurso de ódio se evidencia quando os sistemas algorítmicos das plataformas, ao segmentar audiências, personalizar feeds ou moderar conteúdos, reforçam vieses discriminatórios existentes ou direcionam conteúdos odiosos, extremistas

ou desinformativos a grupos específicos, ampliando seus impactos negativos. Esse fenômeno, por vezes denominado "violência algorítmica" ou "discriminação algorítmica", ocorre quando os algoritmos, treinados com dados históricos enviesados ou otimizados para métricas de engajamento que favorecem conteúdos polarizadores, acabam por perpetuar ou amplificar desigualdades e estereótipos raciais, de gênero, religiosos, políticos, entre outros (NOBLE, 2018; EUBANKS, 2018).

A governança algorítmica, incluindo as decisões sobre o que é mostrado ou ocultado nas plataformas, não pode operar em um vácuo jurídico ou ético. Ela deve estar sujeita aos princípios e garantias constitucionais e legais, especialmente quando afeta direitos fundamentais sensíveis como a honra, a imagem, a privacidade, a igualdade e a dignidade humana. A LGPD oferece ferramentas importantes para exigir transparência, equidade e accountability nesses processos (BIONI, 2021, p. 78).

A aplicação efetiva da Lei Geral de Proteção de Dados (LGPD) aos algoritmos das plataformas, contudo, ainda enfrenta desafios significativos, relacionados à complexidade técnica, à opacidade dos sistemas e à capacidade de fiscalização da Autoridade Nacional de Proteção de Dados (ANPD), que é uma autarquia federal vinculada ao Ministério da Justiça.

Dessa forma, o quadro normativo brasileiro atual, composto principalmente pelo Marco Civil da Internet e pela LGPD, já oferece alguns instrumentos jurídicos relevantes para abordar a questão da responsabilidade das plataformas digitais e os desafios impostos pelo discurso de ódio e pela desinformação online. No entanto, a aplicação desses instrumentos ainda é objeto de intensos debates doutrinários e jurisprudenciais, e há um reconhecimento crescente da necessidade de

atualização legislativa e regulatória para lidar de forma mais eficaz e equilibrada com a complexidade do ambiente digital contemporâneo. A discussão em torno do PL 2630/2020 e a decisão recente do STF sobre o artigo 19 do MCI são reflexos dessa busca por um novo ponto de equilíbrio, que consiga conciliar a proteção da liberdade de expressão com a responsabilização proporcional das plataformas por sua atuação na moderação e amplificação de conteúdos, à luz dos princípios constitucionais e dos direitos humanos. A construção desse novo marco regulatório exigirá um diálogo informado, plural e atento às experiências internacionais, mas também sensível às particularidades do contexto brasileiro.

2.4.4 Moderação de Conteúdo e Censura: Dilemas Jurídicos na Encruzilhada da Liberdade e da Proteção

A expansão exponencial da comunicação digital, predominantemente mediada por plataformas privadas de alcance transnacional e com poder de mercado significativo, tornou incontornável e cada vez mais urgente o debate sobre os limites da moderação de conteúdo e os riscos inerentes de censura, seja ela privada ou estatal. Em uma democracia constitucional robusta, o exercício pleno da liberdade de expressão transcende o mero direito individual de emitir opiniões; ele abarca também o direito de ser ouvido, o acesso equitativo e não discriminatório à esfera pública e, crucialmente, a garantia contra a supressão arbitrária ou desproporcional de discursos, mesmo aqueles considerados controversos, minoritários ou críticos ao poder estabelecido. A moderação de conteúdo – compreendida como o conjunto de práticas de

revisão, filtragem, rotulagem, limitação de alcance (*shadow banning*), desmonetização, suspensão de contas ou exclusão definitiva de mensagens e perfis por parte das plataformas digitais ou, em certos contextos, por autoridades públicas – ocupa, por conseguinte, um lugar intrinsecamente ambíguo e tenso no debate jurídico contemporâneo. Ela se apresenta, simultaneamente, como um mecanismo potencialmente necessário para a proteção de outros direitos fundamentais e para a salvaguarda da própria integridade do debate democrático contra discursos manifestamente ilícitos e prejudiciais, e como um potencial instrumento de opressão, silenciamento e controle indevido da expressão.

Do ponto de vista do direito constitucional brasileiro e comparado, é um consenso largamente estabelecido que a liberdade de expressão, embora fundamental, não constitui um direito absoluto ou ilimitado. Como já extensivamente demonstrado ao longo desta dissertação, o ordenamento jurídico brasileiro, em consonância com os principais tratados internacionais de direitos humanos, impõe limites materiais e formais a esse direito, notadamente para assegurar a proteção da dignidade da pessoa humana (fundamento da República, art. 1º, III, da Constituição Federal de 1988 – CF/88), da honra, da imagem, da intimidade, da privacidade (art. 5º, X, CF/88), da não discriminação (art. 3º, IV, e art. 5º, XLI e XLII, CF/88), da proteção de crianças e adolescentes, da segurança pública e da própria ordem democrática. Nesse sentido, a moderação de conteúdos que se configurem como manifestamente ilícitos – a exemplo de discursos de ódio que incitem à violência, à discriminação ou ao genocídio, pornografia infantil, apologia ao terrorismo, racismo, ameaças diretas e críveis, ou

campanhas coordenadas de desinformação com o intuito de subverter processos eleitorais – não configura, em tese, censura prévia vedada pela Constituição (art. 5º, IX, e art. 220, §2º, CF/88), mas sim uma forma de restrição legítima, necessária e proporcional, desde que observados os devidos processos legais e substantivos. O desafio essencial, contudo, reside na definição precisa dos critérios para identificar tais conteúdos, nos procedimentos adotados para sua remoção ou restrição, e na determinação de quem são os agentes responsáveis e competentes para realizar essa moderação, com quais garantias e sob qual escrutínio.

Atualmente, a vasta maioria das decisões de moderação de conteúdo online – exclusão de postagens, suspensão ou banimento de contas, redução de alcance – é tomada pelas próprias plataformas digitais (Meta, Google, X, TikTok, etc.), com base em seus termos de uso (Termos de Serviço - ToS) e políticas internas de comunidade (TERRA, 2025). Esses documentos, elaborados de forma unilateral pelas empresas, frequentemente extensos, vagos e sujeitos a alterações discricionárias, funcionam como uma espécie de "legislação privada" que governa a expressão de bilhões de usuários em escala global. Essa prática configura uma forma de poder normativo e adjudicatório privado de imensa magnitude, que muitas vezes opera à margem das garantias processuais mínimas consagradas nos ordenamentos jurídicos estatais, como o direito à notificação prévia e específica, o direito à fundamentação da decisão, o direito ao contraditório e à ampla defesa, e o direito a um recurso efetivo perante um órgão imparcial.

Como adverte Kate Klonick (2018), as plataformas se tornaram os

"novos governadores" da expressão online, tomando decisões com profundas implicações para os direitos fundamentais, muitas vezes em ambientes opacos e com pouca *accountability* externa. Essa "privatização da censura", como alguns autores a denominam (POST, 2012), levanta sérias preocupações sobre a arbitrariedade, a inconsistência e o potencial viés ideológico, comercial ou político dessas decisões de moderação, que podem resultar no silenciamento indevido de vozes dissidentes, minoritárias ou críticas.

Essa realidade fática suscita o risco concreto do chamado "efeito silenciador" ou "efeito inibidor" (*chilling effect*), pelo qual indivíduos e grupos deixam de se expressar livremente por receio de sofrerem sanções privadas (remoção de conteúdo, suspensão de conta) de forma arbitrária, imprevisível ou desproporcional. Como destaca Ronald Dworkin (2002, p. 251) em sua defesa de uma concepção robusta da liberdade de expressão, esta "exige não apenas a ausência de censura legal formal, mas também a ausência de ameaças difusas, intimidações e outras formas de pressão indevida que possam inibir o discurso público". Assim, o imenso poder de moderação concentrado nas mãos de poucas empresas de tecnologia precisa ser juridicamente enquadrado e controlado, tanto para evitar abusos e garantir a proteção dos direitos dos usuários, quanto para assegurar a efetividade do pluralismo discursivo e a vitalidade do debate democrático na esfera digital.

Do ponto de vista normativo, o ordenamento jurídico brasileiro ainda carece de mecanismos legais abrangentes e detalhados para o controle democrático da moderação de conteúdo realizada por plataformas

privadas, embora existam importantes iniciativas legislativas em andamento, com destaque para o Projeto de Lei nº 2.630/2020 (BRASIL, 2020), a chamada "Lei das Fake News" ou Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Este projeto, após inúmeras versões e intenso debate público, busca estabelecer um conjunto de regras para a atuação das plataformas, incluindo obrigações de transparência sobre políticas de moderação e sistemas algorítmicos, dever de notificação específica ao usuário em caso de remoção de conteúdo ou suspensão de conta, direito a recurso interno e externo, e a criação de uma entidade autônoma de supervisão. Tais medidas se aproximam da doutrina da "justiça procedimental algorítmica" e do conceito de "devido processo digital" (*digital due process*), que exigem que as decisões automatizadas ou semi-automatizadas de moderação de conteúdo obedeçam a princípios constitucionais fundamentais como a motivação clara e precisa, a razoabilidade, a proporcionalidade, a não discriminação e o direito à ampla defesa e ao recurso efetivo (CITRON, 2008; PASQUALE, 2015).

A experiência europeia com o Digital Services Act (DSA) (UNIÃO EUROPEIA, 2022) também oferece um paradigma relevante, ao impor obrigações de diligência e transparência mais robustas às plataformas, especialmente às de grande dimensão, no combate a conteúdos ilegais e na mitigação de riscos sistêmicos, com mecanismos de supervisão e sanção.

Por outro lado, é necessário reconhecer e mitigar os riscos associados à moderação estatal excessiva ou abusiva, especialmente em contextos políticos de crescente polarização, autoritarismo ou populismo. Tentativas de criminalizar de forma ampla, vaga e imprecisa

manifestações de pensamento tidas como "ofensivas", "antiéticas", "antipatrióticas" ou "desestabilizadoras" podem facilmente transformar o aparato estatal em um censor ideológico, comprometendo a integridade da esfera pública e silenciando vozes críticas. Como lembra enfaticamente Luís Roberto Barroso (2017, p. 122), "não se pode confundir discurso impopular, chocante ou perturbador com discurso inconstitucional ou ilegal; a democracia é, por excelência, o regime das vozes dissonantes, do contraditório e da tolerância com o diferente". A atuação estatal na regulação da expressão online, portanto, deve ser estritamente necessária para a proteção de um interesse público legítimo, idônea para alcançar tal finalidade, proporcional em sentido estrito (ou seja, os benefícios da restrição devem superar os custos para a liberdade de expressão) e guiada por parâmetros legais claros, objetivos e previsíveis, sob pena de violar a cláusula de vedação ao retrocesso democrático e os princípios da legalidade e da segurança jurídica.

Nesse cenário complexo e multifacetado, o dilema jurídico central torna-se dolorosamente evidente: como garantir a efetiva remoção de conteúdos manifestamente odiosos, violentos ou que representem grave ameaça a direitos fundamentais, sem, com isso, inviabilizar o debate público livre, plural e robusto, ou criar um ambiente de vigilância e censura generalizada? Como responsabilizar adequadamente as plataformas digitais por sua omissão culposa ou dolosa, ou mesmo pelo lucro indevido obtido com a circulação de conteúdos ilícitos, sem gerar uma onda de remoções preventivas excessivas (*overblocking*) e uma censura privada indireta que prejudique a diversidade de vozes? Essas

questões cruciais demandam uma abordagem jurídica e regulatória equilibrada, sofisticada e adaptativa, que combine elementos de regulação estatal democrática (com foco em princípios, devido processo e supervisão independente), autorregulação regulada e transparente por parte das plataformas (com códigos de conduta claros, mecanismos de recurso e auditorias externas), e uma vigorosa participação da sociedade civil, da academia e de organismos de defesa de direitos na definição, implementação e fiscalização dessas políticas.

A moderação de conteúdo na era digital situa-se em uma encruzilhada delicada entre a necessidade imperativa de proteger a democracia, a dignidade humana e os grupos vulneráveis contra discursos destrutivos e ilícitos, e o dever fundamental de resguardar a liberdade de expressão como pilar essencial do pluralismo, da inovação e do controle social do poder. A ausência de uma regulação democrática e equilibrada tende a favorecer tanto a omissão ou a atuação arbitrária das plataformas quanto os abusos por parte de atores estatais com agendas autoritárias. O enfrentamento jurídico adequado desse desafio exige, portanto, a construção de um ecossistema regulatório que priorize a procedimentalização das decisões de moderação, a transparência dos critérios e algoritmos envolvidos, a proporcionalidade das medidas restritivas, o controle social e judicial efetivo, e o fortalecimento da educação midiática e digital. Somente assim a moderação de conteúdo poderá se converter, não em um instrumento de censura ou controle, mas em um mecanismo legítimo e necessário para a garantia dos direitos fundamentais e para a promoção de um ambiente online mais seguro, ético

e propício ao florescimento da democracia na era digital.

2.4.5 Casos Emblemáticos de Discurso de Ódio no Brasil: A Construção Jurisprudencial em um Campo Minado

A interpretação jurídica do discurso de ódio, especialmente em sua manifestação virulenta no ambiente digital, tem sido progressivamente delineada através da análise de casos paradigmáticos, tanto nos tribunais constitucionais nacionais quanto nos sistemas internacionais de proteção dos direitos humanos. Essas decisões judiciais, muitas vezes proferidas em contextos de alta tensão social e política, revelam a intrínseca complexidade de equilibrar a liberdade de expressão – pilar de qualquer sociedade democrática – com a imperativa proteção da dignidade da pessoa humana, da igualdade e da própria integridade do processo democrático, todos severamente ameaçados pela velocidade, alcance e potencial de dano exponenciados pelas redes sociais e outras plataformas online. No Brasil, o Supremo Tribunal Federal (STF) tem sido chamado a se pronunciar sobre o tema em diversas oportunidades, construindo uma jurisprudência ainda em consolidação, mas que reflete os desafios impostos por essa nova realidade comunicacional, mesmo diante de uma ausência de legislação específica e abrangente que sistematize o combate ao discurso de ódio online.

Um dos julgados mais emblemáticos e citados do STF nesse contexto é a Ação Direta de Inconstitucionalidade por Omissão (ADO) 26/DF, apreciada em 2019. Nesta ação, discutia-se a mora legislativa em criminalizar especificamente a homofobia e a transfobia. O Tribunal, por expressiva maioria, compreendeu que tais condutas deveriam ser

equiparadas ao crime de racismo, previsto no art. 5º, XLII, da Constituição Federal, até que o Congresso Nacional legisle sobre a matéria. O voto condutor do Ministro Celso de Mello, decano da Corte à época, ressaltou de forma incisiva que “a liberdade de expressão não constitui manto protetor para práticas de ódio que atentam contra os direitos e a dignidade de grupos historicamente vulnerabilizados” (BRASIL, 2019a). Esta decisão representou um marco na proteção dos direitos da população LGBTQIA+ no Brasil, reconhecendo o impacto devastador do discurso de ódio em suas vidas e a necessidade de uma resposta penal efetiva por parte do Estado.

Um episódio de enorme repercussão nacional e internacional, que tensionou ao máximo a relação entre o Poder Judiciário brasileiro e as grandes empresas de tecnologia, ocorreu em agosto de 2024, quando o Ministro Alexandre de Moraes, do STF, no âmbito de inquéritos que investigam a atuação de milícias digitais e a disseminação de discursos antidemocráticos, determinou a suspensão temporária da plataforma X (anteriormente conhecida como Twitter) em todo o território nacional.

A drástica medida foi motivada, segundo a decisão, pelo descumprimento reiterado e deliberado de ordens judiciais anteriores que determinavam a remoção de conteúdos e perfis envolvidos em atividades ilícitas, além da recusa da empresa em nomear um representante legal no Brasil e em fornecer dados essenciais para as investigações. A suspensão, que perdurou por 39 dias, foi condicionada ao cumprimento integral das determinações judiciais, incluindo o pagamento de multas milionárias (BRASIL, 2024; JUSBRASIL, 2024). Essa decisão foi amplamente

debatida, sendo considerada por alguns um marco necessário na responsabilização das plataformas e na defesa da soberania jurisdicional brasileira, enquanto outros a criticaram como uma medida excessiva com potencial de cercear a liberdade de expressão. Independentemente das controvérsias, o episódio foi visto por parte da imprensa internacional como um “precedente democrático” que poderia inspirar outras nações no complexo desafio de regular as *big techs* e combater o uso das redes para a disseminação de discursos de ódio e desinformação que corroem as instituições (EL PAÍS, 2024).

A realidade brasileira tem fornecido, infelizmente, exemplos contundentes da virulência do discurso de ódio online e suas nefastas consequências. Após o resultado das eleições presidenciais de 2022, observou-se uma explosão de manifestações xenófobas e racistas direcionadas à população nordestina, majoritariamente por meio de plataformas digitais. Comentários odiosos, que associavam os nordestinos à ignorância, à miséria e à dependência de programas sociais, inundaram as redes, revelando um preconceito estrutural profundamente arraigado (G1, 2023; BRASIL DE FATO, 2022). Relatórios de organizações da sociedade civil, como a SaferNet, registraram um aumento exponencial nas denúncias de xenofobia online nesse período, evidenciando como o ambiente digital pode ser instrumentalizado para incitar o ódio e a divisão regional (SAFERNET BRASIL, 2023). Esses episódios demonstram a necessidade urgente de mecanismos mais eficazes de prevenção, moderação e responsabilização por tais condutas.

Ademais, a proliferação de discursos de ódio encontra terreno fértil

em espaços menos regulados da internet, como a *dark web* e fóruns anônimos como o 4chan e seus congêneres brasileiros, a exemplo do extinto Dogolachan. Esses ambientes são frequentemente utilizados para o compartilhamento de ideologias extremistas, misóginas, racistas, antissemitas e de incitação à violência, funcionando como verdadeiras incubadoras de radicalização (NAGLE, 2017; SILVA, 2020). A dificuldade de rastreamento e a sensação de impunidade nesses espaços contribuem para a escalada de discursos cada vez mais violentos, que por vezes transbordam para o mundo offline, com consequências trágicas. Um exemplo chocante da materialização desse ódio ocorreu em maio de 2025, quando a polícia do Rio de Janeiro frustrou uma tentativa de atentado a bomba durante um show da cantora Lady Gaga. As investigações apontaram que o plano teria sido articulado em fóruns da *dark web* e motivado por intenso discurso de ódio contra a comunidade LGBTQIA+, da qual a artista é uma proeminente defensora (TERRA, 2025; VEJA, 2025). Este caso, ainda sob investigação, acende um alerta máximo sobre os perigos da radicalização online e a urgência de estratégias de inteligência e combate a esses grupos extremistas.

Esses precedentes, demonstram um esforço contínuo dos órgãos jurisdicionais em construir parâmetros jurídicos que busquem compatibilizar a fundamental liberdade de expressão com a igualmente indispensável proteção contra os efeitos deletérios do discurso de ódio, sobretudo no volátil e complexo ambiente digital. Embora ainda persistam divergências significativas quanto aos critérios precisos para a definição de discurso de ódio, ao grau de responsabilidade das plataformas digitais

e aos modelos regulatórios mais adequados, emerge um consenso crescente de que a dignidade da pessoa humana, o princípio da não discriminação e a própria saúde do debate democrático não podem ser sistematicamente subordinados à livre e irrestrita circulação de conteúdo odioso, especialmente quando este é intencionalmente amplificado por estruturas algorítmicas e impulsionado por modelos de negócio que lucram com a polarização e o sensacionalismo.

2.5 DIREITOS HUMANOS, DEMOCRACIA E O ENFRENTAMENTO AO DISCURSO DE ÓDIO

2.5.1 A Proteção de Grupos Vulneráveis como dever Jurídico e Moral do Estado Democrático de Direito

O enfrentamento eficaz e democraticamente legítimo ao discurso de ódio, fenômeno complexo e multifacetado que corrói as bases da convivência social e ameaça a integridade dos regimes democráticos contemporâneos, exige o reconhecimento inequívoco de que o Estado possui um dever jurídico e moral inalienável de proteger ativamente os grupos vulneráveis contra todas as formas de violência simbólica e discursiva que atentem contra sua dignidade, sua segurança física e psíquica, e sua plena e igualitária participação na vida pública. Este dever fundamental não é uma mera faculdade discricionária do poder público, mas uma obrigação positiva que emana diretamente tanto das normas constitucionais fundantes do Estado Democrático de Direito, quanto dos compromissos internacionais de direitos humanos soberanamente assumidos pela nação. Tal postulado representa um dos principais pilares da concepção de democracia substantiva, que transcende a mera

formalidade dos procedimentos eleitorais para abarcar a promoção efetiva da igualdade material e do respeito à diversidade.

A Constituição da República Federativa do Brasil de 1988, promulgada sob o signo da redemocratização e da superação de um passado autoritário marcado por violações sistemáticas de direitos, estabeleceu um verdadeiro marco civilizatório ao consagrar, logo em seu artigo 1º, inciso III, a dignidade da pessoa humana como um dos fundamentos da República. Essa cláusula pétrea, de valor normativo supremo, irradia seus efeitos por todo o ordenamento jurídico e orienta a interpretação e aplicação de todas as demais normas constitucionais e infraconstitucionais. A dignidade humana, nesse contexto, não é um conceito abstrato ou meramente retórico, mas um princípio jurídico concreto que impõe ao Estado a adoção de políticas públicas, medidas legislativas, administrativas e judiciais capazes de assegurar a todos os indivíduos, sem qualquer distinção, o gozo efetivo e incondicional dos direitos fundamentais. Essa proteção deve ser especialmente reforçada em relação àqueles grupos que, em razão de processos históricos, sociais, econômicos e culturais, sofreram e continuam a sofrer marginalização, preconceito, discriminação e exclusão.

A dignidade da pessoa humana constitui o valor-fonte da ordem constitucional brasileira, servindo de critério de legitimidade para as condutas estatais e privadas e impondo um dever de tutela por parte de todos os poderes constituídos (SARLET, 2022, p. 42).

A omissão do Estado em proteger a dignidade de grupos vulneráveis pode, portanto, configurar uma grave violação constitucional. Entre os principais grupos sociais que, no contexto brasileiro e global,

demandam uma proteção estatal reforçada contra o discurso de ódio e suas manifestações correlatas, destacam-se, de forma não exaustiva:

As populações racializadas, notadamente a população negra, que é alvo persistente de racismo estrutural, institucional e interpessoal, manifestado em discursos que negam sua humanidade, inferiorizam sua cultura ou incitam à violência;

Os povos indígenas e comunidades tradicionais (quilombolas, ribeirinhos, etc.), cujos direitos territoriais, culturais e à autodeterminação são constantemente ameaçados por discursos de ódio que os estigmatizam como obstáculos ao progresso ou como culturalmente inferiores, frequentemente associados a interesses econômicos predatórios;

As pessoas LGBTQIA+ (lésbicas, gays, bissexuais, travestis, transexuais, queers, intersexos, assexuais e outras identidades de gênero e orientações sexuais), que enfrentam níveis alarmantes de violência física, psicológica e simbólica, alimentada por discursos de ódio de natureza religiosa fundamentalista, moralista ou pseudocientífica;

As mulheres, especialmente em razão da misoginia, do sexismo e da violência de gênero, que se manifestam em discursos que as objetificam, inferiorizam, culpam pelas violências sofridas ou incitam ao feminicídio, ainda profundamente entranhados em diversos espaços sociais, políticos e institucionais;

As pessoas com deficiência, que são frequentemente alvo de capacitismo e excluídas de processos de participação social e democrática por meio de discursos que as infantilizam, desumanizam ou as consideram um fardo;

As pessoas em situação de rua, migrantes, refugiados e outras comunidades em contextos de extrema vulnerabilidade socioeconômica, frequentemente desumanizadas e culpabilizadas por sua condição;

As religiões de matriz africana (como Candomblé e Umbanda) e outras minorias religiosas, historicamente atingidas por intolerância religiosa, perseguição e discursos demonizadores.

E, de forma particularmente visível no contexto brasileiro recente, os nordestinos, que se tornam alvos recorrentes de discursos discriminatórios e xenófobos, marcados por estigmas regionais, estereótipos depreciativos e ridicularização cultural, especialmente em períodos eleitorais ou em debates acirrados sobre a identidade nacional, com tais discursos sendo massiva e velozmente disseminados através das redes sociais e aplicativos de mensagem.

É essencial compreender que esses grupos não são "minorias" apenas em um sentido puramente numérico, mas sim minorias políticas e sociais, no sentido sociológico e dos direitos humanos, ou seja, são segmentos da população que, independentemente de seu quantitativo, ocupam posições historicamente desprivilegiadas e subalternizadas nas estruturas de poder social, econômico e político, e que, por isso, estão sistematicamente sujeitos a formas agravadas e interseccionais de exclusão, discriminação, estigmatização e violência. A proteção específica e proativa desses segmentos não representa um privilégio indevido, mas sim uma condição indispensável para a efetivação dos princípios constitucionais da igualdade material (art. 5º, caput e inciso I, CF/88), que exige tratar os desiguais de forma desigual na medida de suas

desigualdades para promover a isonomia real, e do pluralismo político (art. 1º, V, CF/88), que reconhece e valoriza a diversidade de identidades, culturas e modos de vida como elemento constitutivo da nação.

Para além do robusto arcabouço constitucional interno, o Brasil é signatário de um vasto conjunto de tratados e convenções internacionais de direitos humanos que impõem obrigações positivas e inequívocas ao Estado em matéria de prevenção, repressão e reparação do discurso de ódio e da discriminação. A Convenção Internacional sobre a Eliminação de Todas as Formas de Discriminação Racial (CERD), adotada pela ONU em 1965 e incorporada ao ordenamento jurídico nacional com status supralegal (ou mesmo constitucional, conforme certas interpretações) por meio do Decreto nº 65.810/1969, determina em seu artigo 4º que os Estados Partes devem condenar toda propaganda e todas as organizações que se inspirem em ideias ou teorias baseadas na superioridade de uma raça ou de um grupo de pessoas de uma certa cor ou origem étnica, ou que pretendam justificar ou promover o ódio racial e a discriminação racial sob qualquer forma, comprometendo-se a adotar medidas imediatas e positivas destinadas a eliminar toda incitação a tal discriminação, ou atos de discriminação. De forma ainda mais específica e recente, a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, adotada no âmbito da Organização dos Estados Americanos (OEA) em 2013 e ratificada pelo Brasil em 2021 (promulgada pelo Decreto nº 10.932/2022), reforça e detalha essas obrigações, ampliando o escopo da proteção para incluir explicitamente categorias como orientação sexual, identidade de gênero, idade, idioma, religião e

status socioeconômico, entre outras. Esta Convenção estabelece deveres claros para os Estados, como o de adotar legislação que defina e proíba o racismo e a discriminação, incluindo o discurso de ódio, e o de promover políticas de ações afirmativas para garantir a igualdade de oportunidades (ORGANIZAÇÃO DOS ESTADOS AMERICANOS, 2013).

Nesse sentido, a proteção de grupos vulneráveis contra o discurso de ódio transcende a esfera da mera faculdade política ou da conveniência administrativa, configurando-se como um dever jurídico positivo e vinculante do Estado, que deve atuar de forma coordenada e multifacetada por meio de leis penais e civis adequadas, políticas públicas inclusivas e eficazes, programas educacionais de promoção da cultura de direitos humanos e respeito à diversidade, e decisões judiciais que afirmem a prevalência da dignidade e da igualdade sobre manifestações odiosas. Como ressalta com acerto Piovesan (2021, p. 89), a omissão do Estado em prevenir, investigar, punir e reparar atos de discriminação e discurso de ódio pode configurar, por si só, uma violação de direitos humanos por inação (ou *failure to protect*), o que tem sido consistentemente reconhecido e sancionado por cortes internacionais de direitos humanos, como a Corte Interamericana de Direitos Humanos, em diversos casos paradigmáticos.

No plano da jurisprudência pátria, o Supremo Tribunal Federal tem, progressivamente, reconhecido e afirmado essa obrigação estatal de proteção de forma cada vez mais explícita e contundente. No já mencionado julgamento da ADO 26/DF, que equiparou a homofobia e a transfobia ao crime de racismo, a Corte não apenas supriu uma omissão

legislativa inconstitucional, mas também reafirmou que a dignidade da pessoa humana e o princípio da igualdade substancial impõem ao Estado o dever indeclinável de adotar todas as medidas protetivas necessárias contra discursos e condutas discriminatórias, sob pena de legitimar a exclusão, a violência e a perpetuação de um ciclo de opressão. A decisão do STF firmou o entendimento de que a omissão legislativa ou administrativa pode representar uma forma insidiosa de perpetuação da violação de direitos fundamentais, especialmente quando atinge grupos já estigmatizados e vulnerabilizados pela história e pelas estruturas sociais (BRASIL, 2019).

É fundamental destacar que, em contextos democráticos e pluralistas, a proteção robusta contra o discurso de ódio não deve ser erroneamente interpretada como uma exceção arbitrária ou uma ameaça à liberdade de expressão, mas sim como uma condição intrínseca para a sua própria possibilidade e florescimento.

Exige uma esfera pública inclusiva, na qual todos os cidadãos possam participar em condições de igualdade e respeito mútuo, sem medo de serem silenciados, desumanizados ou ameaçados (POST, 2011, p. 75).

O discurso que sistematicamente desumaniza, inferioriza, silencia ou ameaça a existência física e simbólica de determinados grupos não apenas fere direitos individuais e coletivos, mas também corrompe o próprio fundamento deliberativo da democracia, impedindo um debate público racional, informado e respeitoso, essencial para a tomada de decisões coletivas justas. A liberdade de expressão não pode ser um porto seguro para a opressão.

Portanto, o enfrentamento jurídico, institucional e social ao

discurso de ódio deve ser compreendido como parte integrante e indissociável da agenda de proteção e promoção dos direitos humanos no século XXI. Cabe ao Estado, em todas as suas instâncias e poderes, não apenas a tarefa reativa de punir discursos manifestamente ofensivos e ilícitos, mas, sobretudo, o dever proativo de prevenir sua proliferação por meio da educação em direitos humanos desde a primeira infância, da promoção da diversidade cultural e identitária, do fortalecimento de instituições democráticas resilientes e da construção de uma cultura cívica de respeito, empatia e solidariedade. A omissão ou a complacência do Estado nesse campo vital compromete não apenas a segurança e o bem-estar dos indivíduos e grupos diretamente afetados pelo ódio, mas a própria coesão do tecido democrático, a legitimidade das instituições e a promessa de uma sociedade justa, livre e solidária inscrita na Constituição.

2.5.2 Discurso de Ódio como Forma de Violência Simbólica, Estrutural e Instrumento Político-Eleitoral

Na esfera democrática, a palavra é, por excelência, o instrumento primordial de deliberação, construção de identidades coletivas, expressão de dissensos e exercício pleno da cidadania. No entanto, quando a linguagem é pervertida e instrumentalizada para negar a humanidade de outrem, para inferiorizar, estigmatizar ou deslegitimar sujeitos e grupos sociais historicamente marginalizados, ela deixa de ser um veículo de comunicação e converte-se em uma potente arma de dominação simbólica e de perpetuação da violência estrutural. O discurso de ódio, nesse intrincado contexto, não se resume a um mero problema moral individual ou a um desvio comunicacional isolado, mas constitui uma complexa

categoria jurídica, social e política que ativamente reproduz, reforça e legitima hierarquias históricas de poder, minando as bases da igualdade e do respeito mútuo que sustentam o pacto democrático. Ademais, em cenários de acirrada disputa política, o discurso de ódio é frequentemente mobilizado de forma estratégica como uma nefasta arma político-eleitoral, visando desumanizar adversários, inflamar bases radicalizadas e corroer a confiança nas instituições democráticas.

Inspirado nos trabalhos seminais de Pierre Bourdieu, o conceito de violência simbólica revela-se decisiva para a compreensão da dinâmica insidiosa do discurso de ódio. Conforme Bourdieu (1989), refere-se à imposição sutil, porém eficaz, de significados, valores e classificações sociais por meio da linguagem, da cultura, da educação e das instituições, de forma tão profundamente naturalizada e internalizada pelos próprios dominados que sua natureza violenta é frequentemente invisibilizada, mascarada e, por vezes, aceita como legítima ou inevitável. O discurso de ódio, ao reiterar incessantemente estigmas, ao associar determinadas identidades a atributos intrinsecamente negativos (como periculosidade, incapacidade, imoralidade) e ao sistematicamente silenciar, ridicularizar ou invalidar narrativas dissidentes e contra hegemônicas, cumpre precisamente a função de naturalizar a desigualdade, tornando-a parte integrante e incontestável do imaginário social e das práticas institucionais. Essa naturalização dificulta a percepção da injustiça e a mobilização para a sua superação.

Essa violência simbólica, manifestada através do discurso de ódio, não opera no vácuo, mas atua de forma intrinsecamente interligada e

retroalimentada com estruturas sociais profundamente desiguais e historicamente consolidadas, como o racismo estrutural, o patriarcado, a LGBTfobia, o capacitismo, a xenofobia e o colonialismo cultural e epistêmico. Quando um determinado discurso pejorativo é sistemática e massivamente repetido em meios de comunicação de grande alcance, nas redes sociais digitais, em pronunciamentos de figuras públicas ou mesmo em conversas cotidianas — como, por exemplo, a afirmação de que nordestinos são “atrasados” ou “preguiçosos”, que pessoas negras são inerentemente “perigosas” ou “menos inteligentes”, que mulheres são “emocionais demais para cargos de poder” ou “naturalmente submissas”, que indígenas “atrapalham o progresso” ou são “selvagens”, ou que pessoas LGBTQIA+ são “doentes” ou “uma ameaça à família tradicional” —, ocorre uma insidiosa reprodução e um perigoso reforço discursivo da exclusão social, econômica e política, cujos impactos transcendem em muito o plano da ofensa individual para atingir a própria estrutura de oportunidades e o reconhecimento da dignidade desses grupos.

A instrumentalização do discurso de ódio como arma político-eleitoral representa uma das facetas mais preocupantes desse fenômeno na contemporaneidade. Políticos e grupos extremistas, muitas vezes com o auxílio de sofisticadas estratégias de comunicação digital e da disseminação calculada de desinformação, utilizam narrativas odiosas para mobilizar suas bases, demonizar adversários e minorias, criar bodes expiatórios para problemas sociais complexos e, em última instância, minar a confiança no processo democrático e nas instituições. Como alertam Levitsky e Ziblatt (2018), a erosão das normas de tolerância mútua

e de autocontenção institucional, incluindo o repúdio ao discurso de ódio na arena política, é um dos principais indicadores do declínio democrático. No Brasil, as eleições recentes têm sido marcadas por uma escalada preocupante de violência política e pela disseminação de discursos de ódio contra grupos específicos, como mulheres na política, nordestinos, e minorias religiosas, muitas vezes impulsionados por campanhas de desinformação e ataques coordenados em redes sociais (ABRADEP, 2022; GOMES, 2020). Essa tática não apenas envenena o debate público, mas também pode incitar à violência real e criar um clima de medo e intimidação que afasta cidadãos da participação política.

No contexto brasileiro, a naturalização de certas formas de agressão verbal e de estereotipagem negativa é particularmente grave e deletéria, pois ocorre em uma sociedade historicamente marcada por profundas e persistentes desigualdades socioeconômicas, raciais e de gênero, e onde diversos marcadores identitários foram e continuam sendo sistematicamente violados e desrespeitados. A linguagem do ódio opera, nesse cenário, como um poderoso dispositivo de manutenção de privilégios e de hierarquias sociais injustas, promovendo a desumanização do “outro” (o diferente, o minoritário, o dissidente) e a consequente negação de sua legitimidade para ocupar o espaço público em igualdade de condições, para ter suas demandas ouvidas e seus direitos respeitados. Como afirma enfaticamente Angela Davis (2017, p. 61), “quando um grupo é constantemente representado como inferior ou perigoso, não se trata apenas de linguagem, trata-se de uma política de morte simbólica”, que precede e muitas vezes legitima a violência física e a exclusão

material.

Do ponto de vista estritamente jurídico, o discurso de ódio deve ser inequivocamente compreendido como uma flagrante violação de múltiplos direitos fundamentais, com destaque para a dignidade da pessoa humana (fundamento da República, art. 1º, III, CF/88), o princípio da igualdade e da não discriminação (art. 3º, IV, e art. 5º, caput e XLI, CF/88), e a própria liberdade de participação política e de associação (art. 17 e art. 5º, XVII, CF/88). Embora o discurso de ódio se manifeste através da palavra, seu conteúdo e seus efeitos são equivalentes a uma prática concreta de exclusão, opressão e, em muitos casos, incitação à violência. Como destaca a filósofa Martha Nussbaum (2004, p. 104), “a humilhação pública e sistemática de certos grupos destrói a base afetiva do autorrespeito, corroendo o próprio sentido de pertencimento democrático” e a capacidade desses indivíduos de se verem como cidadãos plenos e iguais (NUSBAUM, 2004).

A arquitetura tecnológica e o modelo de negócios das grandes plataformas de redes sociais acabam por potencializar exponencialmente esse fenômeno deletério. O discurso de ódio, que antes poderia ter um alcance mais limitado, deixa de ser meramente pontual ou localizado e torna-se sistêmico, viral e ubíquo, alimentado por algoritmos de recomendação que privilegiam o engajamento a qualquer custo, por mecanismos de recompensa social (curtidas, compartilhamentos) e por dinâmicas de polarização afetiva que conferem uma amplificação desproporcional à retórica do preconceito e da intolerância. Essa lógica é funcional ao modelo de negócios de muitas plataformas, que lucram com

a atenção gerada por conteúdos sensacionalistas, polêmicos e divisivos, ainda que manifestamente danosos à coesão social e à saúde democrática (ZUBOFF, 2020; NOBLE, 2018). Nesse sentido, a violência simbólica torna-se não apenas amplificada, mas também monetizada, escalável e, muitas vezes, protegida por um véu de anonimato ou pseudoanonimato, desafiando severamente os mecanismos tradicionais de responsabilização civil e penal.

O ordenamento jurídico brasileiro, embora ainda carente de uma legislação específica e abrangente sobre o discurso de ódio online, já reconhece, em diversos dispositivos constitucionais e infraconstitucionais, a imperiosa necessidade de coibir essa forma insidiosa de violência. O art. 5º, XLII, da Constituição Federal, por exemplo, determina que “a prática do racismo constitui crime inafiançável e imprescritível, sujeito à pena de reclusão, nos termos da lei”. De forma complementar, o art. 4º, inciso VIII, da mesma Carta Magna, estabelece que a República Federativa do Brasil se rege nas suas relações internacionais, entre outros, pelo princípio do “repúdio ao terrorismo e ao racismo”. A jurisprudência do Tribunal Superior Eleitoral (TSE) também tem avançado no sentido de coibir o uso de discurso de ódio e desinformação como táticas eleitorais, reconhecendo seu potencial de viciar a vontade do eleitor e de comprometer a legitimidade do pleito (TSE, 2022).

Dessa forma, o discurso de ódio deve ser enfrentado de maneira abrangente e coordenada, não apenas como um conjunto de conteúdos abusivos isolados, mas fundamentalmente como um mecanismo estrutural de reprodução de desigualdades e de corrosão dos valores democráticos,

que demanda respostas sistêmicas, envolvendo medidas educativas de longo prazo, legislação clara e proporcional, atuação firme e independente do sistema de justiça, e a promoção de uma cultura de direitos humanos e de respeito à diversidade em todas as esferas da sociedade. É preciso internalizar coletivamente que o ódio não é uma opinião legítima a ser tolerada em nome de uma concepção absolutista de liberdade de expressão, e que permitir sua disseminação indiscriminada e impune compromete a própria arquitetura dos direitos humanos e a viabilidade de uma democracia constitucional verdadeiramente inclusiva e pluralista.

2.5.3 A Jurisprudência do STF e da Corte Interamericana de Direitos Humanos sobre o Discurso de Ódio: Diálogos e Tensões na Proteção de Direitos Fundamentais

O discurso de ódio, enquanto manifestação que ultrapassa os limites da liberdade de expressão para violar a dignidade e os direitos fundamentais de grupos vulneráveis, tem sido objeto de crescente e complexa análise por parte dos tribunais constitucionais e internacionais. No contexto brasileiro e interamericano, o Supremo Tribunal Federal (STF) e a Corte Interamericana de Direitos Humanos (Corte IDH) são relevantes na construção de um corpo jurisprudencial que busca, não sem tensões e desafios, reconhecer a gravidade do fenômeno e a necessidade de impor limites a manifestações ofensivas quando estas implicam discriminação, violência simbólica, incitação à violência real e ameaça à própria dignidade humana e aos fundamentos do Estado Democrático de Direito.

No Brasil, a evolução da jurisprudência do STF sobre a liberdade

de expressão e seus limites revela uma trajetória de progressivo amadurecimento e de crescente sensibilidade aos impactos do discurso de ódio. Inicialmente, a Corte adotou uma postura marcadamente libertária, com forte ênfase na proteção ampla da liberdade de expressão, como ficou evidenciado no emblemático julgamento da Arguição de Descumprimento de Preceito Fundamental (ADPF) 130, em 2009. Nesta decisão, que declarou a não recepção da Lei de Imprensa (Lei nº 5.250/1967) pela Constituição de 1988, o relator, Ministro Ayres Britto, cunhou a célebre frase de que a liberdade de expressão é a “pedra de toque de todas as liberdades”, asseverando que qualquer forma de censura prévia é incompatível com o regime democrático inaugurado pela nova Carta (BRASIL, 2009). Essa posição inicial, embora fundamental para consolidar a liberdade de imprensa e de manifestação no pós-ditadura, foi sendo progressivamente tensionada e nuançada por casos concretos em que o discurso se apresentava não como expressão legítima de ideias, mas como veículo de ódio, exclusão, discriminação e ataque a direitos fundamentais de grupos minoritários.

Um exemplo significativo dessa ponderação entre a liberdade de expressão e outros direitos fundamentais ocorreu no julgamento da Ação Direta de Inconstitucionalidade (ADI) 4.815, em 2015, que discutiu a exigibilidade de autorização prévia da pessoa biografada (ou de seus familiares, em caso de falecimento) para a publicação de biografias. O STF, por unanimidade, julgou a ação procedente para dar interpretação conforme a Constituição aos artigos 20 e 21 do Código Civil, afastando a necessidade de consentimento prévio. A Corte ponderou que a liberdade

de expressão e o direito à informação histórica e cultural prevalecem sobre a exigência de autorização, ressalvando, contudo, a possibilidade de responsabilização posterior por eventuais abusos que configurem violação à honra, à imagem ou à privacidade (BRASIL, 2015a). Embora não se trate de um caso clássico de discurso de ódio, a ADI 4.815 é paradigmática por demonstrar como o STF lida com a colisão entre a liberdade de expressão e direitos da personalidade, estabelecendo que a liberdade não é absoluta e que abusos podem ser coibidos, especialmente quando atingem a dignidade da pessoa. Essa decisão reforça a ideia de que, mesmo em um contexto de ampla proteção à liberdade de manifestação, o ordenamento jurídico prevê mecanismos para coibir excessos e proteger direitos fundamentais que possam ser lesados, um raciocínio que se aplica, com ainda maior ênfase, aos casos de discurso de ódio.

Posteriormente, a jurisprudência do STF avançou de forma mais assertiva no enfrentamento direto ao discurso de ódio. O julgamento do Habeas Corpus (HC) 82.424/RS (Caso Ellwanger), em 2003, embora anterior à ADPF 130, já havia estabelecido um precedente fundamental ao confirmar a condenação de um editor por publicar livros de conteúdo antissemita e negacionista do Holocausto. A Corte entendeu que tais publicações configuravam crime de racismo, rechaçando a alegação de que estariam protegidas pela liberdade de expressão. O STF afirmou que o discurso de ódio racial não é liberdade de expressão, mas sim uma forma de agressão e discriminação que atenta contra a dignidade humana e os valores fundamentais da Constituição (BRASIL, 2003). Mais recentemente, na ADO 26, o Tribunal equiparou a homofobia e a

transfobia ao crime de racismo, reconhecendo o impacto do discurso de ódio na perpetuação da violência contra a população LGBTQIA+ (BRASIL, 2019b).

Em decisões relacionadas a inquéritos que apuram a disseminação de *fake news* e ataques às instituições democráticas (como o Inq. 4.781 e o Inq. 4.874), o STF tem reiterado que a liberdade de expressão não deve ser um meio para práticas antidemocráticas, discriminatórias, ou que violem direitos de terceiros, especialmente quando se trata de ataques coordenados contra minorias políticas e étnicas, ou contra a própria higidez do processo eleitoral e a estabilidade das instituições. A Corte tem enfatizado a necessidade de responsabilização por discursos que incitam à violência, à subversão da ordem democrática ou que promovem o ódio contra grupos vulneráveis, inclusive no ambiente digital (SARLET; NETO, 2021).

No plano internacional, a Corte Interamericana de Direitos Humanos, principal órgão jurisdicional do Sistema Interamericano de Proteção dos Direitos Humanos, adota igualmente uma posição firme e evolutiva contra o discurso de ódio, interpretando a Convenção Americana sobre Direitos Humanos (Pacto de San José da Costa Rica), especialmente seus artigos 13 (Liberdade de Pensamento e Expressão) e 1.1 (Obrigação de Respeitar os Direitos sem discriminação). O artigo 13.5 da Convenção proíbe expressamente “toda propaganda a favor da guerra, bem como toda apologia ao ódio nacional, racial ou religioso que constitua incitação à discriminação, à hostilidade, ao crime ou à violência”. A Corte IDH tem desenvolvido uma jurisprudência que busca compatibilizar a ampla

proteção à liberdade de expressão, essencial para a democracia, com a necessidade de coibir discursos que ameacem direitos de terceiros e a própria ordem pública democrática (CANÇADO TRINDADE, 2018).

Um precedente de grande relevância é o *Caso Fontevecchia e D'Amico vs. Argentina* (2011), no qual a Corte IDH analisou a responsabilidade ulterior decorrente de publicações jornalísticas que afetavam a privacidade de uma figura pública. Embora não se tratasse estritamente de discurso de ódio, a Corte estabeleceu importantes parâmetros sobre os limites da liberdade de expressão, reafirmando que esta não é um direito absoluto e pode ser restringida para proteger direitos de terceiros, desde que as restrições estejam previstas em lei, persigam um objetivo legítimo e sejam necessárias e proporcionais em uma sociedade democrática. No *Caso Kimel vs. Argentina* (2008), a Corte já havia ressaltado que as sanções penais por abuso da liberdade de expressão devem ser aplicadas com extrema cautela, reservando-as para as hipóteses mais graves, como a incitação à violência ou ao ódio (PIOVESAN, 2021).

No *Caso González Lluy et al. vs. Equador* (2015b), embora o foco principal fosse a discriminação contra uma criança vivendo com HIV, a Corte IDH salientou que a perpetuação de estigmas e preconceitos por parte de agentes estatais e pela sociedade em geral equivale a uma forma de violência simbólica e discriminatória, que viola a dignidade humana e o direito à igualdade (CORTE IDH, 2015b). Essa perspectiva é fundamental para compreender como o discurso de ódio, mesmo quando não incita diretamente à violência física imediata, pode criar um ambiente hostil e degradante que impede o pleno exercício de direitos por parte de

grupos vulnerabilizados.

Ambos os tribunais, STF e Corte IDH, apesar de suas distintas naturezas institucionais e dos contextos específicos em que atuam, convergem progressivamente para a ideia central de que a liberdade de expressão, pilar de qualquer sociedade democrática, não pode ser utilizada como um caminho lícito para corroer os fundamentos da própria democracia, do pluralismo, da igualdade e da proteção integral dos direitos humanos. O STF, em particular, tem demonstrado uma crescente disposição em dialogar com a jurisprudência da Corte Interamericana, em consonância com o disposto no art. 5º, §3º, da Constituição Federal, que atribui status de emenda constitucional aos tratados de direitos humanos ratificados pelo Brasil que sigam o rito qualificado, e reconhece a importância do controle de convencionalidade das leis e atos normativos internos (MAZZUOLI, 2021). Esse diálogo jurisprudencial é essencial para fortalecer o sistema de proteção de direitos no Brasil e para alinhar a interpretação constitucional interna aos mais elevados padrões internacionais de direitos humanos.

Portanto, a análise da jurisprudência brasileira e interamericana revela um esforço contínuo e complexo para delimitar os contornos do discurso de ódio e para estabelecer mecanismos eficazes de seu enfrentamento, sem que isso implique um retrocesso na proteção da liberdade de expressão legítima. A repressão a manifestações que comprovadamente incitam ao ódio, ao preconceito, à discriminação e à violência não representa censura indevida, mas sim a afirmação dos princípios da dignidade da pessoa humana, da igualdade material e da

vedação ao retrocesso civilizatório, que são pilares irrenunciáveis do Estado Democrático de Direito e da ordem internacional dos direitos humanos.

2.5.4 Um Modelo Normativo e Jurisprudencial Equilibrado: Desafios e Perspectivas na Era Digital

A construção de um modelo normativo e jurisprudencial que seja simultaneamente robusto na proteção da liberdade de expressão e eficaz no combate ao discurso de ódio representa um dos desafios mais prementes e complexos para as democracias constitucionais no século XXI, especialmente diante da velocidade e do alcance da disseminação de narrativas nocivas no ambiente digital. No contexto brasileiro, essa intrincada tarefa exige não apenas o aprimoramento da legislação existente, mas também uma reflexão crítica sobre o papel das plataformas digitais, a necessidade de maior transparência algorítmica e o fortalecimento de uma cultura de direitos humanos que permeie a sociedade. O objetivo último deve ser o de consolidar um marco regulatório e interpretativo que proteja vigorosamente o dissenso legítimo, a crítica e a pluralidade de ideias, ao mesmo tempo em que coíba de forma efetiva e proporcional as manifestações que visam à exclusão, à discriminação, à inferiorização e à deslegitimação de sujeitos e grupos sociais historicamente vulnerabilizados.

O ponto de partida fundamental para alcançar esse delicado equilíbrio reside na compreensão consolidada de que a liberdade de expressão, como já dito aqui, embora constitua um dos pilares essenciais da ordem democrática e um direito humano fundamental, não é um direito

absoluto ou ilimitado. Conforme estabelecido de forma clara no artigo 13 da Convenção Americana sobre Direitos Humanos (Pacto de San José da Costa Rica), à qual o Brasil é signatário e cujo conteúdo normativo integra o ordenamento jurídico nacional, o direito à liberdade de expressão pode e deve ser restringido por lei em circunstâncias específicas, desde que tais restrições sejam necessárias, proporcionais e estritamente voltadas à proteção dos direitos e da reputação de terceiros, à salvaguarda da segurança nacional, da ordem pública, ou da saúde e da moral públicas. Essa premissa é igualmente ecoada pela jurisprudência consolidada tanto do Supremo Tribunal Federal (STF) quanto da Corte Interamericana de Direitos Humanos (Corte IDH), que reconhecem a possibilidade de ponderação entre a liberdade de expressão e outros direitos fundamentais igualmente relevantes, como a dignidade da pessoa humana, a igualdade e a não discriminação (PIOVESAN, 2021; SARLET; NETO, 2021).

A dificuldade em obter ordens judiciais com a celeridade necessária e a ausência de deveres de cuidado mais proativos por parte das plataformas têm gerado um ambiente de relativa impunidade e de proliferação de conteúdos nocivos (LEMONS, 2019; BIONI, 2021). Essa constatação tem impulsionado um intenso debate público e diversas propostas legislativas em trâmite no Congresso Nacional, com destaque para o Projeto de Lei nº 2630/2020, conhecido como “PL das Fake News”, e outras iniciativas que buscam aprimorar a regulação do ambiente digital.

a) O Debate em Torno do PL 2630/2020 – “Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet”

O Projeto de Lei nº 2630/2020, que propõe a criação da Lei

Brasileira de Liberdade, Responsabilidade e Transparência na Internet, situa-se no cerne desse debate. Com o objetivo declarado de combater a desinformação e regular a atuação das grandes plataformas digitais, o projeto original e suas diversas versões subsequentes introduzem uma série de mecanismos e obrigações. Entre seus principais dispositivos, destacam-se a obrigatoriedade de elaboração e publicação de relatórios de transparência periódicos detalhando as medidas de moderação de conteúdo adotadas; a imposição de procedimentos de moderação de conteúdo mais claros, com dever de justificativa para remoções ou suspensões e a garantia de mecanismos de recurso acessíveis e eficazes para os usuários; o dever de disponibilizar canais de denúncia de conteúdo ilícito que sejam facilmente acessíveis e responsivos; a exigência de identificação de contas de grande alcance e de conteúdo impulsionado ou publicitário, especialmente em períodos eleitorais; e, crucialmente, a previsão de um regime de responsabilidade civil mais robusto para as plataformas em caso de omissão ou negligência diante da disseminação sistemática de conteúdo ilícito ou desinformativo que cause danos a terceiros. Embora o texto tenha sido objeto de intensas disputas políticas e diversas modificações ao longo de sua tramitação, a essência do PL 2630 reside na tentativa de estabelecer um regime de responsabilidade regulada e de deveres de diligência (*duty of care*) para os provedores de aplicação de grande porte, como Meta (Facebook, Instagram, WhatsApp), Google (YouTube), X (antigo Twitter) e TikTok. A proposta também prevê a possibilidade de o poder público, em circunstâncias excepcionais e mediante devido processo legal, exigir a retirada de conteúdos que configurem grave e iminente violação de direitos

fundamentais. As críticas ao projeto, por outro lado, alertam para os riscos de censura, de restrição indevida à liberdade de expressão e de criação de um ambiente regulatório excessivamente oneroso ou sujeito a instrumentalização política (COALIZÃO DIREITOS NA REDE, 2020; GOOGLE, 2023).

b) A Regulação da Inteligência Artificial e seus Impactos no Discurso de Ódio

Outra frente legislativa importante que dialoga diretamente com a problemática do discurso de ódio é a regulação da Inteligência Artificial (IA). Projetos como o 2338/2023 buscam estabelecer um marco legal para o desenvolvimento e uso de sistemas de IA no Brasil.

Esses textos geralmente propõem princípios fundamentais como a centralidade da pessoa humana, a não discriminação algorítmica, a transparência, a explicabilidade das decisões automatizadas, a segurança e a responsabilização por danos causados por sistemas de IA. A relevância dessa discussão para o enfrentamento ao discurso de ódio reside no reconhecimento de que os algoritmos utilizados pelas plataformas para curadoria, recomendação e moderação de conteúdo podem, inadvertida ou deliberadamente, amplificar vieses discriminatórios, criar bolhas de filtro, promover a polarização e facilitar a viralização de conteúdos odiosos e desinformativos (NOBLE, 2018; PASQUALE, 2015). Uma regulação da IA que imponha deveres de avaliação de impacto algorítmico, de auditoria externa e de transparência sobre os critérios utilizados pelos sistemas pode contribuir significativamente para mitigar esses riscos e para responsabilizar as plataformas quando seus sistemas automatizados

favorecerem a disseminação de conteúdos nocivos à dignidade humana, à igualdade e à integridade do debate público.

c) Propostas de Tipificação Penal Específica do Discurso de Ódio

Paralelamente às discussões sobre regulação de plataformas e IA, tramitam no Congresso Nacional diversas propostas que buscam incluir o discurso de ódio de forma mais explícita no Código Penal brasileiro ou criar tipos penais autônomos para criminalizar condutas específicas. Temos como exemplos o PL 385/2022 de autoria do Senador Rogério Carvalho, PL 1961/2022 do Senador Alexandre Silveira, PL 1636/2022 do Senador Randolfe Rodrigues, PL 836/2024 do Senador Carlos Viana e o PL 2821/2022 do Senador Fabiano Contarato. Algumas dessas proposições legislativas pretendem criminalizar a negação ou justificação de crimes contra a humanidade e genocídios reconhecidos, a incitação ao preconceito e à discriminação contra minorias com base em raça, gênero, orientação sexual, religião, origem nacional ou regional, e o uso sistemático e intencional de linguagem depreciativa e desumanizante em meios digitais contra grupos vulneráveis. Essas propostas, no entanto, demandam extremo cuidado técnico e um debate aprofundado para evitar a criação de tipos penais excessivamente amplos ou vagos, que poderiam ser utilizados para criminalizar o dissenso legítimo, a crítica social ou manifestações artísticas e culturais.

O uso de termos como “discurso ofensivo” ou “conteúdo impróprio”, sem uma delimitação precisa dos elementos objetivos e subjetivos do tipo, pode gerar insegurança jurídica e riscos à liberdade de expressão. A jurisprudência da Corte Europeia de Direitos Humanos

(CEDH) em casos como *Handyside vs. Reino Unido* e *Jersild vs. Dinamarca*, bem como a Recomendação Geral nº 35 do Comitê da ONU para a Eliminação da Discriminação Racial (CERD), oferecem critérios importantes para a delimitação do discurso de ódio passível de sanção penal, restringindo-o, em geral, às formas mais graves que incitam à violência, à discriminação ou à hostilidade e que representam um perigo claro e iminente (CERD, 2013; MEYER-PFLUG, 2009).

d) Comparativo com Modelos Internacionais: O DSA Europeu e o NetzDG Alemão.

No cenário internacional, o Regulamento dos Serviços Digitais (Digital Services Act - DSA) da União Europeia, que entrou em vigor progressivamente a partir de 2022, representa um marco importante na regulação de plataformas digitais. O DSA estabelece um conjunto abrangente de obrigações para intermediários online, com regras mais estritas para as plataformas de grande dimensão. Entre as medidas, destacam-se a maior transparência sobre publicidade online e sistemas de recomendação, a obrigação de combater a disseminação de conteúdo ilegal (incluindo discurso de ódio e desinformação), a criação de mecanismos de denúncia e recurso, e a exigência de avaliação de riscos sistêmicos (EUROPEAN COMMISSION, 2022). Outra referência frequentemente citada é a Lei Alemã de Aplicação na Rede (Netzwerkdurchsetzungsgesetz - NetzDG), de 2017, que impõe às redes sociais com mais de dois milhões de usuários na Alemanha a obrigação de remover conteúdo “claramente ilegal” (como incitação ao ódio e difamação) em até 24 horas após a notificação, sob pena de multas elevadas. Embora o NetzDG tenha sido

pioneiro, também recebeu críticas por potencialmente incentivar a remoção excessiva de conteúdo (over-removal) pelas plataformas para evitar sanções, gerando um efeito de “censura privada” (BBC NEWS, 2020).

Do ponto de vista jurisprudencial, é altamente desejável que o STF e os tribunais superiores brasileiros continuem a aprofundar e a consolidar critérios objetivos e reiterados para a distinção entre a manifestação crítica ou contundente, ainda que protegida pela liberdade de expressão, e o discurso de ódio, que extrapola esses limites e atenta contra direitos fundamentais. A incorporação de elementos de análise como (i) o contexto da manifestação, (ii) a intenção do emissor (dolo específico de odiar, discriminar ou incitar), (iii) as características do público-alvo (se pertencente a grupo historicamente vulnerabilizado), (iv) o conteúdo e a forma da mensagem (grau de virulência, potencial de dano), e (v) o impacto provável ou real sobre a participação democrática e a segurança dos grupos atingidos, pode contribuir para decisões mais consistentes e previsíveis. Esse esforço de uniformização interpretativa, que já se observa em alguns julgados, é basilar para reduzir a litigiosidade e garantir maior segurança jurídica a todos os envolvidos.

Ademais, o Poder Judiciário brasileiro deve fortalecer e intensificar sua articulação com a rica jurisprudência da Corte Interamericana de Direitos Humanos, que tem oferecido parâmetros sólidos e progressistas sobre os limites da expressão odiosa e sobre o dever positivo do Estado em prevenir, investigar, sancionar e reparar os danos causados por tais condutas. O artigo 5º, §2º e §3º, da Constituição Federal, ao prever a

incorporação dos direitos e garantias expressos em tratados internacionais em que o Brasil seja parte e o status supralegal (ou mesmo constitucional, a depender do rito de aprovação) desses tratados, legitima e impõe essa integração hermenêutica e o controle de convencionalidade das leis e das decisões judiciais internas (MAZZUOLI, 2021).

No plano legislativo, é fundamental que toda e qualquer tipificação penal ou regulação administrativa seja rigorosamente guiada pelos princípios da legalidade estrita, da necessidade, da proporcionalidade e da adequação constitucional, e que esteja organicamente articulada com políticas públicas abrangentes voltadas à educação em direitos humanos, à promoção da alfabetização midiática e digital, e à valorização da diversidade e do pluralismo. O enfrentamento ao discurso de ódio não se esgota na aplicação de sanções; exige, antes e acima de tudo, a construção de uma cultura democrática genuinamente inclusiva, que reconheça, respeite e proteja a pluralidade de vozes, identidades e perspectivas que compõem o multifacetado espaço público nacional.

Cabe ressaltar a importância da autorregulação responsável por parte das próprias plataformas digitais, sempre em estrita conformidade com os marcos constitucionais e legais de proteção aos direitos humanos. As empresas de tecnologia que operam no Brasil devem adotar, de forma transparente e proativa, mecanismos eficazes de moderação de conteúdo, baseados em termos de uso claros, públicos e não discriminatórios, com a disponibilização de mecanismos de recurso justos e céleres, explicações compreensíveis sobre as decisões algorítmicas que impactam a visibilidade e o alcance dos conteúdos, e a publicação de relatórios

periódicos de transparência detalhados e auditáveis. Essas medidas, quando acompanhadas de uma fiscalização institucional independente e efetiva por parte de órgãos estatais competentes, podem contribuir significativamente para evitar tanto a censura arbitrária e desproporcional quanto a omissão conivente e danosa diante da proliferação do discurso de ódio e da desinformação.

2.5.5 Educação Digital Crítica, Regulação Democrática das Plataformas e Promoção de uma Cultura de paz como pilares para o enfrentamento ao discurso de ódio

A resposta institucional e social ao fenômeno complexo e multifacetado do discurso de ódio não pode, sob pena de ineficácia, limitar-se exclusivamente à repressão jurídica ou à responsabilização *ex post facto* dos perpetradores. Embora os mecanismos de sanção penal e civil sejam componentes necessários de uma estratégia abrangente, eles se mostram inerentemente insuficientes para transformar as profundas estruturas sociais, culturais, econômicas e comunicacionais que historicamente sustentam e retroalimentam a intolerância, o preconceito, a discriminação e a disseminação deliberada de desinformação. Por essa razão fundamental, a construção de um modelo robusto, resiliente e duradouro de enfrentamento ao discurso de ódio exige um investimento estratégico e contínuo em abordagens preventivas, pedagógicas e emancipatórias, ancoradas em três pilares interdependentes e sinérgicos: o fomento de uma educação digital crítica e orientada para os direitos humanos, o estabelecimento de uma regulação democrática e transparente das tecnologias de informação e comunicação, com especial atenção às

plataformas digitais, e a promoção ativa e transversal de uma cultura de paz, respeito à diversidade e valorização da alteridade.

O primeiro e mais essencial desses pilares é a educação digital em e para os direitos humanos, compreendida não como um mero treinamento técnico para o uso de ferramentas digitais, mas como um processo formativo integral que visa desenvolver a capacidade crítica dos cidadãos para compreender, analisar e interagir de forma consciente, ética e responsável com o ecossistema informacional contemporâneo. Conforme aponta o sociólogo Boaventura de Sousa Santos (2009), não há cidadania plena e efetiva sem a capacidade crítica de decodificar os discursos que circulam nos múltiplos espaços públicos, de identificar as relações de poder subjacentes a eles e de resistir ativamente à manipulação simbólica e à dominação cultural (SANTOS, 2009).

Nesse sentido, é urgente e inadiável a implementação de programas abrangentes e sistemáticos de alfabetização midiática, informacional e digital em todos os níveis de ensino, desde a educação básica até o ensino superior, bem como em iniciativas de educação não formal e comunitária. Tais programas devem capacitar os cidadãos, especialmente crianças e jovens, a identificar as características e os impactos do discurso de ódio, a reconhecer práticas discriminatórias sutis e explícitas, a discernir entre liberdade de opinião legítima e incitação à violência ou à discriminação, a verificar a veracidade das informações antes de compartilhá-las e a utilizar as tecnologias digitais para promover narrativas de respeito, inclusão e justiça social. Iniciativas como o programa Educamídia, desenvolvido pelo Instituto Palavra Aberta, e as ações da Safernet Brasil na promoção da

cidadania digital e no combate a crimes cibernéticos, são exemplos relevantes de esforços nesse sentido no contexto brasileiro (EDUCAMÍDIA, [s.d.]; SAFERNET BRASIL, [s.d.]).

O Programa Mundial de Educação em Direitos Humanos, coordenado pela Organização das Nações Unidas (ONU) e pela UNESCO, recomenda expressamente que os Estados membros promovam, de forma transversal nos currículos escolares e em campanhas públicas de conscientização, o ensino de valores fundamentais como igualdade, dignidade humana, diversidade cultural, não discriminação e resolução pacífica de conflitos (ONU, 2012). No Brasil, esse objetivo pode e deve ser fortalecido por meio de políticas públicas articuladas entre o Ministério da Educação, o Ministério dos Direitos Humanos e da Cidadania, os sistemas estaduais e municipais de ensino, os conselhos de direitos, as defensorias públicas, as organizações da sociedade civil e os meios de comunicação, visando criar ambientes escolares e comunitários que cultivem a empatia, o pensamento crítico, o diálogo intercultural e o respeito incondicional à diferença.

Paralelamente ao investimento em educação, é fundamental avançar no desenvolvimento de mecanismos de regulação democrática das tecnologias de informação e comunicação, com um foco particular nas grandes plataformas digitais que hoje desempenham um papel central na mediação do debate público e na circulação de informações. A experiência regulatória de outras democracias, como a Alemanha com o NetzDG (Network Enforcement Act), a França com a Lei Avia (embora parcialmente invalidada pelo Conselho Constitucional) e, mais

recentemente, a União Europeia com o Digital Services Act (DSA) e o Digital Markets Act (DMA), demonstra que é possível e necessário adotar normas que estabeleçam deveres de transparência, diligência e responsabilidade para as plataformas, sem que isso implique censura prévia ou violação da liberdade de expressão (EUROPEAN COMMISSION, 2022; LEMONS, 2019).

Esses modelos regulatórios, ainda que com suas particularidades e desafios, geralmente envolvem a obrigação de as plataformas implementarem sistemas mais eficazes e transparentes de moderação de conteúdo ilícito (incluindo discurso de ódio), de fornecerem informações claras aos usuários sobre o funcionamento de seus algoritmos de recomendação e de publicidade, de estabelecerem canais de denúncia e recurso acessíveis e céleres, e de cooperarem com autoridades competentes na investigação de crimes e na proteção de direitos.

No Brasil, o debate em torno do PL 2630/2020 e de outras propostas de regulação de plataformas reflete essa busca por um equilíbrio entre a proteção da liberdade de expressão e a necessidade de coibir abusos e responsabilizar os atores que lucram com a disseminação de conteúdos nocivos. Uma regulação democrática eficaz deve ser construída com ampla participação social, envolvendo não apenas o Estado, mas também a academia, as organizações da sociedade civil, os especialistas em tecnologia e direitos humanos, e as próprias plataformas, em um modelo de governança multissetorial e dialógico (BODÓ et al., 2021). É imprescindível que qualquer marco regulatório seja orientado pelos princípios constitucionais da legalidade, necessidade, proporcionalidade,

devido processo legal, e que tenha como objetivo primordial a proteção dos direitos fundamentais e o fortalecimento do ambiente democrático.

O terceiro pilar, intrinsecamente ligado aos anteriores, é a promoção ativa e contínua de uma cultura de paz. A paz, como nos lembra Norberto Bobbio (2004, p. 34), não é meramente a ausência de guerra ou de conflito direto, mas a presença de justiça social, de igualdade de oportunidades, de reconhecimento recíproco e de respeito à dignidade de todos. Nesse sentido, uma cultura de paz pressupõe o enfrentamento sistemático de todas as formas de violência, incluindo as violências simbólicas, discursivas e institucionais que alimentam a exclusão, a discriminação e a desumanização do outro.

A UNESCO tem sido uma defensora proeminente da educação para a paz, os direitos humanos e a compreensão internacional, destacando que a paz duradoura se constrói nas mentes dos homens e mulheres através da educação, da ciência, da cultura e da comunicação (UNESCO, 2021). Isso implica fomentar valores como a tolerância (não no sentido de condescendência, mas de respeito ativo pela diversidade), a solidariedade, a cooperação, a empatia e a capacidade de resolver conflitos de forma não violenta. A promoção de uma cultura de paz envolve desconstruir estereótipos negativos, combater preconceitos arraigados, valorizar a diversidade cultural e religiosa, e promover narrativas que celebrem a interdependência humana e a riqueza do pluralismo. Iniciativas de justiça restaurativa, mediação de conflitos, diálogo intercultural e inter-religioso, e a promoção de uma memória histórica crítica e inclusiva são ferramentas importantes nesse processo.

É importante compreender que esses três pilares – educação digital crítica, regulação democrática e cultura de paz – não são soluções isoladas, mas componentes de uma estratégia integrada e de longo prazo. O discurso de ódio floresce em terrenos férteis de ignorância, desinformação, ressentimento, desigualdade social e desconfiança nas instituições – elementos que só podem ser efetivamente superados por meio de uma formação cidadã contínua, pluralista e inclusiva, que valorize a alteridade não como uma ameaça, mas como um princípio constitutivo e enriquecedor da experiência democrática.

A responsabilidade por essa transformação é compartilhada: cabe ao Estado garantir o acesso universal à educação de qualidade e criar marcos regulatórios justos e eficazes; às instituições de ensino, incorporar a educação para os direitos humanos e a cidadania digital em seus currículos e práticas pedagógicas; às plataformas digitais, assumir sua responsabilidade social e adotar práticas mais éticas e transparentes; aos meios de comunicação, promover um jornalismo responsável e comprometido com a verdade e o interesse público; e a cada cidadão, cultivar a empatia, o pensamento crítico e o compromisso com a construção de uma sociedade mais justa, pacífica e igualitária.

O Brasil, enquanto Estado Democrático de Direito, fundado na dignidade da pessoa humana e no pluralismo político, e como signatário de diversos tratados internacionais de proteção aos direitos humanos, tem o dever constitucional e ético de implementar essas diretrizes por meio de políticas públicas consistentes e efetivas, marcos normativos coerentes e decisões judiciais que reflitam um compromisso inequívoco com a

inclusão social e a proteção dos mais vulneráveis. O enfrentamento ao discurso de ódio, portanto, transcende a esfera puramente jurídica para se configurar como um projeto civilizatório fundamental, no qual o direito, a educação, a tecnologia e a política devem operar em profunda harmonia e sinergia para garantir um futuro onde a liberdade de expressão e a convivência pacífica e respeitosa possam florescer conjuntamente.

CAPÍTULO 03

METODOLOGIA

3 METODOLOGIA

3.1 TIPO DE PESQUISA

A presente dissertação caracteriza-se como uma pesquisa de natureza jurídica, com foco na análise de conceitos, normas e decisões judiciais. Considerando a complexidade do tema abordado – a proteção dos direitos humanos frente ao discurso de ódio nas redes sociais e a liberdade de expressão –, a pesquisa se insere no campo do Direito Constitucional e dos Direitos Humanos, com interfaces com o Direito Digital e a Sociologia do Direito.

3.2 ABORDAGEM DA PESQUISA

Quanto à abordagem, a pesquisa adota uma perspectiva qualitativa. Este tipo de abordagem é fundamental para a compreensão aprofundada dos fenômenos sociais e jurídicos, permitindo a interpretação de significados, valores e contextos que não seriam capturados por métodos quantitativos. A natureza qualitativa da pesquisa é essencial para analisar as nuances do discurso de ódio, as implicações da liberdade de expressão e a complexidade da responsabilização das plataformas digitais, que envolvem aspectos éticos, morais e sociais, além dos estritamente legais.

Com relação ao enfoque lógico, a pesquisa é predominantemente dedutiva. Partindo de premissas gerais – como os princípios constitucionais da liberdade de expressão e da dignidade da pessoa humana, bem como os compromissos internacionais de proteção aos direitos humanos –, busca-se analisar casos específicos e a jurisprudência

para verificar a aplicação e a efetividade dessas normas e princípios no contexto do discurso de ódio nas redes sociais no Brasil. A abordagem dedutiva permite a construção de argumentos sólidos e a formulação de conclusões que se derivam logicamente das premissas e da análise dos dados coletados.

3.3 PROCEDIMENTOS TÉCNICOS

Os procedimentos técnicos adotados para a realização desta pesquisa incluem:

Pesquisa Documental: Consistiu na análise de documentos oficiais, como leis, decretos, projetos de lei, relatórios governamentais e de organizações não governamentais, que abordam a liberdade de expressão, o discurso de ódio e a regulação das redes sociais no Brasil e em âmbito internacional. Esta etapa foi fundamental para compreender o arcabouço normativo e as políticas públicas existentes sobre o tema.

Pesquisa Bibliográfica: Envolveu o levantamento e a análise de vasta literatura científica, incluindo livros, artigos acadêmicos, teses e dissertações, de autores nacionais e estrangeiros, que tratam dos conceitos de liberdade de expressão, discurso de ódio, direitos humanos, direito digital e responsabilidade das plataformas. A pesquisa bibliográfica permitiu a construção de um referencial teórico sólido e a identificação das principais correntes de pensamento sobre o tema.

Pesquisa Comparada: Realizou-se uma análise comparativa de modelos regulatórios e jurisprudenciais adotados em outros países e sistemas jurídicos (especialmente o europeu e o interamericano de direitos

humanos) para lidar com o discurso de ódio online. Esta abordagem permitiu identificar boas práticas, desafios comuns e soluções inovadoras que podem servir de subsídio para o contexto brasileiro.

Pesquisa Jurisprudencial: Consistiu na análise de decisões de tribunais superiores brasileiros, com destaque para o Supremo Tribunal Federal (STF), e da Corte Interamericana de Direitos Humanos (Corte IDH), que abordam o conflito entre liberdade de expressão e discurso de ódio, bem como a responsabilização das plataformas digitais. A análise crítica da jurisprudência foi essencial para compreender a evolução do entendimento jurídico sobre o tema e as tendências de responsabilização.

Análise Crítica Pautada na Hermenêutica dos Direitos Fundamentais: Todos os dados e informações coletados foram submetidos a uma análise crítica, utilizando a hermenêutica dos direitos fundamentais como lente interpretativa. Esta abordagem permitiu ir além da mera descrição dos fatos e normas, buscando compreender os significados, os valores e os princípios subjacentes às questões jurídicas e sociais envolvidas, e propor soluções que garantam a efetiva proteção dos direitos humanos.

3.4 PERÍODO E ETAPAS DA PESQUISA

A pesquisa foi desenvolvida entre agosto de 2024 e agosto de 2025, período em que foram realizadas as seguintes etapas:

Fase Exploratória (Agosto - Outubro de 2024): Caracterizada pela leitura preliminar de textos e documentos sobre o tema, a fim de familiarizar-se com o objeto de estudo, identificar as principais questões e

delimitar o problema de pesquisa. Nesta fase, foram realizadas buscas iniciais em bases de dados acadêmicas e jurídicas para mapear a literatura existente.

Fase de Levantamento Bibliográfico e Documental (Novembro de 2024 - Fevereiro de 2025): Consistiu na coleta sistemática de dados por meio da pesquisa bibliográfica e documental. Foram consultadas bases de dados como Scielo, Google Scholar, periódicos jurídicos especializados, repositórios de teses e dissertações, além de documentos oficiais de órgãos governamentais e internacionais. A seleção dos materiais foi pautada pela relevância e atualidade em relação ao tema da dissertação.

Fase de Análise e Interpretação dos Dados (Março - Junho de 2025): Nesta etapa, os dados coletados foram organizados, categorizados e submetidos à análise crítica. A interpretação foi guiada pela hermenêutica dos direitos fundamentais, buscando identificar padrões, contradições, lacunas e tendências na legislação, doutrina e jurisprudência. Foram elaborados fichamentos, resumos e sínteses dos materiais para subsidiar a escrita da dissertação.

Fase de Redação e Revisão (Julho - Agosto de 2025): Corresponde à escrita dos capítulos da dissertação, com base na análise e interpretação dos dados. Após a redação inicial, o texto foi submetido a um processo de revisão e aprimoramento, com foco na clareza, coerência, coesão e rigor acadêmico, além da adequação às normas da ABNT.

3.5 INSTRUMENTOS DE COLETA DE DADOS

Os instrumentos de coleta de dados utilizados nesta pesquisa foram

essencialmente documentais e bibliográficos, dada a natureza da abordagem qualitativa e dedutiva. Não houve coleta de dados primários por meio de questionários, entrevistas ou observação direta, uma vez que o foco da pesquisa reside na análise de fontes secundárias e na interpretação de textos jurídicos e acadêmicos. Os principais instrumentos foram:

Legislação: Códigos, leis específicas (como o Marco Civil da Internet, Lei Caó), decretos e regulamentos pertinentes ao tema da liberdade de expressão, discurso de ódio e regulação de plataformas digitais no Brasil. Inclui-se também a análise de propostas legislativas em tramitação que visam aprimorar o arcabouço normativo.

Doutrina: Obras de juristas, filósofos e sociólogos do direito que abordam os fundamentos teóricos da liberdade de expressão, os limites ao discurso, o conceito de discurso de ódio, a dignidade da pessoa humana, a responsabilidade civil e penal no ambiente digital, e a regulação de plataformas. Foram consultados autores clássicos e contemporâneos, nacionais e internacionais.

Jurisprudência: Decisões de tribunais superiores brasileiros (Supremo Tribunal Federal - STF, Superior Tribunal de Justiça - STJ) e da Corte Interamericana de Direitos Humanos (Corte IDH). A análise incluiu acórdãos, votos e ementas, buscando identificar os argumentos jurídicos, os precedentes e as tendências interpretativas sobre o tema.

Relatórios e Documentos de Organismos Internacionais e ONGs: Relatórios da Organização das Nações Unidas (ONU), UNESCO, Conselho da Europa, SaferNet Brasil, entre outros, que oferecem dados,

análises e recomendações sobre o discurso de ódio, a desinformação e a proteção dos direitos humanos no ambiente digital.

Artigos Científicos e Teses/Dissertações: Publicações em periódicos acadêmicos e repositórios universitários que aprofundam aspectos específicos do tema, oferecendo diferentes perspectivas e resultados de pesquisas anteriores. A seleção desses materiais foi feita com base em sua relevância, rigor metodológico e contribuição para o debate.

3.6 ANÁLISE DOS DADOS

A análise dos dados coletados nesta pesquisa foi realizada por meio de uma abordagem qualitativa e interpretativa, pautada na análise de conteúdo e na hermenêutica jurídica. Não foram empregadas técnicas de análise estatística, uma vez que o objetivo não era quantificar fenômenos, mas sim compreender em profundidade os significados, as relações e as implicações dos conceitos, normas e decisões jurídicas. As etapas da análise incluíram:

Leitura Exaustiva e Flutuante: Inicialmente, todos os documentos e textos foram lidos de forma atenta para uma compreensão geral do conteúdo e para identificar os temas centrais e as categorias de análise relevantes para a pesquisa.

Categorização e Codificação: Os dados foram organizados em categorias temáticas predefinidas (como liberdade de expressão, discurso de ódio, responsabilidade das plataformas, jurisprudência do STF, jurisprudência da Corte IDH, educação digital, etc.) e outras que emergiram da própria leitura. A codificação permitiu agrupar informações

semelhantes e identificar padrões e divergências entre as fontes.

Análise Hermenêutica: A interpretação dos textos jurídicos e acadêmicos foi realizada sob a ótica da hermenêutica dos direitos fundamentais. Isso envolveu a busca pelo sentido e alcance das normas e decisões, considerando o contexto histórico, social e político em que foram produzidas, bem como os princípios e valores constitucionais e internacionais de direitos humanos. A análise buscou desvendar as tensões e os dilemas presentes no debate sobre o discurso de ódio e a liberdade de expressão.

Comparação e Contraste: Os diferentes pontos de vista apresentados na doutrina, as abordagens regulatórias em distintas jurisdições e as decisões judiciais foram comparados e contrastados para identificar convergências, divergências e lacunas. Esta etapa foi importante para a construção de uma argumentação crítica e para a proposição de um modelo regulatório equilibrado.

Síntese e Articulação: Por fim, os resultados da análise foram sintetizados e articulados de forma coerente e lógica, construindo-se os argumentos que sustentam as conclusões da dissertação. A redação buscou apresentar uma análise aprofundada e crítica do tema, contribuindo para o avanço do debate jurídico.

3.7 CONSIDERAÇÕES ÉTICAS

A realização desta pesquisa observou rigorosamente os princípios éticos inerentes à produção científica, especialmente no campo das Ciências Jurídicas e Sociais. Embora a pesquisa não envolva a coleta de

dados diretamente de seres humanos (entrevistas, questionários, etc.), o que dispensaria a submissão a um Comitê de Ética em Pesquisa, foram adotadas as seguintes diretrizes:

Respeito à Propriedade Intelectual: Todas as fontes consultadas, sejam elas bibliográficas, documentais ou jurisprudenciais, foram devidamente citadas e referenciadas, seguindo as normas da ABNT. O plágio, em qualquer de suas formas, foi veementemente evitado, garantindo a originalidade e a integridade da pesquisa.

Transparência e Clareza: A metodologia utilizada foi descrita de forma clara e detalhada, permitindo a reprodutibilidade da pesquisa por outros estudiosos. Os procedimentos de coleta e análise de dados foram explicitados, garantindo a transparência do processo de construção do conhecimento.

Objetividade e Imparcialidade: A análise dos dados foi conduzida com o máximo de objetividade e imparcialidade possível, buscando apresentar os diferentes pontos de vista e as evidências de forma equilibrada. A pesquisa não se pautou por preconceitos ou ideologias pré-concebidas, mas sim pela busca da verdade científica e pela construção de argumentos sólidos e fundamentados.

Relevância Social: A escolha do tema e a condução da pesquisa foram guiadas pela relevância social do problema do discurso de ódio nas redes sociais e sua interface com a liberdade de expressão e os direitos humanos. A pesquisa buscou contribuir para o debate público e para a proposição de soluções que beneficiem a sociedade como um todo, especialmente os grupos mais vulneráveis.

Ao seguir essas considerações éticas, esta dissertação busca não apenas contribuir para o avanço do conhecimento jurídico, mas também reforçar o compromisso com a integridade, a responsabilidade e o impacto social positivo da pesquisa acadêmica.

CAPÍTULO 04

RESULTADOS E DISCUSSÕES

4 RESULTADOS E DISCUSSÕES

4.1 ANÁLISE DOS RESULTADOS

4.1.1 A Natureza Multifacetada do Discurso de Ódio

A análise qualitativa dos dados, fundamentada na hermenêutica dos direitos fundamentais e acompanhada da revisão bibliográfica e documental, revelou uma compreensão aprofundada da natureza do discurso de ódio. Longe de ser uma mera manifestação ofensiva, controversa ou chocante, o discurso de ódio se desvela como uma forma insidiosa e perniciosa de violência simbólica e estrutural. Seu propósito primordial é a marginalização, a inferiorização, a desumanização e o silenciamento de grupos sociais específicos, atuando como um corrosivo que mina os próprios alicerces da democracia plural e da convivência pacífica. Essa constatação é basilar, pois desloca o debate de uma perspectiva meramente individual para uma dimensão coletiva e sistêmica, onde o dano não se restringe ao indivíduo diretamente atingido, mas se estende à coesão social e à estrutura democrática. A pesquisa demonstrou, de forma inequívoca, que a liberdade de expressão, embora seja um direito fundamental inalienável e uma condição *sine qua non* para o florescimento do debate público e da esfera democrática, não pode ser concebida como um direito absoluto e ilimitado. Seus limites são intrínsecos e explicitamente delineados tanto na Constituição Federal de 1988 quanto nos tratados internacionais de direitos humanos dos quais o Brasil é signatário. Consequentemente, o discurso de ódio é manifestamente incompatível com os valores e princípios que regem a dignidade da pessoa

humana, a igualdade, a não discriminação, a segurança pública e a paz social. A compreensão desses limites é vital para evitar que a liberdade de expressão seja instrumentalizada para fins que contradizem os próprios fundamentos de uma sociedade justa e equitativa.

4.1.2 O Impacto Desproporcional em Grupos Historicamente Vulneráveis

Um dos achados mais contundentes da pesquisa reside na constatação de que o discurso de ódio opera como uma poderosa e eficaz ferramenta social de exclusão, perpetuando e aprofundando desigualdades históricas e estruturais. Sua incidência e seus efeitos são desproporcionalmente sentidos por grupos que já se encontram em situação de vulnerabilidade social, econômica e política. A pesquisa identificou, de forma clara, que a população negra, mulheres, pessoas LGBTQIA+, povos indígenas, pessoas com deficiência, migrantes e refugiados são alvos frequentes e preferenciais dessas manifestações. No contexto brasileiro, essa lista se expande para incluir também os nordestinos e os adeptos de religiões de matriz africana, que historicamente enfrentam preconceitos e discriminações arraigados. Essa constatação não é meramente descritiva; ela sublinha a urgência e a imperatividade da adoção de respostas jurídicas, institucionais e sociais que sejam não apenas eficazes, mas também profundamente sensíveis e adaptadas a essas vulnerabilidades específicas. A abordagem interseccional torna-se, assim, indispensável para compreender a complexidade das opressões e para formular estratégias de combate ao discurso de ódio que sejam verdadeiramente inclusivas e protetivas.

4.1.3 A Convergência Jurisprudencial na Repressão ao Ódio

A investigação aprofundada da jurisprudência do Supremo Tribunal Federal (STF) e da Corte Interamericana de Direitos Humanos (Corte IDH) revelou uma notável e consistente evolução no tratamento jurídico do discurso de ódio. Ambas as cortes têm convergido para o entendimento de que a repressão a tais manifestações, quando devidamente balizada por critérios de proporcionalidade e razoabilidade, não representa, de forma alguma, uma negação ou um cerceamento ilegítimo da liberdade de expressão. Pelo contrário, essa repressão é reconhecida como uma condição necessária para a garantia da efetividade universal da própria liberdade de expressão e, mais fundamentalmente, para a proteção da dignidade e dos direitos de todos os indivíduos e grupos sociais. Casos emblemáticos no STF, como o julgamento do HC 82.424 (conhecido como Caso Ellwanger), que criminalizou o antissemitismo, a Ação Direta de Inconstitucionalidade por Omissão (ADO) 26, que reconheceu a criminalização da LGBTfobia, e o Inquérito das Fake News (Inq. 4.781), são marcos jurisprudenciais que solidificam essa interpretação. No âmbito da Corte IDH, julgados como *Kimel vs. Argentina* reforçam essa compreensão, estabelecendo precedentes importantes para a limitação da liberdade de expressão em face do discurso de ódio. Essa convergência demonstra uma maturidade jurídica na compreensão de que a liberdade de expressão não pode ser um salvo conduto para a incitação à violência e à discriminação.

4.1.4 Lacunas Normativas e a Urgência de Aprimoramento Legislativo

Apesar da existência de instrumentos legais importantes no ordenamento jurídico brasileiro, como a Lei nº 7.716/1989 (conhecida como Lei Caó), que define os crimes resultantes de preconceito de raça ou de cor, e o Marco Civil da Internet (Lei nº 12.965/2014), que estabelece princípios, garantias, direitos e deveres para o uso da internet no Brasil, a pesquisa identificou lacunas significativas. Essas lacunas impedem uma resposta jurídica eficaz e abrangente à complexidade e à dinamicidade do discurso de ódio no ambiente digital. O artigo 19 do Marco Civil da Internet, que trata da responsabilização civil das plataformas digitais por conteúdos de terceiros, é um ponto crítico, mesmo após o julgado do STF. A redação atual desse dispositivo tem gerado debates e controvérsias, sendo frequentemente apontada como um obstáculo à responsabilização efetiva das plataformas pela disseminação de conteúdos ilegais e danosos. Além disso, a ausência de uma tipificação penal clara e precisa do discurso de ódio no Código Penal brasileiro dificulta a aplicação da lei e a punição efetiva dos infratores. Essa imprecisão normativa pode levar a interpretações subjetivas e à impunidade, comprometendo a eficácia do sistema jurídico na proteção dos direitos humanos. Tais pontos, portanto, demandam uma urgente atualização normativa e um debate legislativo qualificado e participativo, que envolva diversos setores da sociedade e especialistas na área.

As propostas para o aprimoramento das políticas públicas, da regulação e da atuação judicial no enfrentamento ao discurso de ódio no Brasil, conforme delineado na pesquisa, incluem um investimento robusto

e contínuo em políticas públicas de educação digital e direitos humanos, com a inserção transversal de temas como pensamento crítico, alfabetização midiática e informacional, empatia e respeito à diversidade nos currículos escolares. A revisão, através do Congresso Nacional, do artigo 19 do Marco Civil da Internet é fundamental para estabelecer um regime de responsabilidade civil das plataformas digitais que contemple deveres de diligência (*duty of care*), transparência algorítmica e moderação de conteúdo mais proativa e eficaz, inspirando-se em modelos internacionais como o Digital Services Act (DSA) europeu, mas adaptado à realidade brasileira. A discussão sobre a tipificação penal específica do discurso de ódio deve ser baseada em critérios objetivos que considerem a intenção do agente, o conteúdo da mensagem, o contexto, o público-alvo e o potencial de dano, em conformidade com a jurisprudência da Corte IDH, evitando a criminalização de manifestações legítimas ou críticas. A atuação judicial deve ser pautada por parâmetros consistentes, e o fomento à autorregulação responsável e à correção, com participação da sociedade civil, é essencial para um modelo equilibrado.

4.2 DISCUSSÃO DAS IMPLICAÇÕES

4.2.1 A Ponderação de Direitos Fundamentais e a Essência da Democracia

A tensão intrínseca entre a liberdade de expressão e a imperiosa necessidade de combater o discurso de ódio constitui, sem dúvida, um dos eixos centrais e mais desafiadores desta dissertação. Os resultados da pesquisa, ao serem analisados sob uma ótica crítica e aprofundada,

reforçam de maneira categórica a compreensão de que, embora a liberdade de expressão seja um direito fundamental inalienável e um pilar insubstituível para a sustentação de qualquer regime democrático, ela não pode ser concebida como um direito absoluto e ilimitado. Pelo contrário, sua efetividade e legitimidade encontram balizas e limites claros na proteção de outros direitos fundamentais igualmente essenciais, como a dignidade da pessoa humana, o direito à não discriminação e a igualdade material. A jurisprudência consolidada tanto do Supremo Tribunal Federal (STF) quanto da Corte Interamericana de Direitos Humanos (Corte IDH) tem sido uníssona e coerente ao afirmar que a repressão ao discurso de ódio não se configura como um ato de censura ou um cerceamento ilegítimo da liberdade de manifestação do pensamento. Ao invés disso, é reconhecida como uma medida necessária e proporcional para garantir a efetividade da própria liberdade de expressão para todos os membros da sociedade, especialmente para os grupos historicamente vulneráveis que são sistematicamente alvo de tais manifestações odiosas. Essa perspectiva é essencial, pois ressalta que a liberdade de expressão, em uma sociedade democrática, não pode ser utilizada como uma proteção para a propagação de ideologias que negam a própria humanidade de determinados grupos, subvertendo o propósito original do direito.

Essa ponderação de valores, deve ser guiada por princípios basilares como a proporcionalidade e a razoabilidade. Não se trata, portanto, de uma supressão arbitrária da liberdade de pensamento ou de opinião, mas sim de uma intervenção legítima e necessária para coibir manifestações que, por sua natureza intrínseca, incitam à violência, ao

preconceito, à discriminação ou à exclusão social, transfigurando-se em verdadeiras formas de violência simbólica e estrutural. A distinção entre um discurso meramente ofensivo ou controverso, que pode e deve ser tolerado em uma sociedade democrática pluralista, e o discurso de ódio, que demanda uma resposta firme e inequívoca do Estado e da sociedade, é um desafio hermenêutico e prático de suma importância. Essa distinção exige a formulação e a aplicação de parâmetros claros, objetivos e transparentes para sua correta identificação e aplicação. A ausência de tais parâmetros pode, lamentavelmente, conduzir a interpretações subjetivas e, conseqüentemente, a abusos que comprometam tanto a liberdade de expressão legítima quanto a proteção incondicional dos direitos humanos. A complexidade reside em traçar essa linha tênue, garantindo que a repressão ao ódio não se torne um instrumento para silenciar vozes dissidentes ou minoritárias, mas sim uma ferramenta para proteger a integridade e a dignidade de todos.

4.2.2 Desafios e Propostas para o Aprimoramento do Marco Regulatório Brasileiro

Os resultados da pesquisa apontam para uma necessidade premente e urgente de aprimoramento do marco normativo brasileiro no que tange ao combate eficaz ao discurso de ódio no ambiente digital. Embora a Lei Caó (Lei nº 7.716/1989) e o Marco Civil da Internet (Lei nº 12.965/2014) representem avanços legislativos importantes e pioneiros em suas respectivas épocas, a complexidade e a dinamicidade vertiginosa do ambiente digital contemporâneo exigem a implementação de mecanismos mais específicos, abrangentes e sistemáticos. A insuficiência do artigo 19

do Marco Civil da Internet, que versa sobre a responsabilização das plataformas digitais por conteúdos de terceiros, emerge como um ponto crítico e um dos principais gargalos regulatórios. A pesquisa sugere, de forma contundente, que a revisão desse artigo é imperativa e deve estabelecer um regime de responsabilidade civil das plataformas que contemple, de forma explícita e vinculante, deveres de diligência (*duty of care*), transparência algorítmica e mecanismos de moderação de conteúdo que sejam mais proativos, eficazes e ágeis, especialmente em relação a conteúdos manifestamente ilegais e danosos. Essa revisão poderia inspirar-se em modelos regulatórios internacionais mais avançados, como o Digital Services Act (DSA) europeu, mas sempre com a devida adaptação e contextualização à realidade jurídica, social e cultural brasileira. A experiência europeia, por exemplo, demonstra a viabilidade de se impor obrigações claras às plataformas sem necessariamente comprometer a liberdade de expressão.

Outra implicação de relevância inquestionável é a discussão aprofundada e tecnicamente embasada sobre a tipificação penal específica do discurso de ódio. A ausência de uma definição clara, precisa e exaustiva no Código Penal brasileiro tem sido um fator que dificulta sobremaneira a aplicação da lei e a punição efetiva dos infratores, gerando um sentimento de impunidade e desproteção para as vítimas. A proposta de tipificação deve ser cuidadosamente elaborada, baseando-se em critérios objetivos que considerem a intenção do agente (dolo específico), o conteúdo da mensagem (incitação à violência, discriminação, etc.), o contexto em que a manifestação ocorre, o público-alvo a que se destina e o potencial real de

dano que o discurso pode causar. Essa abordagem deve estar em plena conformidade com a jurisprudência consolidada da Corte IDH, evitando-se, a todo custo, a criminalização de manifestações legítimas, críticas ou satíricas que não configurem incitação ao ódio. Além disso, a pesquisa aponta para a necessidade da criação de uma entidade reguladora independente ou do fortalecimento e aparelhamento de órgãos já existentes. Essa entidade teria a função de fiscalizar o cumprimento das obrigações legais pelas plataformas, garantindo a participação ativa da sociedade civil e a proteção efetiva dos direitos dos usuários, conferindo maior legitimidade e eficácia ao processo regulatório. A governança da internet e a regulação de plataformas digitais são temas complexos que exigem soluções inovadoras e colaborativas.

4.2.3 A Responsabilidade Ampliada das Plataformas Digitais na Era da Algoritmização

As redes sociais e as plataformas digitais, em um curto espaço de tempo, transformaram-se em verdadeiras arenas públicas contemporâneas, reconfigurando de maneira profunda as dinâmicas comunicacionais, sociais e políticas em escala global. Contudo, essa evolução vertiginosa trouxe consigo um desafio de proporções igualmente grandiosas: a amplificação exponencial do discurso de ódio, impulsionado, em grande medida, por algoritmos que, em sua busca incessante por maximizar o engajamento e a viralização, acabam por priorizar conteúdos polarizadores e, muitas vezes, danosos. A pesquisa destaca, com base em evidências robustas, que o ambiente digital, caracterizado por sua instantaneidade, alcance massivo e um anonimato relativo, potencializa de forma alarmante

a disseminação e o impacto deletério do discurso de ódio. Os algoritmos de recomendação e moderação de conteúdo das grandes plataformas digitais, frequentemente opacos em seu funcionamento e orientados por lógicas de maximização de lucros e engajamento, podem, de forma inadvertida ou até mesmo deliberada, contribuir para a formação de “bolhas de filtro” e “câmaras de eco”. Esses fenômenos isolam os usuários em universos informacionais restritos, amplificando narrativas extremistas e conspiratórias, em detrimento da qualidade, da pluralidade e da racionalidade do debate democrático. A arquitetura dessas plataformas, portanto, não é neutra; ela molda o tipo de conteúdo que é consumido e, conseqüentemente, o tipo de discurso que prevalece.

A responsabilização das plataformas digitais emerge, assim, como um ponto crucial e inadiável para a construção de um ambiente digital mais seguro e equitativo. A dissertação aponta, de forma categórica, que o atual regime de responsabilidade, especialmente o artigo 19 do Marco Civil da Internet, é manifestamente insuficiente para lidar com a complexidade e a escala do fenômeno do discurso de ódio online. É imperativo que as plataformas assumam um papel muito mais proativo e responsável na moderação de conteúdo e na prevenção da disseminação do discurso de ódio. Isso implica a implementação de deveres de diligência rigorosos, transparência algorítmica efetiva e mecanismos eficazes de remoção de conteúdo ilegal e danoso, com prazos e procedimentos claros. A discussão sobre modelos de correção, que envolvam a participação ativa da sociedade civil, de especialistas independentes e de órgãos públicos na definição de diretrizes e na fiscalização de seu cumprimento, surge como

uma alternativa promissora e democrática. Essa abordagem colaborativa busca equilibrar a liberdade de expressão com a proteção dos direitos humanos no ambiente digital, reconhecendo que a solução não reside apenas na regulação estatal, mas também na corresponsabilidade de todos os atores envolvidos. A governança da internet é um desafio global que exige soluções multinível e multissetoriais.

4.2.4 O Papel Essencial da Jurisprudência na Construção de Limites e Garantias

A análise exaustiva da jurisprudência do Supremo Tribunal Federal (STF) e da Corte Interamericana de Direitos Humanos (Corte IDH) é de fundamental importância para compreender a evolução e a consolidação do tratamento jurídico do discurso de ódio. Os resultados da pesquisa demonstram, de forma inequívoca, que ambas as cortes têm se posicionado de maneira consistente e convergente no sentido de que a repressão ao discurso de ódio não configura uma violação ilegítima da liberdade de expressão. Pelo contrário, é reconhecida como uma medida necessária e proporcional para proteger a dignidade humana e outros direitos fundamentais, como a igualdade, a não discriminação e a segurança pública. Casos paradigmáticos como o HC 82.424 (Caso Ellwanger) no STF, que resultou na criminalização do antissemitismo, e a Ação Direta de Inconstitucionalidade por Omissão (ADO) 26, que reconheceu a criminalização da LGBTfobia por omissão legislativa, são marcos jurídicos que consolidam a interpretação de que a liberdade de expressão não pode, em hipótese alguma, servir de escudo para a incitação à discriminação, à violência ou ao ódio contra grupos específicos. Esses

julgados demonstram o compromisso do Poder Judiciário em salvaguardar os valores constitucionais e os direitos humanos.

Na esfera internacional, a Corte IDH, em julgados de grande repercussão como *Kimel vs. Argentina* e *Álvarez Ramos vs. Venezuela*, tem reforçado a compreensão de que a liberdade de expressão, embora seja um pilar democrático, possui limites intrínsecos e que o discurso de ódio, por sua natureza intrínseca e seu potencial deletério, não está protegido por essa garantia. A Corte tem enfatizado a necessidade imperiosa de os Estados adotarem medidas efetivas para combater o discurso de ódio, sem, contudo, cair em arbitrariedades ou excessos que possam cercear a liberdade de expressão legítima e plural. Essa convergência de entendimentos entre as cortes nacional e internacional oferece um arcabouço sólido e coerente para a construção de um modelo regulatório que concilie, de forma equilibrada e justa, a proteção à liberdade de expressão com o combate efetivo ao discurso de ódio. Esse modelo deve ser pautado pelos princípios da proporcionalidade e da razoabilidade, garantindo que as intervenções sejam estritamente necessárias e adequadas aos fins que se propõem. A atuação judicial, portanto, deve ser orientada por parâmetros objetivos e consistentes, considerando o contexto específico da manifestação, a gravidade do discurso, a intencionalidade do emissor, o impacto real e potencial sobre grupos vulneráveis e o potencial de incitação à violência ou à discriminação. A capacitação contínua de magistrados e operadores do direito é fundamental para a aplicação desses parâmetros em um cenário digital em constante mutação.

CONSIDERAÇÕES FINAIS

CONSIDERAÇÕES FINAIS

A presente dissertação de mestrado debruçou-se sobre uma problemática central e cada vez mais premente no contexto das democracias contemporâneas: como o ordenamento jurídico brasileiro pode efetivamente equilibrar a salvaguarda da liberdade de expressão, pilar fundamental do Estado Democrático de Direito, com a imperiosa necessidade de combater o discurso de ódio disseminado, sobretudo, por meio das redes sociais digitais, tudo isso em estrito respeito aos princípios constitucionais e aos compromissos internacionais de proteção aos direitos humanos assumidos pelo Brasil? A hipótese norteadora que guiou a investigação foi a de que a efetiva proteção aos direitos humanos no país não apenas permite, mas demanda a construção de parâmetros jurídicos mais claros, robustos e eficazes para a limitação da liberdade de expressão quando esta transborda para o campo do discurso de ódio. Tal empreitada exige não somente a elaboração de normas mais específicas e adequadas à complexidade do fenômeno, mas também uma maior e mais efetiva responsabilização das plataformas digitais, que desempenham um papel crucial na arquitetura e na dinâmica do ecossistema informacional contemporâneo.

Ao longo dos capítulos, a partir de uma análise aprofundada da doutrina nacional e internacional, da legislação pertinente e da jurisprudência consolidada do Supremo Tribunal Federal (STF) e da Corte Interamericana de Direitos Humanos (Corte IDH), constatou-se que o discurso de ódio transcende a mera expressão ofensiva, controversa ou

chocante. Configura-se, em verdade, como uma sofisticada e perniciosa forma de violência simbólica e estrutural, cujo objetivo primordial é marginalizar, inferiorizar, desumanizar e silenciar determinados grupos sociais, minando, por conseguinte, os próprios alicerces da democracia plural e da convivência pacífica. A liberdade de expressão, embora consagrada como direito fundamental e condição substancial para o debate público e a formação da vontade popular, não é, e nunca foi, um direito absoluto ou ilimitado. Conforme reiteradamente afirmado pelas mais altas cortes, ela encontra limites iminentes e explícitos na própria Constituição Federal de 1988 e nos tratados internacionais de direitos humanos, sendo manifestamente incompatível com discursos que atentem contra a dignidade da pessoa humana, a igualdade, a não discriminação, a segurança pública e a paz social (SARLET; NETO, 2021; PIOVESAN, 2021).

As principais conclusões que emergiram desta pesquisa podem ser sintetizadas nos seguintes pontos nodais:

A liberdade de expressão, em uma perspectiva constitucional e convencional, deve ser interpretada de forma harmônica e sistêmica com os demais direitos fundamentais. Sua invocação não pode servir de blindagem para justificar ou legitimar práticas de incitação à violência, ao preconceito, à discriminação ou à exclusão social. A ponderação de valores, guiada pelos princípios da proporcionalidade e da razoabilidade, é ferramenta hermenêutica indispensável nesse processo (ALEXY, 2015).

Nesse contexto, o discurso de ódio opera como uma potente tecnologia social de exclusão e de manutenção de desigualdades históricas,

afetando de maneira desproporcional grupos historicamente vulnerabilizados, como a população negra, mulheres, pessoas LGBTQIA+, povos indígenas, pessoas com deficiência, migrantes, refugiados e, no contexto brasileiro, também os nordestinos e adeptos de religiões de matriz africana. Essa constatação exige a adoção de respostas jurídicas, institucionais e sociais específicas e sensíveis a essas vulnerabilidades (DAVIS, 2017; MBEMBE, 2018).

Adicionalmente, o ambiente digital, com suas características de instantaneidade, alcance massivo e anonimato relativo, potencializa exponencialmente a disseminação e o impacto do discurso de ódio. Os algoritmos de recomendação e moderação de conteúdo das grandes plataformas digitais, muitas vezes opacos e orientados por lógicas de maximização do engajamento e da viralização, podem inadvertida ou deliberadamente contribuir para a criação de bolhas de filtro, câmaras de eco e para a amplificação de narrativas extremistas, em detrimento da qualidade e da pluralidade do debate democrático (ZUBOFF, 2020; NOBLE, 2018).

Diante desse cenário, a jurisprudência do Supremo Tribunal Federal, em casos emblemáticos como o HC 82.424 (Caso Ellwanger), a ADO 26 (criminalização da LGBTfobia) e o Inquérito das Fake News (Inq. 4.781), e da Corte Interamericana de Direitos Humanos, em julgados como *Kimel vs. Argentina*, tem evoluído consistentemente no sentido de reconhecer que a repressão ao discurso de ódio, quando devidamente balizada, não representa uma negação da liberdade de expressão, mas, ao contrário, uma condição essencial para a sua efetividade universal e para a

proteção da dignidade de todos os indivíduos e grupos.

Contudo, a legislação brasileira, embora conte com instrumentos importantes como a Lei nº 7.716/1989 (Lei Caó) e o Marco Civil da Internet (Lei nº 12.965/2014), ainda carece de mecanismos mais específicos, abrangentes e sistemáticos para lidar com a complexidade e a dinamicidade do discurso de ódio no ambiente digital. A insuficiência do artigo 19 do MCI no que tange à responsabilização das plataformas e a ausência de uma tipificação penal clara e precisa do discurso de ódio são lacunas que demandam urgente atualização normativa e um debate legislativo qualificado e participativo.

Diante desse diagnóstico, e em consonância com os objetivos desta dissertação, apresentam-se as seguintes propostas para o aprimoramento das políticas públicas, da regulação e da atuação judicial no enfrentamento ao discurso de ódio no Brasil:

No âmbito das Políticas Públicas e da Educação: É necessário o investimento robusto e contínuo em políticas públicas de educação digital em e para os direitos humanos, com a inserção transversal nos currículos escolares, desde a educação básica, de temas como pensamento crítico, alfabetização midiática e informacional, empatia, respeito à diversidade, combate à desinformação e reconhecimento de discursos discriminatórios e de ódio. Campanhas de conscientização pública, utilizando linguagem acessível e diversos canais de comunicação, também são fundamentais para fomentar uma cultura de respeito e não violência online e offline. O fortalecimento de programas como o Educamídia e o apoio a iniciativas da sociedade civil que promovem a cidadania digital são estratégias

promissoras (UNESCO, 2021).

No âmbito da Regulação e da Legislação: Urge a adoção de medidas legislativas claras, precisas e proporcionais que aprimorem o arcabouço normativo brasileiro. Isso inclui: (a) a revisão do artigo 19 do Marco Civil da Internet, para estabelecer um regime de responsabilidade civil das plataformas digitais que contemple deveres de diligência (*duty of care*), transparência algorítmica e moderação de conteúdo mais proativa e eficaz, especialmente em relação a conteúdos manifestamente ilegais e danosos, inspirando-se em modelos como o Digital Services Act (DSA) europeu, mas adaptado à realidade brasileira; (b) a discussão aprofundada e tecnicamente embasada sobre a tipificação penal específica do discurso de ódio, com base em critérios objetivos que considerem a intenção do agente, o conteúdo da mensagem, o contexto, o público-alvo e o potencial de dano, em conformidade com a jurisprudência da Corte IDH, evitando-se a criminalização de manifestações legítimas ou críticas; (c) a criação de uma entidade reguladora independente ou o fortalecimento de órgãos existentes para fiscalizar o cumprimento das obrigações legais pelas plataformas, garantindo a participação da sociedade civil e a proteção dos direitos dos usuários.

No âmbito da Atuação Judicial: O Poder Judiciário desempenha um papel primordial na interpretação e aplicação das normas, devendo pautar sua atuação por parâmetros objetivos e consistentes, como o contexto da manifestação, a gravidade do discurso, a intencionalidade do emissor, o impacto sobre grupos vulneráveis e o potencial de incitação à violência ou à discriminação. É fundamental que os magistrados e demais operadores

do direito busquem uma formação continuada sobre as dinâmicas do ambiente digital e a jurisprudência internacional de direitos humanos, aplicando o controle de convencionalidade e evitando decisões contraditórias, excessivamente restritivas ou politicamente motivadas. A promoção do diálogo entre as cortes e a academia pode contribuir para a construção de uma jurisprudência mais sólida e previsível.

Fomento à Autorregulação Responsável e à Corregulação: Incentivar as plataformas digitais a desenvolverem e aprimorarem seus mecanismos de autorregulação, com termos de serviço claros, políticas de moderação transparentes e processos de recurso justos e acessíveis, é uma via complementar importante. Modelos de correção, que envolvam a participação da sociedade civil, de especialistas independentes e de órgãos públicos na definição de diretrizes e na fiscalização de seu cumprimento, podem oferecer um caminho equilibrado e democrático.

Por fim, esta pesquisa reconhece que o tema aqui abordado é intrinsecamente dinâmico e complexo, estando em constante evolução com o rápido avanço das tecnologias de comunicação e os contínuos rearranjos do espaço público digital. Nesse sentido, diversas e instigantes linhas futuras de investigação se apresentam como necessárias e promissoras, tais como:

A análise aprofundada do impacto da inteligência artificial generativa, dos *deepfakes* e de outras tecnologias emergentes na produção, disseminação e amplificação de discursos de ódio e desinformação, bem como os desafios regulatórios e éticos daí decorrentes.

O estudo do papel e das dinâmicas das redes sociais

descentralizadas, das plataformas baseadas em *blockchain* e de outros ambientes digitais menos centralizados na propagação de conteúdos odiosos e na dificuldade de moderação efetiva.

A investigação comparada e crítica entre os modelos regulatórios de enfrentamento ao discurso de ódio adotados por países do Sul Global e os padrões normativos e jurisprudenciais desenvolvidos no contexto europeu e norte-americano, buscando soluções adaptadas às realidades locais.

A pesquisa interdisciplinar sobre os efeitos psicossociais, emocionais e comportamentais do discurso de ódio em indivíduos e grupos vulneráveis, com um enfoque interseccional que considere as múltiplas dimensões da discriminação (raça, gênero, orientação sexual, classe, etc.).

O desenvolvimento de métricas e metodologias mais eficazes para monitorar a prevalência e o impacto do discurso de ódio online, bem como para avaliar a efetividade das diferentes estratégias de prevenção e combate.

Reafirma-se, portanto, a convicção de que o discurso de ódio representa não um excesso tolerável ou um subproduto inevitável da liberdade de expressão, mas sua própria negação e perversão, na medida em que busca silenciar, excluir e aniquilar a dignidade do outro. Enfrentá-lo de maneira firme, equilibrada e democraticamente orientada é, pois, um compromisso inadiável e um imperativo ético-jurídico de qualquer ordem constitucional que se pretenda verdadeiramente pluralista, inclusiva e protetora da dignidade inerente a cada pessoa humana.



REFERÊNCIAS BIBLIOGRÁFICAS

REFERÊNCIAS BIBLIOGRÁFICAS

ABRADEP (Associação Brasileira de Direito Eleitoral e Político). **Direitos políticos, liberdade de expressão e discurso de ódio**. v. V. Brasília: ABRADEP, 2022.

ALEXY, Robert. **Teoria dos direitos fundamentais**. Tradução de Virgílio Afonso da Silva. 2. ed. São Paulo: Malheiros, 2015.

ARENDT, Hannah. **A Condição Humana**. Tradução de Roberto Raposo. 8. ed. Rio de Janeiro: Forense Universitária, 1997.

ÁVILA, Humberto. **Teoria dos Princípios: da definição à aplicação dos princípios jurídicos**. 12. ed. São Paulo: Malheiros, 2011.

BARROSO, Luís Roberto. **A dignidade da pessoa humana no direito constitucional contemporâneo: natureza jurídica, conteúdos mínimos e critérios de aplicação**. Belo Horizonte: Fórum, 2010.

BARROSO, Luís Roberto. **Interpretação e aplicação da Constituição: fundamentos de uma dogmática constitucional transformadora**. 8. ed. São Paulo: Saraiva Educação, 2017.

BBC NEWS. **A controversa lei alemã que inspira projeto de lei das Fake News**. BBC News Brasil, 26 ago. 2020. Disponível em: <https://www.bbc.com/portuguese/brasil-53914408>. Acesso em: 6 maio 2025.

BIONI, Bruno Ricardo. **Proteção de dados pessoais: a função e os limites do consentimento**. 2. ed. São Paulo: Forense, 2021.

BITENCOURT, Cezar Roberto. **Tratado de Direito Penal: Parte Especial 2**. 21. ed. São Paulo: Saraiva Educação, 2021.

BOBBIO, Norberto. **A era dos direitos**. Tradução de Carlos Nelson Coutinho. Rio de Janeiro: Elsevier, 2004.

BODÓ, Balázs et al. **Tech Platforms and Democratic Discourse: How**

to Fix Social Media by Regulation. Amsterdam: Institute for Information Law (IViR), University of Amsterdam, 2021.

BONAVIDES, Paulo. **Curso de Direito Constitucional.** 20. ed. São Paulo: Malheiros, 2004.

BOURDIEU, Pierre. **O poder simbólico.** Tradução de Fernando Tomaz. Rio de Janeiro: Bertrand Brasil, 1989.

BRASIL. **Constituição da República Federativa do Brasil de 1988.** Brasília, DF: Presidência da República, [2023]. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 4 maio 2025.

BRASIL. **Decreto nº 678, de 6 de novembro de 1992.** Promulga a Convenção Americana sobre Direitos Humanos (Pacto de São José da Costa Rica), de 22 de novembro de 1969. Brasília, DF: Presidência da República, [1992]. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto/d0678.htm. Acesso em: 4 maio 2025.

BRASIL. **Decreto nº 10.932, de 10 de janeiro de 2022.** Promulga a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, firmada pela República Federativa do Brasil, em La Antigua, Guatemala, em 5 de junho de 2013. *Diário Oficial da União*, Brasília, DF, 11 jan. 2022. Seção 1, p. 1. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2022/decreto/D10932.htm. Acesso em: 6 maio 2025.

BRASIL. **Decreto nº 65.810, de 8 de dezembro de 1969.** Promulga a Convenção Internacional sobre a Eliminação de todas as Formas de Discriminação Racial. *Diário Oficial da União*, Brasília, DF, 10 dez. 1969.

BRASIL. **Lei nº 7.716, de 5 de janeiro de 1989.** Define os crimes resultantes de preconceito de raça ou de cor. Brasília, DF: Presidência da República, [1989]. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l7716.htm. Acesso em: 4 maio 2025.

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011.** Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal. Brasília, DF: Presidência da República, [2011]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 4 maio 2025.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014.** Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília, DF: Presidência da República, [2014]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm. Acesso em: 4 maio 2025.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, [2018]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 4 maio 2025.

BRASIL. **Projeto de Lei nº 2630/2020.** Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Senado Federal. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/141944>. Acesso em: 4 maio 2025.

BRASIL. **Projeto de Lei nº 2338/2023.** Dispõe sobre o uso da Inteligência Artificial. Senado Federal. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 6 maio 2025.

BRASIL. Supremo Tribunal Federal. **Ação Direta de Inconstitucionalidade 4.815/DF.** Relatora: Ministra Cármen Lúcia. Julgamento em 10 jun. 2015a.

BRASIL. Supremo Tribunal Federal. **Ação Direta de Inconstitucionalidade por Omissão nº 26/DF.** Relator: Min. Celso de Mello. Brasília, DF, 13 de junho de 2019.

BRASIL. Supremo Tribunal Federal. **Arguição de Descumprimento de Preceito Fundamental nº 130/DF.** Relator: Min. Carlos Ayres Britto. Brasília, DF, 30 de abril de 2009.

BRASIL. Supremo Tribunal Federal. **Arguição de Descumprimento de Preceito Fundamental nº 187/DF**. Relator: Min. Celso de Mello. Brasília, DF, 15 de junho de 2011.

BRASIL. Supremo Tribunal Federal. **Decisão sobre suspensão da plataforma X. Agosto de 2024**.

BRASIL. Supremo Tribunal Federal. **Habeas Corpus nº 82.424/RS**. Relator: Min. Moreira Alves. Redator para Acórdão: Min. Maurício Corrêa. Brasília, DF, 17 de setembro de 2003.

BRASIL. Supremo Tribunal Federal. **Inquérito nº 4.781/DF**. Relator: Min. Alexandre de Moraes. Em andamento.

BRASIL. Supremo Tribunal Federal. **Mandado de Injunção 4.733/DF**, de 13 de junho de 2019. Disponível em: <https://portal.stf.jus.br/processos/detalhe.asp?incidente=4239576>. Acesso em: 23 maio 2022.

BRASIL. Supremo Tribunal Federal. **Recurso Extraordinário nº 511.961/SP**. Relator: Min. Carlos Ayres Britto. Brasília, DF, 17 de junho de 2009.

BRASIL. Supremo Tribunal Federal. **Informação à Sociedade: RE 1.037.396 (Tema 987) e 1.057.258 (Tema 533), Responsabilidade de plataformas digitais por conteúdo de terceiros**. Relator: Min. Dias Toffoli. Repercussão Geral. STF, Brasília, 26 jun 2025. Disponível em: https://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/Informac807a771oa768SociedadeArt19MCI_vRev.pdf. Acesso em: 10 agosto 2025.

BRASIL. Supremo Tribunal Federal. **STF define parâmetros para responsabilização de plataformas por conteúdos de terceiros**. STF, Brasília, 26 jun 2025. Disponível em: <https://noticias.stf.jus.br/postsnoticias/stf-define-parametros-para-responsabilizacao-de-plataformas-por-conteudos-de-terceiros/>. Acesso em: 14 agosto 2025.

BRASIL DE FATO. **Xenofobia contra nordestinos revela a forte presença do racismo no Brasil, dizem especialistas**. Brasil de Fato, São

Paulo, 7 out. 2022. Disponível em: <https://www.brasildefato.com.br/2022/10/07/xenofobia-contra-nordestinos-revela-a-forte-presenca-do-racismo-no-brasil-dizem-especialistas>. Acesso em: 12 maio 2025.

CANÇADO TRINDADE, Antônio Augusto. **O legado da Declaração Universal de 1948 e o futuro da proteção internacional dos direitos humanos**. Brasília: FUNAG, 2018.

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

CAMPOS, Tiago Soares. "**Misoginia**". Brasil Escola, 2025. Disponível em: <https://brasilecola.uol.com.br/sociologia/misoginia.htm>. Acesso em 18 de agosto de 2025.

CALHOUN, Craig (Ed.). **Habermas and the Public Sphere**. Cambridge, MA: MIT Press, 1992.

CETIC.BR. **TIC Domicílios - 2024 Indivíduos**. Cetic.BR. Disponível em: <https://cetic.br/pt/tics/domicilios/2024/individuos/>. Acesso em 12 de julho de 2025.

CITRON, Danielle Keats. **Technological Due Process**. Washington University Law Review, v. 85, n. 6, p. 1249-1313, 2008.

COALIZÃO DIREITOS NA REDE. **Nota Técnica sobre os problemas centrais do PL 2630/2020**. 23 jun. 2020. Disponível em: <https://direitosnarede.org.br/2020/06/23/nota-tecnica-da-coalizacao-direitos-na-rede-sobre-problemas-centrais-do-pl-2630-2020/>. Acesso em: 16 maio 2025.

COMPARATO, Fábio Konder. **A afirmação histórica dos direitos humanos**. 3. ed. rev. e ampl. São Paulo: Saraiva, 2003

CONSELHO DA EUROPA. **Convenção para a Proteção dos Direitos Humanos e das Liberdades Fundamentais, conforme emendada pelos Protocolos nº 11 e nº 14**. Roma, 4 nov. 1950. Disponível em: https://www.echr.coe.int/documents/convention_por.pdf. Acesso em: 4

maio 2025.

CORTE INTERAMERICANA DE DIREITOS HUMANOS. **Cadernos de Jurisprudência: Liberdade de Expressão**. San José, 2022. Disponível em:

https://www.corteidh.or.cr/sitios/libros/todos/docs/cuadernillo36_2022_port1.pdf. Acesso em: 14 agosto 2025

CORTE INTERAMERICANA DE DIREITOS HUMANOS. **Caso Álvarez Ramos vs. Venezuela**. Sentença de 30 de agosto de 2019 (Exceções Preliminares, Mérito, Reparações e Custas). Série C N° 380. Disponível em:

https://www.corteidh.or.cr/docs/casos/articulos/seriec_380_esp.pdf. Acesso em: 4 maio 2025.

CORTE INTERAMERICANA DE DIREITOS HUMANOS. **Caso González Lluy e outros vs. Equador**. Sentença de 1 de setembro de 2015 (Exceções Preliminares, Mérito, Reparações e Custas). Série C N° 298. Disponível em:

https://www.corteidh.or.cr/docs/casos/articulos/seriec_298_esp.pdf. Acesso em: 4 maio 2025.

CORTE INTERAMERICANA DE DIREITOS HUMANOS. **Caso Kimel vs. Argentina**. Sentença de 2 de maio de 2008 (Mérito, Reparações e Custas). Série C N° 177. Disponível em:

https://www.corteidh.or.cr/docs/casos/articulos/seriec_177_esp.pdf. Acesso em: 4 maio 2025.

CORTE INTERAMERICANA DE DIREITOS HUMANOS. **Caso Fontevecchia e D'Amico vs. Argentina**. Sentença de 29 de novembro de 2011 (Mérito, Reparações e Custas). Série C N° 238. Disponível em:

https://www.corteidh.or.cr/docs/casos/articulos/seriec_238_esp.pdf. Acesso em: 4 maio 2025.

CORTE INTERAMERICANA DE DIREITOS HUMANOS. **Caso Usón Ramírez vs. Venezuela**. Sentença de 20 de novembro de 2009 (Exceção Preliminar, Mérito, Reparações e Custas). Série C N° 207. Disponível em:

https://www.corteidh.or.cr/docs/casos/articulos/seriec_207_esp.pdf. Acesso em: 4 maio 2025.

CORTE INTERAMERICANA DE DIREITOS HUMANOS. **Opinião Consultiva OC-5/85, de 13 de novembro de 1985.** A Colegição Obrigatória de Jornalistas (Arts. 13 e 29 da Convenção Americana sobre Direitos Humanos). Série A N° 5. Disponível em: https://www.corteidh.or.cr/docs/opiniones/seriea_05_por.pdf. Acesso em: 4 maio 2025.

DAVIS, Angela. **Mulheres, raça e classe.** Tradução de Heci Regina Candiani. São Paulo: Boitempo, 2017.

DIAS, Matheus Felipe Gomes. **"O que é racismo?"**; Brasil Escola, 2025. Disponível em: <https://brasilecola.uol.com.br/o-que-e/o-que-e-sociologia/o-que-e-racismo.htm>. Acesso em 18 de agosto de 2025.

DONEDA, Danilo. **Da privacidade à proteção de dados pessoais.** São Paulo: Thomson Reuters Brasil, 2020.

DWORKIN, Ronald. **Levando os direitos a sério.** Tradução de Nelson Boeira. São Paulo: Martins Fontes, 2002.

ECO, Umberto. **Entrevista ao jornal La Stampa.** Turim, 10 jun. 2015. Disponível em: <https://www.lastampa.it/cultura/2015/06/11/news/umberto-eco-con-i-social-parola-a-legioni-di-imbecilli-1.35250428/>. Acesso em: 23 maio 2025.

EDUCAMÍDIA. **O que é educação midiática?** [S.l.: s.n., s.d.]. Disponível em: <https://educamidia.org.br/educacao-midiatica>. Acesso em: 8 maio 2025.

EL PAÍS. **Brasil contra X: un precedente democrático.** El País, Opinião, 14 out. 2024. Disponível em: <https://elpais.com/opinion/2024-10-14/brasil-contra-x-un-precedente-democratico.html>. Acesso em: 9 maio 2025.

ESTADOS UNIDOS, Congresso. **Declaração da Independência.** 4 de jul. de 1776. Disponível em: <https://www.archives.gov/founding-docs/declaration>. Acesso em 14 agosto 2025.

EUBANKS, Virginia. **Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.** New York: St. Martin's Press, 2018.

EUROPEAN COMMISSION. **Digital Services Act: Commission welcomes political agreement on rules ensuring a safe and accountable online environment.** Brussels, 23 Apr. 2022. Disponível em: https://ec.europa.eu/commission/presscorner/detail/en/IP_22_2545. Acesso em: 6 maio 2025.

EUROPEAN COMMISSION. **Digital Services Act: Questions and Answers.** Brussels, 2022. Disponível em: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348. Acesso em: 6 maio 2025.

FOUCAULT, Michel. **A coragem da verdade: o governo de si e dos outros II.** Curso no Collège de France (1983-1984). Tradução de Eduardo Brandão. São Paulo: Martins Fontes, 2006.

FRANÇA. **Déclaration des Droits de l'Homme et du Citoyen.** 1789. Disponível em: <https://www.conseil-constitutionnel.fr/le-bloc-de-constitutionnalite/declaration-des-droits-de-l-homme-et-du-citoyen-de-1789>. Acesso em: 4 maio 2025.

FRASER, Nancy. **Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy.** Social Text, n. 25/26, p. 56-80, 1990.

G1. **Xenofobia contra nordestinos na época da eleição fez número de denúncias disparar na internet, mostra pesquisa.** G1, Tecnologia, 8 fev. 2023. Disponível em: <https://g1.globo.com/tecnologia/noticia/2023/02/08/xenofobia-contra-nordestinos-na-epoca-da-eleicao-fez-numero-de-denuncias-disparar-na-internet-mostra-pesquisa.ghtml>. Acesso em: 6 maio 2025.

GERMANY. **Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act - NetzDG).** 1 September 2017. Disponível em: https://www.bmj.de/SharedDocs/Downloads/DE/Gesetzgebung/RefE/NetzDG_engl.pdf?__blob=publicationFile&. Acesso em: 10 maio 2025.

GILLESPIE, Tarleton. **Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media**. New Haven: Yale University Press, 2018.

GOMES, Fabrício Vasconcelos; SALVADOR, João Pedro Favaretto; LUCCAS, Victor Nóbrega. **Discurso de ódio: desafios jurídicos**. São Paulo: Atlas, 2020.

GOMES, Wilson. **A praga da desinformação política no Brasil: causas, consequências e antídotos**. Revista USP, São Paulo, n. 125, p. 13-26, abr./jun. 2020.

GOOGLE. **Como o PL 2630 pode piorar a sua internet**. Google Blog, 27 abr. 2023. Disponível em: <https://blog.google/intl/pt-br/novidades/iniciativas/como-o-pl-2630-pode-piorar-a-sua-internet/>. Acesso em: 2 maio 2025.

GRECO, Luis. **A função pública da comunicação digital e os limites da neutralidade de rede**. In: LEITE, George Salomão; LEMOS, Ronaldo (Orgs.). *Direito Digital: Debates Contemporâneos*. 2. ed. São Paulo: Thomson Reuters Brasil, 2019.

HABERMAS, Jürgen. **Mudança estrutural da esfera pública: investigações quanto a uma categoria da sociedade burguesa**. Tradução de Flávio Beno Siebeneichler. Rio de Janeiro: Tempo Brasileiro, 2003.

INGLATERRA. **Declaração de Direitos (Bill of Rights, 1689)**. In: *Internet History Sourcebooks Project. Modern History Sourcebook*. Fordham University, New York. Disponível em: <http://www.fordham.edu/halsall/mod/1689billofrights.asp>. Acesso em: 14 agosto 2025.

INGLATERRA. **Magna Carta Libertatum, 1215**. Disponível em: <https://www.archives.gov/exhibits/featured-documents/magna-carta/translation.html>. Acesso em: 18 ago. 2025.

JUSBRASIL. **STF determina suspensão do X, antigo Twitter, no Brasil**. Jusbrasil Notícias, 21 ago. 2024. Disponível em: <https://www.jusbrasil.com.br/artigos/stf-determina-suspensao-total-do-x->

antigo-twitter-no-brasil-entenda-a-decisao-que-impoe-bloqueio-multas-e-responsabiliza-grupo-economico/2701717701. Acesso em: 3 maio 2025.

KANT, Immanuel. **Fundamentação da Metafísica dos Costumes**. Tradução de Paulo Quintela. Lisboa: Edições 70, 1986.

KLONICK, Kate. **The New Governors: The People, Rules, and Processes Governing Online Speech**. Harvard Law Review, v. 131, n. 6, p. 1598-1670, Apr. 2018.

LEMONS, Taylor. **Regulating Hate Speech in the Digital Age: A Comparative Analysis of the United States and Germany**. Washington University Global Studies Law Review, v. 18, n. 2, p. 435-465, 2019.

LEMOS, Ronaldo; MARQUES, Marcela. **Marco Civil da Internet: Construção e Aplicação**. São Paulo: Almedina Brasil, 2021.

LEVITSKY, Steven; ZIBLATT, Daniel. **Como as democracias morrem**. Rio de Janeiro: Zahar, 2018.

LOCKE, John. **Carta sobre a tolerância**. Tradução de Luiz João Baraúna. São Paulo: Martins Fontes, 2000.

MAFEI, Rafael. **Liberdade de expressão: as fronteiras entre o político e o jurídico**. São Paulo: Saraiva Educação, 2016.

MAZZUOLI, Valerio de Oliveira. **Curso de Direito Internacional Público**. 2. ed. São Paulo: Forense, 2019.

MAZZUOLI, Valerio de Oliveira. **O controle jurisdicional da convencionalidade das leis**. 6. ed. São Paulo: Thomson Reuters Brasil, 2021.

MBEMBE, Achille. **Necropolítica**. Tradução de Renata Santini. 3. ed. São Paulo: N-1 Edições, 2018.

MENDES, Gilmar Ferreira; BRANCO, Paulo Gustavo Gonet. **Curso de Direito Constitucional**. 15. ed. São Paulo: Saraiva Educação, 2020.

MENDES, Rafael Pereira da Silva. **"Homofobia"**; Brasil Escola, 2025.

Disponível em: <https://brasilecola.uol.com.br/psicologia/homofobia.htm>. Acesso em 18 de agosto de 2025.

MENDES, Rafael Pereira da Silva. "**Liberdade de expressão**"; Brasil Escola, 2025. Disponível em: <https://brasilecola.uol.com.br/sociologia/liberdade-de-expressao.htm>. Acesso em 18 de agosto de 2025.

MEYER-PFLUG, Samantha Ribeiro. **Liberdade de expressão e discurso de ódio**. São Paulo: Saraiva Educação, 2009.

MILL, John Stuart. **Sobre a liberdade**. Tradução de Maria Lúcia de Oliveira. São Paulo: Martins Fontes, 2000.

MILTON, John. **Areopagítica: um discurso pela liberdade de imprensa sem licença de aprovação**. Tradução de Célia Berrettini. São Paulo: Martins Fontes, 2003.

MOREIRA, Adilson. **Tratado de direito antidiscriminatório**. São Paulo: Contracorrente, 2020.

NAÇÕES UNIDAS. Comitê para a Eliminação da Discriminação Racial (CERD). **Recomendação Geral nº 35: Combate ao discurso de ódio racista**. 26 set. 2013. (CERD/C/GC/35).

NAGLE, Angela. **Kill all normies: Online culture wars from 4chan and Tumblr to Trump and the alt-right**. Winchester: Zero Books, 2017.

NAHAS, Isabela. **Meta permitirá discurso de ódio para aumentar lucros, alertam especialistas da USP**. Jornal da USP, 10 jan. 2025. Disponível em: <https://jornal.usp.br/diversidade/meta-permitira-discurso-de-odio-para-aumentar-lucros-alertam-especialistas-da-usp/>. Acesso em 3 maio 2025.

NOBLE, Safiya Umoja. **Algorithms of Oppression: How Search Engines Reinforce Racism**. New York: New York University Press, 2018.

NUSSBAUM, Martha C. **Hiding from humanity: disgust, shame, and**

the law. Princeton: Princeton University Press, 2004.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU). Alto Comissariado das Nações Unidas para os Direitos Humanos (ACNUDH). **Programa Mundial de Educação em Direitos Humanos – Segunda Fase (2010-2014).** Plano de Ação. Genebra: ONU, 2012.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU). **Convenção Internacional sobre a Eliminação de Todas as Formas de Discriminação Racial.** Nova Iorque, 21 de dezembro de 1965. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000139390>. Acesso 10 agosto 2025.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU). **Convenção sobre a Eliminação de Todas as Formas de Discriminação contra a Mulher.** Nova Iorque, 18 de dezembro de 1979. Disponível em: https://assets-compromissoeatitude-igp.sfo2.digitaloceanspaces.com/2012/11/SPM2006_CEDAW_portugues.pdf. Acesso 10 agosto 2025.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU). **Convenção sobre os Direitos das Pessoas com Deficiência.** 2006. Disponível em: https://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=424-cartilha-c&category_slug=documentos-pdf&Itemid=30192. Acesso em: 08 agosto 2025.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU). **Declaração Universal dos Direitos Humanos.** Paris, 1948. Disponível em: <https://www.un.org/en/about-us/universal-declaration-of-human-rights>. Acesso em: 4 maio 2025.

ORGANIZAÇÃO DOS ESTADOS AMERICANOS (OEA). **Convenção Americana sobre Direitos Humanos (Pacto de San José da Costa Rica).** San José, 1969. Disponível em: https://www.cidh.oas.org/basicos/portugues/c.convencao_americana.htm. Acesso em: 2 maio 2025.

ORGANIZAÇÃO DOS ESTADOS AMERICANOS (OEA). **Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas**

Correlatas de Intolerância. La Antigua, Guatemala, 5 jun. 2013. Disponível em: <https://www.oas.org/pt/cidh/mandato/basicos/discriminacioneintolerancia.pdf>. Acesso em: 6 maio 2025.

PARISER, Eli. **O Filtro Invisível: O que a internet está escondendo de você.** Tradução de Diego Alfaro. Rio de Janeiro: Zahar, 2012.

PIOVESAN, Flávia. **Direitos humanos e o direito constitucional internacional.** 20. ed. São Paulo: Saraiva Educação, 2021.

PASQUALE, Frank. **The Black Box Society: The Secret Algorithms That Control Money and Information.** Cambridge, MA: Harvard University Press, 2015.

POST, Robert. Participatory Democracy and Free Speech. *In*: BREYER, Stephen (Ed.). **Constitutionalism and Democracy.** New York: New York University Press, 2011. p. 68-92.

POST, Robert. **The Classic First Amendment Tradition Under Stress: Freedom of Speech and the University.** *In*: POST, Robert C. (Ed.). **Democracy, Expertise, and Academic Freedom: A First Amendment Jurisprudence for the Modern State.** New Haven: Yale University Press, 2012. p. 20-57.

RECUERO, Raquel; SOARES, Felipe. **O Discurso Desinformativo sobre a Cura do COVID-19 no Twitter: Estudo de caso.** *E-Compós*, v. 24, 2021.

RIBEIRO, Manoel Horta et al. **Auditing radicalization pathways on YouTube.** *In*: **Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency.** New York: ACM, 2019. p. 131-141. Disponível em: <https://dl.acm.org/doi/10.1145/3351095.3372879>. Acesso em: 4 maio 2025.

SAFERNET BRASIL. **Denúncias de crimes de discurso de ódio e de imagens de abuso sexual infantil na internet têm crescimento em 2022.** Salvador: SaferNet Brasil, 2023. (Disponível em <https://new.safernet.org.br/content/denuncias-de-crimes-de-discurso-de->

odio-e-de-imagens-de-abuso-sexual-infantil-na-internet#:~:text=Den%C3%BAncias%20de%20neonazismo%20%C3%A0%20Safernet,9004%20den%C3%BAncias%20registradas%20em%202020. Acesso em: 12 maio 2025

SAFERNET BRASIL. **Quem somos**. [S.l.: s.n., s.d.]. Disponível em: <https://new.safernet.org.br/content/quem-somos>. Acesso em: 12 maio 2025.

SANTOS, Boaventura de Sousa. **A crítica da razão indolente: contra o desperdício da experiência**. Para um novo senso comum: a ciência, o direito e a política na transição paradigmática. 3. ed. São Paulo: Cortez, 2009. (Coleção Para um Novo Senso Comum; v. 1).

SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais: uma teoria geral dos direitos fundamentais na perspectiva constitucional**. 14. ed. Porto Alegre: Livraria do Advogado Editora, 2022.

SARLET, Ingo Wolfgang. **Dignidade da pessoa humana e direitos fundamentais na Constituição Federal de 1988**. 14. ed. Porto Alegre: Livraria do Advogado, 2021.

SARLET, Ingo Wolfgang; MARINONI, Luiz Guilherme; MITIDIERO, Daniel. **Curso de Direito Constitucional**. 10. ed. São Paulo: Saraiva Educação, 2021.

SARLET, Ingo Wolfgang; NETO, Cláudio Pereira de Souza. **Discurso de ódio e liberdade de expressão nas democracias constitucionais: uma análise a partir da jurisprudência do STF**. Revista de Investigações Constitucionais, Curitiba, v. 8, n. 1, p. 219-247, jan./abr. 2021.

SARMENTO, Daniel. **Dignidade da Pessoa Humana: Conteúdo, Trajetórias e Metodologia**. Belo Horizonte: Fórum, 2016.

SILVA, Daniel Neves. "**Xenofobia**"; Brasil Escola, 2024. Disponível em: <https://brasilecola.uol.com.br/sociologia/xenofobia.htm>. Acesso em 18 de agosto de 2025.

SILVA, David T. G. da. **Dogolachan e o extremismo de direita online**

no Brasil. Revista Sociedade e Estado, Brasília, v. 35, n. 1, p. 267-293, jan./abr. 2020.

SILVA, Virgílio Afonso da. **O proporcional e o razoável.** Revista dos Tribunais, São Paulo, v. 798, p. 23-50, abr. 2002. Republicado em: SILVA, Virgílio Afonso da. **Direitos fundamentais: conteúdo essencial, restrições e eficácia.** 2. ed. São Paulo: Malheiros, 2010.

SUNSTEIN, Cass R. **#Republic: Divided Democracy in the Age of Social Media.** Princeton: Princeton University Press, 2017.

TERRA. **O que se sabe sobre ameaça de bomba no show da Lady Gaga no Rio.** Terra, São Paulo, 5 maio 2025. Disponível em: <https://www.terra.com.br/noticias/brasil/o-que-se-sabe-sobre-ameaca-de-bomba-no-show-da-lady-gaga-no-rio,759170beeb2e5dc1a8bee319f1bbd836qpltn9o6.html> Acesso em: 5 maio 2025.

TERRA. **STF forma maioria para responsabilizar redes sociais por conteúdos publicados por usuários.** Terra, São Paulo, 11 junho 2025. Disponível em: <https://www.terra.com.br/byte/stf-forma-maioria-para-responsabilizar-redes-sociais-por-conteudos-publicados-por-usuarios,63ac3b0ae6d9b292cf61923c66bcd90d67fubi3f.html> Acesso em: 10 agosto 2025.

TRIBUNAL EUROPEU DE DIREITOS HUMANOS (TEDH). **Case of Erbakan v. Turkey.** Application no. 59405/00. Judgment of 6 July 2006. Disponível em: <https://hudoc.echr.coe.int/eng#%7B%22itemid%22:%5B%22001-76232%22%5D%7D>. Acesso em: 4 maio 2025.

TRIBUNAL EUROPEU DE DIREITOS HUMANOS (TEDH). **Case of Handyside v. The United Kingdom.** Application no. 5493/72. Judgment of 7 December 1976. Disponível em: <https://hudoc.echr.coe.int/eng?i=001-57499>. Acesso em: 4 maio 2025.

TRIBUNAL EUROPEU DE DIREITOS HUMANOS (TEDH). **Case of Garaudy v. France.** Application no. 65831/01. Decision of 24 June 2003. Disponível em:

<https://hudoc.echr.coe.int/eng#%7B%22itemid%22:%5B%22001-23829%22%5D%7D>. Acesso em: 4 maio 2025.

TRIBUNAL EUROPEU DE DIREITOS HUMANOS (TEDH). **Case of Jersild v. Denmark**. Application no. 15890/89. Judgment of 23 September 1994. Disponível em: <https://hudoc.echr.coe.int/eng?i=001-57891>. Acesso em: 4 maio 2025.

UNESCO. **Educação para a cidadania global: preparando alunos para os desafios do século XXI**. Brasília: UNESCO, 2021. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000234311>. Acesso em: 13 maio 2025.

UNIÃO EUROPEIA. **Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022, relativo a um mercado único para os serviços digitais e que altera a Diretiva 2000/31/CE (Regulamento dos Serviços Digitais)**. Jornal Oficial da União Europeia, L 277, 27 out. 2022. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:32022R2065>. Acesso em: 4 maio 2025.

UNITED NATIONS. **UN Strategy and Plan of Action on Hate Speech**. 18 June 2019. Disponível em: <https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20Plan%20of%20Action%20on%20Hate%20Speech%2018%20June%20SYNOPSIS.pdf>. Acesso em: 4 maio 2025.

UNITED KINGDOM. **Online Safety Act 2023**. c. 50. Disponível em: <https://www.legislation.gov.uk/ukpga/2023/50/contents/enacted>. Acesso em: 1 maio 2025.

VAN DIJCK, José; POELL, Thomas; DE WAAL, Martijn. **The Platform Society: Public Values in a Connective World**. New York: Oxford University Press, 2018.

VAN DIJK, Pieter; VAN HOOFF, Fried; VAN RIJN, Arjen; ZWAAK, Leo (Eds.). **Theory and Practice of the European Convention on Human Rights**. 5th ed. Cambridge: Intersentia, 2018.

VEJA. **Plano de ataque a bomba em show da Lady Gaga visava matar crianças e LGBTs.** Veja, São Paulo, 6 maio 2025. Disponível em: <https://veja.abril.com.br/coluna/o-som-e-a-furia/plano-de-ataque-a-bomba-em-show-da-lady-gaga-visava-matar-criancas-e-lgbts/>. Acesso em: 15 maio 2025.

VENTURINI, Jamila. **Plataformas digitais e regulação democrática: os desafios da responsabilização por conteúdo ilícito.** In: CANABARRO, Diego; DONEDA, Danilo; JUNIOR, Walter B. (Orgs.). Regulação de plataformas digitais: desafios e caminhos. São Paulo: Comitê Gestor da Internet no Brasil - CGI.br, 2020. p. 85-106.

WALDRON, Jeremy. **The Harm in Hate Speech.** Cambridge, MA: Harvard University Press, 2012.

WARDLE, Claire; DERAKHSHAN, Hossein. **Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking.** Strasbourg: Council of Europe, 2017. Disponível em: <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c>. Acesso em: 20 maio 2025.

WU, Tim. **The Attention Merchants: The Epic Scramble to Get Inside Our Heads.** New York: Alfred A. Knopf, 2017.

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

ÍNDICE REMISSIVO

A

Adaptação, 31

Ampliação, 16

B

Bibliográfico, 157

Bibliográficos, 158

C

Cidadãos, 18

Civilizações, 26

Códigos, 114

Comunicação, 17

Comunidade, 28

Constituição, 17

Controversas, 38

Convenção, 124

Cultura, 22

D

Deficiência, 36, 121

Democrático, 19, 119

Democratização, 16

Desumaniza, 125

Didática, 29

Digitais, 11

Dignidade, 16

Dinâmicas, 16

Dinamicidade, 19

Direitos, 11

Discurso, 11

Ditatorial, 16

Divisão, 117

Documento, 35

E

Ecosistema, 114

Efetivação, 36

Emblemáticos, 19

Engajamento, 17

Enunciação, 38

Excelência, 126

Exponenciados, 115

Expressão, 11

F

Fenômeno, 16

Fraternidade, 30

Fundamentais, 11, 125

G

Gênero, 36

Genética, 31

Globalização, 31

H

Hermenêutica, 156

Histórica, 25

Hostilidade, 36

Humana, 17

I

Igualdade, 124

Imperativo, 34

Inalienáveis, 27

Incompatível, 33

Indivíduos, 17

Inegociáveis, 37

Insidiosa, 127

Interdependência, 25

Intrínseco, 25

J

Judiciais, 115

Jurídico, 119

L

Liberdade, 17

Liberdade, 11

M

Massiva, 16

Mecanismos, 28

Migrantes, 122

Multifacetado, 119

Mundiais, 30

N

Narrativas, 128

Natureza, 17

O

Odiosos, 117

Online, 115

P

Pejorativo, 128

Pilares, 25

Premissas, 154

Prerrogativas, 25

Procedimentos, 155

Progresso, 121

Q

Qualitativa, 159

R

Regulação, 22

Regulação, 11

Regulatórios, 155

Relevância, 159

Religião, 25

Repressão, 123

Revolução, 26

S

Sistemática, 128

Sociedade, 17

Sociedades, 25

Solidária, 17

U

Universais, 27

Universalidade, 25

V

Vulnerabilizados, 116

PROTEÇÃO AOS DIREITOS HUMANOS FRENTE AO DISCURSO DE ÓDIO NAS REDES SOCIAIS E À LIBERDADE DE EXPRESSÃO

Revista REASE chancelada pela Editora Arché.

São Paulo- SP.

Telefone: +55(11) 5107- 0941

<https://periodicorease.pro.br>

contato@periodicorease.pro.br

**PROTEÇÃO AOS DIREITOS HUMANOS FRENTE AO DISCURSO
DE ÓDIO NAS REDES SOCIAIS E À LIBERDADE DE
EXPRESSÃO**

IBR



9786560542419