

## LIMITES ENTRE A LIBERDADE DE EXPRESSÃO E O DISCURSO DE ÓDIO NAS REDES SOCIAIS À LUZ DO DIREITO PENAL

Willian Ribeiro da Silva<sup>1</sup>  
Vinícius Leite Ferreira<sup>2</sup>

**RESUMO:** O presente trabalho analisa os limites entre a liberdade de expressão e o discurso de ódio nas redes sociais à luz do Direito Penal brasileiro em comparação com o Direito Internacional. Em um contexto social marcado pelo aumento significativo de crimes de ódio na internet, que ultrapassaram 87 mil casos no Brasil apenas em 2025, busca-se compreender até que ponto o Estado pode atuar de forma repressiva sem violar garantias constitucionais fundamentais. A pesquisa parte de revisão bibliográfica e jurisprudencial, examinando marcos relevantes como a ADPF 187 e o Tema 533 do Supremo Tribunal Federal, bem como os axiomas do Garantismo Penal. O estudo discute o papel e a responsabilidade das plataformas digitais na prevenção e remoção de conteúdos ilícitos, especialmente diante do crescimento de discursos extremistas impulsionados pelo anonimato e pela rápida circulação de informações. Pretende-se, assim, contribuir para o debate acerca da necessidade de uma atuação estatal equilibrada, através da tipicidade estrita que seja capaz de proteger o ambiente democrático e os direitos fundamentais sem restringir indevidamente a liberdade de expressão.

1

**Palavras-Chave:** Liberdade de expressão. Discurso de ódio. Redes sociais. Direito Penal. Responsabilidade das plataformas.

**ABSTRACT:** This study analyzes the boundaries between freedom of expression and hate speech on social media under Brazilian Criminal Law in comparison with International Law. In a social context marked by a significant increase in online hate crimes, which exceeded 87,000 cases in Brazil in 2025 alone, it seeks to understand to what extent the State can act repressively without violating fundamental constitutional guarantees. The research is based on a bibliographical and jurisprudential review, examining relevant milestones such as ADPF 187 and Theme 533 of the Brazilian Supreme Court, as well as the axioms of Penal Guaranteeism. The study discusses the role and responsibility of digital platforms in preventing and removing illicit content, especially in light of the growth of extremist discourse driven by anonymity and the rapid circulation of information. Thus, it intends to contribute to the debate concerning the need for balanced State action, through strict typicality capable of protecting the democratic environment and fundamental rights without unduly restricting freedom of expression.

**Keywords:** Freedom of expression. Hate speech. Social media. Criminal law. Platform liability.

<sup>1</sup> Estudante de Direito da Faculdade Santo Antônio, São José dos Campos, SP.

<sup>2</sup> Professor Orientador. Mestre em Direito - UNIFIEO/SP. Pós-graduado em Direito Penal e Processual Penal - FAAP e Damásio de Jesus. Professor em pós-graduação na FAAP. Faculdade Santo Antônio, São José dos Campos, SP.

## I. INTRODUÇÃO

A sociedade contemporânea encontra-se diante de uma linha tênue que coloca, de um lado, o direito fundamental à liberdade de expressão e, de outro, o avanço progressivo dos discursos de ódio e da intolerância, condutas que afetam diretamente o princípio da dignidade da pessoa humana. Esse panorama revela que o ambiente digital, outrora idealizado como um espaço de democratização da fala, experimenta um crescimento expressivo de manifestações hostis, transformando-se frequentemente em um terreno fértil para a intolerância e para a violência virtual.

O crescente número de manifestações de ódio e discursos radicais nas redes sociais desperta a preocupação de que a reiteração dessas manifestações abra precedentes para a formação de uma estrutura social intolerante. Essa ausência de sensibilidade nos remete a teoria da banalidade do mal, de Hannah Arendt, segundo a qual a ausência de reflexão ética pode conduzir indivíduos comuns à prática de atos de extrema crueldade. No ambiente virtual, a arquitetura das plataformas impulsionada por algoritmos que promovem a polarização e a criação de bolhas de radicalização, aliado ao chamado “efeito da desinibição tóxica”, potencializa e legitima essas agressões à dignidade humana.

A relevância jurídica da discussão também se manifesta no recente julgamento da ADPF 187, no qual o Supremo Tribunal Federal consolidou o entendimento de que o direito à liberdade de expressão deve ser amplamente protegido em um Estado Democrático de Direito, ainda que as manifestações possam contrariar a opinião pública ou defender ideias controversas. Essa decisão estabeleceu parâmetros jurídicos para distinguir a livre manifestação do pensamento, do discurso de ódio, permitindo compreender até que ponto o Direito Penal pode atuar sem ferir os princípios constitucionais da dignidade da pessoa humana e da liberdade de expressão.

Além disso, o debate ganhou novos contornos com a análise da constitucionalidade do art. 19 da lei nº 12.965/2014 (Marco Civil da internet), em que o julgamento do Tema 533 pelo Supremo Tribunal Federal discutiu a responsabilidade das plataformas digitais na manutenção de conteúdos ilícitos na internet. Ao analisar o Recurso Extraordinário 1.037.396, o STF entendeu que, embora a liberdade de expressão seja fundamento essencial do Estado Democrático de Direito, ela não pode servir de escudo para discursos evidentemente ilícitos, como manifestações de ódio, racismo, incitação à violência ou ataques à integridade das instituições democráticas. Esse precedente contribui de maneira significativa para o presente

estudo, pois reforça a necessidade de redefinir os limites jurídicos entre o exercício legítimo da liberdade de expressão e a repressão estatal às condutas que violam direitos fundamentais, especialmente no ambiente virtual, em que o anonimato e a rápida circulação de informações intensificam os efeitos do discurso de ódio.

Diante da complexidade em se rastrear esse limite no cenário nacional, faz-se indispensável recorrer ao Direito Comparado, o qual evidencia que o enfrentamento ao *hate speech* se divide em duas grandes vertentes mundiais. De um lado, encontra-se o modelo norte-americano, fortemente respaldado pela jurisprudência da Primeira Emenda, que adota uma postura maximalista da liberdade de expressão ao tolerar discursos profundamente ofensivos, desde que não representem um perigo claro e iminente. De outro lado, o modelo europeu, aqui personificado pelo direito alemão, fundamenta-se no conceito de "democracia militante", uma construção histórica do pós-guerra voltada a proteger a dignidade humana como valor supremo e absoluto, o que leva a sua legislação a criminalizar severamente o incitamento ao ódio público e a impor obrigações mais rígidas de moderação às redes sociais.

Assim, a pesquisa utiliza método dedutivo, com abordagem qualitativa, mediante revisão bibliográfica, análise jurisprudencial e estudo legislativo. O objetivo central é analisar, sob a ótica do Garantismo Penal e do princípio da intervenção mínima, até que ponto o discurso de ódio pode ser reprimido no Direito Penal brasileiro sem que se viole o direito constitucional à liberdade de expressão. Para tanto, o trabalho examina os fatores da desinibição tóxica e da polarização algorítmica no ambiente virtual, confronta as respostas jurisprudenciais do STF na ADPF 187 e no Tema 533 (RE 1.037.396), e propõe uma articulação entre a estrita legalidade penal e a governança digital responsável.

3

## 2 O DIREITO FUNDAMENTAL À LIBERDADE DE EXPRESSÃO

### 2.1 Conceito e evolução

O que diferencia o ser humano no reino animal é a capacidade de comunicação, adaptação e convivência social. Para viabilizar essa vida em coletividade tornou-se fundamental o estabelecimento de um conjunto de regras e princípios capaz de harmonizar a natureza humana com as exigências da vida em sociedade. Tal ordenamento visa disciplinar a conduta individual sem, contudo, acabar com a esfera de liberdade inerente a cada sujeito, garantindo o equilíbrio entre o arbítrio pessoal e a harmonia coletiva.

Nesse sentido, a liberdade de expressão emerge como direito fundamental de primeira dimensão, constituindo a prerrogativa de todo cidadão manifestar, exprimir e divulgar livremente seus pensamentos, ideias, convicções e críticas, sem qualquer tipo de censura prévia ou discriminações arbitrárias. Historicamente, este direito evoluiu de uma resistência às formas absolutistas de poder para se consolidar como um dos pilares do Estado Democrático de Direito.

Sob a ótica do ordenamento jurídico brasileiro, a Constituição Federal de 1988 consagrou esse direito em seu Artigo 5º, incisos IV e IX, estabelecendo a livre manifestação do pensamento (vedado o anonimato) e a liberdade de atividade intelectual, artística, científica e de comunicação, independentemente de censura ou licença. Conforme ensina José Afonso da Silva (2023), a liberdade de pensamento é o núcleo de onde irradiam as demais liberdades de expressão, sendo essencial para a autonomia individual e fiscalização do exercício do poder político.

Reforçando essa compreensão, Norberto Bobbio sustenta que a democracia não se limita a participação apenas através do voto, mas pressupõe um conjunto de “regras do jogo” que assegurem ao cidadão uma escolha consciente. Entre tais condições, a liberdade de opinião e de imprensa assume papel central, uma vez que, sem o acesso a fontes diversificadas de informação e sem a prerrogativa de contestar o poder, o consenso que sustenta o regime democrático se torna meramente artificial, resultando em uma vontade popular manipulada.

Dessa forma, a evolução do conceito de liberdade de expressão revela que ela não se esgota na proteção do indivíduo, mas cumpre uma função social e democrática essencial. Como aponta a doutrina de Ingo Wolfgang Sarlet (2021), os direitos fundamentais, ao garantirem a livre circulação de ideias, permitem o pluralismo político e a participação ativa dos cidadãos na esfera pública, elementos sem os quais a democracia torna-se puramente formal e desprovida de substância.

4

## 2.2 A Liberdade de Expressão como Pilar da Democracia

Ao assegurar o exercício dos direitos sociais e individuais logo em seu preâmbulo e fundamentos iniciais, a Constituição Federal de 1988 estabelece as bases de um Estado Democrático de Direito no qual a liberdade de expressão se consolida como garantia fundamental indispensável, assegurando ao cidadão o espaço necessário para a livre manifestação de seus pensamentos e convicções. Embora a decorrente pluralidade de ideias desperte conflitos naturais, o debate não deve ser evitado, visto que é justamente por meio da

constante exposição dos argumentos ao contraditório que as posições consistentes se solidificam, ao passo que as formulações frágeis se dissolvem.

Nessa perspectiva, John Stuart Mill, em sua obra “sobre a liberdade”, argumenta que o silenciamento de qualquer opinião é um “mal peculiar”, pois rouba da humanidade a oportunidade de trocar o erro pela verdade ou, no caso de uma opinião parcialmente correta, de refinar o entendimento através do conflito de percepções. Segundo o autor, se toda a humanidade, menos uma pessoa, tivesse uma opinião, e apenas essa pessoa tivesse a opinião contrária, a humanidade não teria mais direito de silenciar essa pessoa do que ela teria de silenciar a humanidade. (MILL, 2000, p. 33).

A liberdade de expressão cumpre, portanto, um papel pedagógico na democracia ao superar os dogmas que por vezes são repetidos sem a compreensão real de seus fundamentos, além de permitir ao indivíduo enquanto portador de direitos e deveres o exercício de sua autonomia individual e construção crítica.

No entanto, o direito à liberdade de expressão não possui caráter absoluto. A própria Constituição em seu art. 5º, IV, ao garantir a livre manifestação do pensamento, veda expressamente o anonimato. De igual modo, o art. 220 garante que a manifestação do pensamento não sofrerá restrição sob qualquer forma, desde que observado o disposto na própria constituição. Assim, o exercício desse direito deve coexistir harmonicamente com outros pilares, como a dignidade da pessoa humana e o pluralismo político.

5

Nesses termos, para a doutrina dominante, falar em direito de expressão ou de pensamento não é falar em direito absoluto de dizer tudo aquilo ou fazer tudo aquilo que se quer. De modo lógico-implícito a proteção constitucional não se estende à ação violenta. Nesse sentido, para a corrente majoritária de viés axiológico, a liberdade de manifestação é limitada por outros direitos e garantias fundamentais como a vida, a integridade física, a liberdade de locomoção. Assim sendo, embora haja liberdade de manifestação, essa não pode ser usada para manifestação que venham a desenvolver atividades ou práticas ilícitas (antisemitismo, apologia ao crime etc...) (FERNANDES, 2011, p. 279).

Dessa forma, a resolução de eventuais colisões entre a liberdade de expressão e outros direitos de mesma estatura constitucional exige a aplicação do Princípio da Ponderação, também compreendido sob a ótica da proporcionalidade, visto que, diante da inexistência de direitos fundamentais de caráter absoluto, cabe ao intérprete jurídico realizar um sopesamento no caso concreto para harmonizar os bens em conflito sem anular um em detrimento do outro. É por meio dos subelementos da adequação, necessidade e proporcionalidade em sentido estrito que essa técnica permite identificar o limite necessário para proteger a dignidade alheia ou o

interesse coletivo, protegendo o núcleo essencial da liberdade de manifestação sem que isso resulte na asfixia do debate público ou em censura prévia.

### 2.3 A Proibição da Censura e a Responsabilidade Posterior

Embora a técnica da ponderação possibilite a harmonização entre direitos fundamentais em conflito, a liberdade de expressão não possui caráter absoluto. Seu exercício encontra limites quando passa a comprometer a convivência social ou a violar direitos e garantias de terceiros. Nesse sentido, leciona Barcellos:

O exercício de um direito, porém, não pode chegar ao ponto de inviabilizar a vida em sociedade ou de violar direitos de terceiros; daí porque, em tese, não se pode descartar a possibilidade de restringir o exercício de direitos fundamentais, em primeiro lugar ao legislador e também ao juiz, quando este estiver diante de conflitos normativos insuperáveis. (BARCELLOS, 2022, p.199).

No cenário de omissão legislativa ou de lacunas em áreas cinzentas da legislação, a atuação do Poder Judiciário alcança o dever de integração do ordenamento jurídico, uma vez que, conforme disposto na Lei de Introdução às Normas do Direito Brasileiro (LINDB), o magistrado não pode se eximir de decidir sob pretexto de lacuna na lei, devendo socorrer-se da analogia, dos costumes e dos princípios gerais do direito para dar concretude à justiça. Essa intervenção torna-se indispensável para coibir condutas que, embora nem sempre estejam claramente delimitadas pelo legislador, ultrapassam os limites da livre manifestação do pensamento e ingressam no campo da ilicitude, a exemplo do que ocorre nos discursos de ódio e nos crimes contra a honra.

Todavia, a responsabilização do indivíduo que praticou a infração deve ocorrer a *posteriori*. Pois como já foi decidido pelo Supremo Tribunal Federal no julgamento da ADPF 130, qualquer cerceamento prévio à manifestação do pensamento constitui censura, prática terminantemente vedada pelo ordenamento brasileiro.

A rejeição da censura prévia não decorre apenas de uma limitação imposta ao poder estatal, mas da compreensão de que o livre fluxo de ideias constitui elemento indispensável à democracia. Quando o Estado assume o papel de tutor do debate público, determinando de antemão quais ideias são dignas de circulação, opera-se uma grave violação da autonomia individual, pois a censura atenta contra a dignidade humana ao retirar do indivíduo a capacidade de exercer sua faculdade de julgamento crítico, tratando a sociedade não como um conjunto de

cidadãos livres e conscientes, mas como uma massa incapaz de formar seus próprios juízos de valor e que necessita de proteção estatal contra os próprios pensamentos.

Além disso, a experiência histórica demonstra que mecanismos de censura tendem a ser utilizados como instrumentos de silenciamento de vozes divergentes e de grupos minoritários. Não há, na prática, um censor plenamente neutro ou imune a interesses políticos, ideológicos ou institucionais. Por essa razão, a vedação constitucional às restrições prévias atua como importante garantia contra o arbítrio, preservando o debate público e assegurando a liberdade de crítica, especialmente em relação aos agentes que possuem poder político, econômico ou social.

No contexto contemporâneo, os riscos à liberdade de expressão nem sempre se apresentam de forma explícita, manifestando-se, muitas vezes, por meio do denominado *chilling effect*, ou efeito resfriador da liberdade de expressão. Esse fenômeno ocorre quando normas excessivamente vagas ou sanções desproporcionais geram insegurança quanto aos limites do discurso permitido, o que leva os indivíduos a optarem pelo silêncio em vez da livre manifestação do seu pensamento.

Como consequência, o debate público passa a ser menos plural e perde a capacidade de cumprir sua função democrática, uma vez que a teoria do *marketplace of ideas*, pressupõe justamente o livre confronto entre diferentes pontos de vista como o meio ideal para que erros sejam contestados e verdades constantemente testadas e reafirmadas. Contudo, quando o medo decorrente desse “efeito resfriador” substitui a liberdade de participação dos usuários no ambiente virtual, o resultado prático é o empobrecimento do debate democrático, o que acaba por favorecer o conformismo social e limitar o desenvolvimento crítico da sociedade no espaço público digital.

Dessa forma, a edição de normas voltadas a coibir o discurso de ódio e a violência no ambiente virtual necessita de um amplo debate no âmbito do Poder Legislativo, cuja atuação deve se pautar na construção de definições claras e objetivas que afastem ambiguidades. Essa precisão regulatória irá cumprir o duplo papel de assegurar ao cidadão o direito à proteção contra manifestações violentas e discriminatórias na rede e, simultaneamente, garantir que ele compreenda com segurança os limites legais de sua conduta, evitando que o indivíduo se autocensure ou que seja submetido a restrições arbitrárias por parte do poder público.

### 3 O DISCURSO DE ÓDIO NAS REDES SOCIAIS

#### 3.1 Definição de *Hate Speech*

O termo *hate speech* ou discurso de ódio até o momento não possui uma definição universal no âmbito do direito internacional dos direitos humanos; contudo a ONU define o discurso de ódio da seguinte forma:

Qualquer tipo de comunicação, seja oral, escrita ou comportamental, que ataque ou utilize linguagem pejorativa ou discriminatória em referência a uma pessoa ou grupo com base em quem são, ou seja, com base em sua religião, etnia, nacionalidade, raça, cor, ascendência, gênero ou outro fator de identidade. (ONU, 2019)

Para autores como Andrew Altman, o discurso de ódio não é apenas uma expressão, mas uma conduta que trata o outro como um subordinado moral, atribuindo-lhe um valor inferior. Assim, a linguagem racista, sexista e discriminatória, além de manifestar o ódio a determinado indivíduo ou grupo, atua para subordiná-los ao tratá-los como moralmente inferiores. Nessa perspectiva, o discurso é uma ação que gera danos psicológicos severos às vítimas, resultando em quadros de ansiedade, depressão e até tendências suicidas.

É fundamental fazer a distinção do discurso de ódio que tem o *animus* discriminatório, da opinião ou crítica que tem o *animus criticandi* (intenção de criticar). Enquanto a opinião e a crítica recaem sobre ideias, comportamentos, obras ou políticas, buscando fomentar o debate público ou manifestar uma percepção subjetiva, o discurso de ódio direciona-se à identidade intrínseca do indivíduo ou grupo, excluindo e silenciando os grupos vulneráveis.

Mari Matsuda afirma que as vítimas possuem sua liberdade limitada pelos seus ofensores, pois em decorrência dos discursos de ódio elas abandonam seus estudos, deixam seus trabalhos, evitam determinados lugares públicos, etc. É evidente que o discurso de ódio ultrapassa o mero discurso, uma percepção subjetiva de uma ofensa ou “palavra mal falada”, e atinge seus efeitos de forma real e concreta, afetando diretamente a vida da vítima e sua percepção como cidadão.

As vítimas possuem a sua liberdade pessoal restringidas. Para evitar receberem mensagens de ódio, as vítimas chegam a deixar seus trabalhos, de seus estudos, deixam as suas casas, evitam determinados lugares públicos, cerceiam o seu próprio direito de expressão, e, de certa forma, modificam o seu comportamento. O destinatário de mensagens de ódio luta com a sua tormenta interna. A resposta do subconsciente é rejeitar a sua identidade como membro do grupo que é vítima. Como escritores, que retratam a experiência afro-americana, tem percebido, o preço de se dissociar da própria raça é muitas vezes a própria sanidade. (MATSUDA, Tradução nossa.)

Conclui-se que o impacto do discurso de ódio transcende a mera ofensa pessoal, configurando uma violação sistemática da dignidade e da isonomia das vítimas. Dessa forma, distinguir o debate da desumanização revela-se fundamental para a preservação do Estado Democrático de Direito, pois enquanto a opinião e a crítica fortalecem o pluralismo no mercado de ideias, o *hate speech* visa aniquilar o próprio sujeito do discurso.

### 3.2 O Fenômeno da Polarização e Anonimato

O ambiente virtual provocou uma alteração profunda na dinâmica da comunicação humana, através da sua capacidade e velocidade de disseminação de conteúdo, o acesso à fala e à informação foi democratizado. Contudo o discurso de ódio também foi amplamente difundido, pois, enquanto em períodos anteriores da história as barreiras geográficas e temporais limitavam a disseminação de preconceitos, violência e discursos de ódio, a era digital permite que uma ofensa alcance usuários do mundo inteiro em intervalos de poucos segundos.

Aliado a isso, a ausência de contato visual, em conjunto com o anonimato e a sensação de impunidade, bloqueia a empatia do usuário. Este, ao proferir mensagens discriminatórias ou discursos de ódio, não se sensibiliza ou tem noção do impacto que isso pode causar na outra pessoa, visto que, no ecossistema virtual, não é possível perceber a reação de dor e sofrimento da vítima.

Sob essa ótica, John Suler propõe que esse fenômeno ocorre por meio do conceito de “Desinibição Tóxica”, sugerindo que a persona digital difere da persona física devido a seis fatores psicológicos interligados que reduzem o controle impulsivo do indivíduo. O primeiro deles é o anonimato dissociativo, processo pelo qual o sujeito separa suas ações online de sua identidade real sob a premissa de não ser conhecido. A este soma-se a invisibilidade, em que a ausência de contato visual direto elimina o constrangimento físico imediato. Há também a assincronia, caracterizada pelo hiato temporal entre o envio da mensagem e a resposta, o que evita o impacto da reação imediata da vítima.

Ademais, o autor aponta a introjeção solipsista, em que o usuário cria uma voz imaginária para o interlocutor, esvaziando a empatia, e a imaginação dissociativa, que leva o indivíduo a enxergar a rede como um ambiente fictício ou um jogo onde as regras sociais comuns não se aplicam. Por fim, ocorre a minimização da autoridade, visto que o status social e as linhas hierárquicas tradicionais perdem o peso no ecossistema digital, enfraquecendo as

restrições comportamentais e abrindo espaço para manifestações extremas de raiva, ódio e violência.

As pessoas dizem e fazem coisas no mundo cibernético que não diriam ou fariam no mundo face a face. Elas se sentem mais livres, menos contidas. Esse fenômeno é o que chamei de efeito de desinibição online. Em sua forma tóxica, ele se manifesta através de linguagem rude, críticas severas, raiva, ódio e até ameaças. (Suler, 2004 p.321)

É evidente que essa desinibição online é prejudicial ao sentimento de coletividade, pois a ausência de empatia, fomentada pela interface das telas e pelo distanciamento físico, resulta em um esvaziamento da humanidade percebida no outro, induzindo o indivíduo a praticar atos de violência e crueldade.

Ao analisar a dinâmica das redes sociais, Cass Sustein observa que um dos principais fatores que influencia os crescentes discursos de ódio na internet é o fator da polarização e a “arquitetura da escolha”. Para ele, a democracia depende de “encontros não planejados” com ideias divergentes, pois ao nos inserirmos em bolhas na internet que filtram apenas o que gostamos e confirmam nossas crenças prévias os algoritmos destroem o tecido social e nos levam a posições muito mais radicais das que tínhamos inicialmente.

Em uma democracia, é crucial que as pessoas sejam expostas a materiais que não escolheram com antecedência. Encontros não planejados e inesperados são centrais para a democracia. [...] Quando os indivíduos se fecham em seus próprios casulos informativos, o discurso de ódio e a violência digital encontram um terreno fértil, pois o 'outro' deixa de ser um concidadão e passa a ser um inimigo. (Sustein, 2017)

Percebe-se que tanto o conceito de Suler da “desinibição online”, quanto o conceito de Sunstein de “arquitetura da escolha” são extremamente relevantes para a compreensão do mecanismos psicológicos e algoritmos que resultam no constante aumento de casos de violência na internet. Segundo levantamento da SaferNet, o Brasil enfrentou um aumento de 28,4% nas denúncias de crimes cibernéticos em 2025, totalizando 87.689 registros únicos, um aumento de 19.403 casos face a 2024.

Este cenário de polarização e crescente violência no ambiente digital, acende um alerta e nos remete ao conceito de “banalidade do mal” de Hannah Arendt, pois, se no século XX a burocracia estatal era o véu que impedia o indivíduo de enxergar a humanidade da vítima, no século XXI a interface digital cumpre papel análogo. Essa automação na entrega de conteúdo, aliada a um fluxo incessante de estímulos que apenas reforçam visões preexistentes, remove o esforço de reflexão do usuário e, sob essa lógica, transforma o internauta em um "agente

administrativo do ódio", que profere discursos violentos e endossa políticas de cancelamento com a mesma naturalidade burocrática com que Eichmann gerenciava os trens da morte.

Ao analisar o perfil de um dos responsáveis pela "Solução Final", Hannah Arendt propõe uma reflexão perturbadora: Eichmann não era um psicopata, mas um homem "terrivelmente normal". Em sua obra *Eichmann em Jerusalém*, a autora utiliza esse caso para demonstrar como a sociedade de massas desumaniza o indivíduo, criando uma multidão incapaz de exercer o julgamento crítico.

O problema com Eichmann era precisamente que muitos eram como ele, e que esses muitos não eram pervertidos nem sádicos, que eram, e ainda são, terrivelmente e assustadoramente normais. Do ponto de vista de nossas instituições legais e de nossos padrões morais de julgamento, essa normalidade era muito mais aterradora do que todas as atrocidades juntas. (Hannah Arendt, *Eichmann em Jerusalém: Um Relato sobre a Banalidade do Mal*)

Nessa seara, os discursos de ódio não decorrem de uma maldade extraordinária dos usuários, mas da superficialidade de uma massa que, imersa em bolhas de confirmação, abre mão da faculdade de julgar. A velocidade das informações, aliada à ausência de pensamento crítico desumaniza tanto o agressor quanto a vítima, ao criar uma multidão que não se sente responsável pelas consequências devastadoras de suas postagens.

A percepção de impunidade, alicerçada na premissa de que a rede é um espaço “extrajudicial” somente será dissipada quando houver a presença estatal no ambiente digital. Tal intervenção deve vir acompanhada de políticas públicas voltadas à desconstrução da polarização, fomentando uma cultura de humanização pautada na responsabilidade ética do usuário e no reconhecimento da dignidade da pessoa humana como limite intransponível às interações.

Desse modo, a presença de mecanismos estatais que combatam a impunidade em um tempo condizente com a velocidade da rede é fundamental para reduzir significativamente a proporção e a disseminação do dano. Cabe, contudo, ressaltar conforme abordado anteriormente que essa atuação estatal deve vir acompanhada de limites claros e específicos para que não seja utilizada como meio de censura.

## 4 LIMITES JURÍDICOS E A ATUAÇÃO DO STF

### 4.1 ADPF 187 e o "Campo das Ideias"

Diante dessa complexa fronteira que coloca de um lado a liberdade de expressão e de outro o dever-punir do Estado, nos deparamos com o julgamento da ADPF 187 do Supremo Tribunal Federal, em que foi julgado a legalidade e a legitimidade de eventos como a marcha da maconha. O julgamento definiu que a defesa pública de entorpecentes não constitui apologia ao crime, mas o exercício legítimo das liberdades de expressão, reunião e petição.

A liberdade de expressão é vista como um instrumento para a participação dos cidadãos da vida política do Estado e o direito de reunião serve como meio para que as minorias possam propagar as suas ideias, mesmo que sejam impopulares, chocantes ou contrárias à vontade da maioria.

Conforme colocou em seu voto, o ministro Celso de Melo dispõe que:

Decerto, inexistiria qualquer razão para que os direitos de liberdade de expressão, de reunião e de manifestação fossem alçados a tal condição caso seu âmbito normativo garantisse, exclusivamente, a exteriorização de concepções compartilhadas pela ampla maioria da sociedade ou pela política em vigor. Se para isso servissem, comporiam uma inimaginável categoria de “direitos desnecessários”; não seriam, pois, verdadeiros direitos.

12

Tal entendimento nos remete ao ideal de liberdade abordado anteriormente na teoria de John Stuart Mill, em que é fundamental garantir a liberdade de expressão justamente para que as minorias tenham voz, pois ao abafar o pensamento da minoria, além de violar seus direitos da personalidade, o Estado estaria impedindo que a sociedade evoluísse seu senso crítico e se desenvolvesse.

No caso em tela, o posicionamento do STF foi o de que a marcha da maconha, não estava fazendo apologia ao crime, mas tão somente a defesa da legalização. Ao propor uma revisão nas leis (*Abolitio criminis*), os manifestantes estão utilizando de um mecanismo legítimo do direito de crítica, de protesto e de petição. Enquanto o crime de apologia, consiste em enaltecer, elogiar ou justificar publicamente um fato criminoso.

Como a liberdade de expressão é um direito constitucional, além de ter um dever negativo de não interferir, o Estado tem a obrigação positiva de garantir que a reunião ocorra com segurança, inclusive protegendo os participantes contra grupos opositores que tentem marginalizar, tumultuar ou impedir a manifestação.

É evidente pelo entendimento do STF que a marcha da maconha tem a finalidade exclusiva de debate e crítica ao modelo proibicionista, diferente do que seria se o caráter fosse de incitação ao crime. Durante o voto o ministro Celso de Melo ressalta que a liberdade de expressão não é absoluta.

É certo que o direito à livre expressão do pensamento não se reveste de caráter absoluto, pois sofre limitações de natureza ética e de caráter jurídico. É por tal razão que a incitação ao ódio público contra qualquer pessoa, povo ou grupo social não está protegida pela cláusula constitucional que assegura a liberdade de expressão. Cabe lembrar, neste ponto, a própria Convenção Americana sobre Direitos Humanos (Pacto de São José da Costa Rica), cujo Art. 13, § 5º, exclui, do âmbito de proteção da liberdade de manifestação do pensamento, “ 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.

Portanto, o entendimento é de que a proteção constitucional abrange o debate de ideias e resguarda as opiniões minoritárias. Contudo, tal proteção não legitima a prática de crimes durante a manifestação tampouco a incitação a delitos; caso isso ocorra, a conduta ultrapassa os limites da legalidade e se configura como infração penal.

Assim, quando falamos do ambiente digital que é uma extensão do espaço público de fala, as opiniões minoritárias estão protegidas constitucionalmente desde que estas não sejam utilizadas como escudo para o abuso de direito, ou para a propagação de discursos que, sob o pretexto de opinião, atentem contra a dignidade humana ou a ordem pública.

#### 4.2 Marco Civil da internet e Tema 533 do STF

A Lei nº 12.965/2014, denominada Marco Civil da Internet, consolidou-se como o principal marco regulatório destinado a estabelecer princípios, garantias, direitos e deveres para o uso da rede no Brasil, surgindo em um contexto de crescente expansão do ambiente digital no qual a ausência de parâmetros específicos exigia uma disciplina adequada para as relações virtuais. Frequentemente reconhecida pela doutrina como a “Constituição da Internet”, a norma tem como objetivo central conferir maior segurança jurídica a esse ecossistema, assegurando a proteção da privacidade dos usuários, a neutralidade da rede e, sobretudo, a salvaguarda da liberdade de expressão.

Nesse cenário, o legislador estruturou um modelo jurídico voltado à preservação da livre circulação de ideias e informações, buscando impedir práticas de censura privada ou remoções arbitrárias de conteúdo pelas plataformas digitais. A lógica adotada pelo Marco Civil partiu da

premissa de que o ambiente virtual deveria permanecer aberto ao pluralismo de opiniões e ao exercício das liberdades comunicativas, sem intervenções indevidas por parte dos provedores.

Entretanto, a rápida evolução das redes sociais e o aumento da violência digital evidenciaram que o problema hoje vai muito além do comportamento individual do usuário; ele alcança diretamente a responsabilidade das empresas que gerenciam essas redes. No ordenamento jurídico brasileiro, o principal eixo normativo desse debate encontra-se no art. 19 da lei nº 12.965/2014, esse dispositivo foi pensado justamente para garantir o fluxo livre de opiniões, determinando que os provedores só seriam responsabilizados civilmente por publicações de terceiros se, após ordem judicial específica, deixassem de promover a indisponibilização do material apontado como ilícito, conforme expressa o texto legal:

Art. 19. Com o intuito de assegurar a liberdade de expressão e impedir a censura, o provedor de aplicações de internet somente poderá ser responsabilizado civilmente por danos decorrentes de conteúdo gerado por terceiros se, após ordem judicial específica, 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, ressalvadas as disposições legais em contrário (BRASIL, 2014).

É a partir desse ponto que o debate ganha uma nova urgência. No julgamento do Recurso Extraordinário (RE) 1.037.396, que deu origem ao Tema 533 de Repercussão Geral, o Supremo Tribunal Federal foi chamado a analisar a constitucionalidade desse artigo.

O intuito central da discussão foi definir se a exigência da remoção de conteúdos apenas mediante ordem judicial, é constitucional ou se cria um cenário de impunidade por parte das plataformas que viola os direitos da personalidade. Além disso, buscou-se conferir parâmetros legais para a remoção desse conteúdo, de tal forma que não cerceiem a liberdade de expressão e pensamento, mas que sejam eficazes para mitigar os danos causados por publicações que violem expressamente a legislação e a dignidade humana. Conforme tese fixada pelo tribunal:

O provedor de aplicações de internet será responsabilizado civilmente, nos termos do art. 21 do MCI, pelos danos decorrentes de conteúdos gerados por terceiros em casos de crime ou atos ilícitos, sem prejuízo do dever de remoção do conteúdo. Aplica-se a mesma regra nos casos de contas denunciadas como inautênticas.

Em sua decisão, o STF estabeleceu que, se um conteúdo ofensivo já foi declarado ilícito pela justiça, os provedores têm o dever de remover as cópias idênticas assim que notificados, sem a necessidade de uma nova ordem judicial para cada cópia. Contudo, a responsabilidade da empresa não é mais exclusivamente reativa (apenas após o juiz decidir), mas passa a envolver um dever de cuidado e análise de falhas sistêmicas.

Com a decisão, a plataforma passa a ser responsabilizada se deixar de adotar medidas preventivas ou de remoção adequadas ao “estado da técnica”, violando o dever de atuar de forma transparente e cautelosa para evitar falhas nos mecanismos de segurança. Essa responsabilidade, em casos que envolvam atividades econômicas, como, por exemplo, anúncios e impulsionamentos pagos, é presumida, cabendo à empresa provar que agiu com diligência e rapidez na retirada do conteúdo ilícito para afastar a condenação.

Embora o STF não tenha adotado a responsabilidade objetiva das plataformas, a Corte passou a reconhecer hipóteses de culpa presumida decorrente da omissão no dever de diligência, e da existência de falhas sistêmicas nos mecanismos de prevenção e remoção de conteúdos manifestamente ilícitos. Para que a cobrança seja possível e a resposta das plataformas seja mais efetiva, o STF fixou hipóteses específicas em que a omissão do provedor pode ensejar responsabilização civil independentemente de ordem judicial prévia.

Esse rol de exceções abrange, primeiramente, as condutas e os atos antidemocráticos, bem como os crimes de terrorismo ou seus atos preparatórios. Da mesma forma, a dispensa de provocação judicial estende-se aos conteúdos que configurem induzimento, instigação ou auxílio ao suicídio e à automutilação, além da incitação à discriminação em razão de raça, cor, etnia, religião, procedência nacional, sexualidade ou identidade de gênero. Por fim, a responsabilização imediata por omissão incide sobre a difusão de crimes sexuais contra vulneráveis, pornografia infantil e demais infrações graves contra crianças e adolescentes, englobando também o tráfico de pessoas e os crimes praticados contra a mulher em razão da condição do sexo feminino, inclusive materiais que propaguem o ódio ou a aversão às mulheres.

A fixação desses critérios objetivos pelo Supremo Tribunal Federal não serve apenas como um catálogo de proibições, mas como um pilar de segurança jurídica no ambiente digital. Ao definir categorias taxativas de conteúdos graves que não necessitam de ordem judicial para remoção, o STF reduz a margem de discricionariedade das plataformas e, simultaneamente, garante ao usuário seu exercício legítimo ao evitar que ele sofra restrição ou censura pela retirada de conteúdos que não estejam no rol taxativo.

Ademais, como a propagação de conteúdos nocivos na rede ocorre de forma exponencial, fenômeno esse que é agravado pela “desinibição tóxica” e pela “arquitetura da escolha”, aguardar o trâmite do moroso processo judicial ensejaria um dano irreversível. Desta forma, ao fixar critérios objetivos para a remoção imediata de conteúdos nas redes, o STF conferiu maior

celeridade e efetividade ao enfrentamento de crimes graves no ecossistema digital. Afinal, como nos advertiu Rui Barbosa, “Nada se parece tanto com a injustiça como a justiça tardia”. (Barbosa, 1921, p. 41).

Em última análise, a transição de um modelo puramente reativo para um sistema baseado em deveres de cuidado e análise de falhas sistêmicas confere maior transparência à sociedade, pois ao identificar quais condutas ensejam intervenção imediata, os usuários têm um ambiente mais previsível e ético.

Conclui-se, portanto, que a releitura constitucional do art. 19 do Marco Civil da Internet promovida pelo STF representa uma tentativa de compatibilizar a proteção da liberdade de expressão com a tutela da dignidade humana no ambiente digital. Ao estabelecer parâmetros objetivos para proteger a livre circulação de ideias minoritárias, bem como deveres de cuidado proporcionais ao risco da atividade econômica das plataformas, o Tribunal sinaliza uma transição para modelos regulatórios capazes de garantir um ambiente virtual efetivamente voltado à promoção da cidadania.

#### 4.3 Direito comparado

No âmbito do direito internacional, o posicionamento sobre liberdade de expressão e discurso de ódio pode variar bastante. Constata-se um contraste nítido entre os sistemas que adotam uma postura mais restritiva de tutela da dignidade humana, como o Brasil, a Alemanha e a Espanha, e a tradição estadunidense, marcada por uma proteção maximalista que limita a intervenção estatal a cenários estritamente excepcionais.

Os Estados Unidos adotaram a teoria liberal e utilitarista da liberdade de expressão por meio da doutrina do *marketplace of ideas*, ou livre mercado de ideias, partindo da premissa de que, em uma sociedade democrática, a verdade e a razão só podem ser alcançadas se a todos for assegurado o direito de expor e debater publicamente seus pontos de vista sem a interferência do Estado. Diante dessa construção filosófica e jurisprudencial, o direito americano não impõe restrições prévias à manifestação do pensamento, deixando de limitar a liberdade de expressão mesmo nos casos que envolvem o discurso de ódio.

Diferentemente do direito brasileiro, cuja base é essencialmente positivada (*civil law*), o direito nos Estados Unidos rege-se pela “*common law*”, fundamentando-se nos precedentes judiciais vinculantes. Assim, o posicionamento adotado como regra é definido pelas

interpretações da Suprema Corte sobre a Constituição de 1787, em especial a primeira emenda, cujo texto consolidou a proibição de restrições estatais ao livre fluxo de ideias e passou a funcionar como o principal parâmetro para a solução de conflitos envolvendo o discurso público.

A jurisprudência norte-americana tem dois precedentes emblemáticos que ilustram essa proteção quase absoluta à manifestação do pensamento. O primeiro é o caso *Brandenburg vs. Ohio*, no qual a Suprema Corte Americana declarou a inconstitucionalidade de uma lei estadual que criminalizava a defesa abstrata da violência como método de reforma política. Ao invalidar a condenação de um líder da Ku Klux Klan por discursos de ódio contra negros e judeus, a corte fixou a tese do “*imminent lawless action*” (ação ilegal iminente) na qual o Estado não pode restringir manifestações extremistas ou discriminatórias, salvo quando houver incitação concreta, iminente e provável à prática de atos ilegais.

O segundo marco é o caso *National Socialist Party of America vs. Village of Skokie*, no qual a Suprema Corte assegurou a um grupo neonazista o direito de marchar publicamente em uma comunidade com grande número de sobreviventes do holocausto. O tribunal entendeu que a mera exibição de símbolos e a defesa de ideias ofensivas não autorizam, por si sós, a censura ou restrição estatal prévia, constituindo, embora de forma controversa, um exercício legítimo da liberdade de expressão sob a Primeira Emenda.

Diferentemente do posicionamento dos EUA, o Brasil e países europeus como Espanha e Alemanha impõem maiores restrições à liberdade de expressão. Como exemplo dessa limitação, o artigo 130 do Código Penal Alemão (*Volksverhetzung*) criminaliza condutas capazes de perturbar a paz pública ao incitar o ódio ou sugerir medidas violentas e arbitrárias contra segmentos da população. O mesmo dispositivo pune agressões à dignidade humana manifestadas por meio de insultos ou difamações, vedando publicações de teor discriminatório baseadas em raça, nacionalidade, religião ou origem étnica, além de proibir de forma taxativa qualquer manifestação de aprovação, negação ou minimização do Holocausto.

O sistema jurídico espanhol, embora reconheça que o direito à liberdade de expressão seja a base de um estado onde predomina o pluralismo político e a tolerância, congrega do posicionamento de que este direito não é absoluto e nem ilimitado. O Tribunal Constitucional já se pronunciou diversas vezes no sentido de que ideias ou opiniões de caráter ofensivo e insultuosos, que objetivem a menosprezar ou discriminar indivíduos ou grupos por qualquer

condição ou circunstância pessoal, ética ou social, não estarão protegidas pelo direito à liberdade de expressão (ARAÚJO, 2018).

Em suma, a análise do direito comparado revela que o tratamento do discurso de ódio não é uniforme, refletindo as cicatrizes históricas e os valores político-sociais de cada nação, o que evidencia que a regulação do tema ultrapassa a mera técnica jurídica para se consolidar como uma escolha valorativa profunda. Enquanto o modelo norte-americano se alicerça na Primeira Emenda para garantir que o Estado não atue como árbitro da verdade, os sistemas europeu e brasileiro adotam uma postura mais restritiva, estabelecendo que a manifestação do pensamento não pode ser instrumentalizada para destruir o próprio regime democrático ou violar a dignidade alheia, em uma clara divisão entre a primazia da liberdade individual e a preservação da paz social e dignidade humana das minorias.

## 5 O DIREITO PENAL COMO *ULTIMA RATIO*

### 5.1 Tipicidade e discurso de ódio

O ordenamento jurídico brasileiro não prevê o discurso de ódio como tipo penal autônomo, razão pela qual sua repressão deve ocorrer por meio da subsunção a tipos penais já existentes, em observância ao princípio da legalidade estrita, previsto no art. 5º, XXXIX, da Constituição Federal e no art. 1º do Código Penal.

Nesse cenário, a principal via de tipificação do discurso de ódio no Brasil é através da Lei nº 7.716/89, que pune crimes resultantes de preconceito de raça ou cor. Com o advento da Lei nº 14.532/2023, essa proteção foi ampliada ao equiparar o crime de injúria racial ao crime de racismo, não apenas majorando as penas cominadas, mas conferindo ao delito a imprescritibilidade e inafiançabilidade.

Importante destacar que tanto o racismo como a injúria racial podem representar manifestações de discurso de ódio, dependendo do contexto e da finalidade discriminatória da conduta. Enquanto o racismo possui natureza coletiva ou difusa atingindo determinado grupo ou comunidade, a injúria racial incide sobre a honra subjetiva de uma pessoa específica.

Como afirma a doutrina, a prática do discurso de ódio, com seu caráter externo de incitar a discriminação, a hostilidade ou a violência, configura, na maioria das vezes, infração penal, como o racismo (previsto na Lei n. 7.716/89). Essa conduta poderá também produzir consequências civis em razão da lesão à honra, à intimidade ou à imagem das pessoas, devendo sua repressão ser feita com cautela. Conforme leciona Rogério Greco, os crimes contra a honra

previstos nos arts. 138, 139 e 140 do Código Penal visam proteger a honra subjetiva e objetiva do indivíduo, ao passo que o discurso de ódio, em regra, atinge a dignidade de grupos historicamente vulneráveis. Dessa forma, a simples tipificação como crime contra a honra pode ser insuficiente para incorporar a gravidade estrutural dessas condutas.

A relevância desse enquadramento foi consolidada pelo Supremo Tribunal Federal no julgamento do HC 82.424/RS (caso Ellwanger), no qual se firmou o entendimento de que manifestações de conteúdo antissemita configuram crime de racismo, ainda que disfarçadas sob a forma de opinião ou manifestação intelectual. Na ocasião, a Corte firmou o entendimento de que a liberdade de expressão não possui caráter absoluto, e não legitima práticas discriminatórias que violem princípios da dignidade da pessoa humana (art. 1º, III, da CF) e da igualdade (art. 5º, caput, da CF), fixando a seguinte tese: “a prática de racismo abrange a discriminação contra os judeus e que a liberdade de expressão não protege o discurso de ódio.”

Para além do racismo e das ofensas à honra, o discurso de ódio atrai a incidência de outras figuras típicas quando abandona o campo das ideias para fomentar ilícitos, a exemplo da incitação e da apologia ao crime (arts. 286 e 287 do Código Penal). Para tentar coibir essa propagação no ambiente virtual, em que o discurso atinge maior alcance em um curto espaço de tempo, o legislador editou a Lei nº 14.811, que inseriu uma causa de aumento de pena de metade se a incitação pública for praticada em redes sociais, canais de transmissão ou aplicativos de mensageria instantânea.

No entanto, quando essa agressão deixa de ser um estímulo genérico e assume caráter sistemático, direcionado a um usuário específico, a hostilidade se subsume ao crime de perseguição (*cyberstalking*), tipificado no artigo 147-A do Código Penal pela Lei nº 14.132. Para tutelar a privacidade e a integridade psicológica dos ataques reiterados, a norma pune a invasão da esfera pessoal do indivíduo, agressão que, nos cenários extremos de linchamento digital, pode inclusive resultar no induzimento ou instigação ao suicídio e à automutilação, sendo crimes tipificados no art. 122 do Código Penal.

Embora o ordenamento jurídico brasileiro ofereça mecanismos para a repressão de condutas ilícitas mesmo na ausência de leis específicas, faz-se necessária a atuação efetiva do Poder Legislativo para a construção de normas objetivas e critérios claros que venham acompanhados da premissa de que o Estado Democrático de Direito não admite a punição do

pensamento, mas da conduta que gera dano concreto a bens jurídicos fundamentais. Nesse sentido pondera Luiz Luisi que:

[...] o postulado em causa expressa a exigência de que as leis penais, especialmente as de natureza incriminadora, sejam claras e o mais possível certas e precisas. Trata-se de um postulado dirigido ao legislador vetando ao mesmo a elaboração de tipos penais com a utilização de expressões ambíguas, equívocas e vagas de modo a ensejar diferentes e mesmo contrastantes entendimentos. O princípio da determinação taxativa preside, portanto, a formulação da lei penal, a exigir qualificação e competência do legislador, e o uso por este de técnica correta e de uma linguagem rigorosa e uniforme. (Luisi, os princípios constitucionais penais. P.18)

Para que haja segurança jurídica, reduza-se a margem de discricionariedade do julgador e, conseqüentemente, a divergência nas decisões, a criação das leis deve observar pilares como a legalidade estrita, a dignidade humana, a exclusividade da responsabilização ulterior, o princípio do dano, o princípio da intervenção mínima e o princípio da taxatividade, através da descrição de condutas específicas, evitando que termos vagos permitam interpretações extensivas que venham a cercear a crítica política ou a liberdade de manifestação do pensamento.

Tal entendimento segue a mesma lógica adotada pelo STF no julgamento do Tema 533, em que a fixação de parâmetros objetivos serviu como antídoto ao subjetivismo, garantindo que a defesa da dignidade humana avance em harmonia com a segurança jurídica e a previsibilidade das decisões.

## 5.2 Garantismo penal

O garantismo penal brasileiro é um modelo jurídico que vincula as garantias individuais previstas na Constituição à aplicação do Direito Penal, repudiando a “coisificação” do ser humano e exigindo que o indivíduo seja tratado como um fim em si mesmo e não como meio para os interesses do Estado.

Nesse contexto Luigi Ferrajoli define o Direito Penal não apenas como um instrumento de punição, mas como um sistema de garantias voltado a minimizar a violência estatal e maximizar a proteção do cidadão. Em sua obra *Direito e Razão*, o autor sustenta que a legitimidade da pena depende da observância rigorosa de um sistema de limites estruturado em dez axiomas fundamentais, os quais condicionam o direito de punir à estrita legalidade (*nullum crimen sine lege*), à necessidade da intervenção penal (*nulla lex sine necessitate*) e à lesividade ou ofensividade ao bem jurídico (*nulla necessitas sine injuria*). Esse encadeamento lógico exige que

o dano decorra de uma conduta humana exteriorizada (*nulla injuria sine actione*) pautada pela culpabilidade (*nulla actio sine culpa*), cuja verificação pressupõe um juízo legítimo (*nulla culpa sine iudicio*) orientado pelo princípio acusatório (*nullum iudicium sine accusatione*), pelo ônus da prova cabal (*nulla accusatio sine probatione*) e pelo pleno exercício do contraditório e da ampla defesa (*nulla probatio sine defensione*).

Sob a ótica garantista, a validade de uma norma penal depende de sua estrita submissão a esses axiomas, sendo que o Estado só pode intervir na liberdade do indivíduo através de regras prévias, claras e objetivas (princípio da taxatividade), reduzindo o espaço para o arbítrio e garantindo que a jurisdição seja um ato de conhecimento (provas) e não um ato de vontade (opinião do juiz).

Dessa forma, é evidente que o discurso de ódio, ao não ter limites claros e objetivos na legislação penal brasileira, torna-se um terreno fértil para o subjetivismo judicial, em que a convicção pessoal do magistrado sobre o que é "ofensivo" pode acabar por silenciar divergências políticas ou críticas sociais legítimas, provocando, conseqüentemente, censura ou repressão desproporcional. Outrossim, o garantismo penal relaciona-se diretamente ao princípio da Intervenção Mínima, segundo o qual o Direito Penal deve atuar como *última ratio* (o último recurso) do ordenamento jurídico. Isso significa que a criminalização de condutas só é legítima quando todos os outros ramos do Direito por exemplo direito Civil ou Digital, se revelam insuficientes para proteger um bem jurídico vital. A *última ratio* evita a "inflação penal", garantindo que a força mais violenta do Estado seja reservada apenas para as ofensas mais graves contra a convivência social.

Nas palavras do Prof. Cezar Roberto Bittencourt:

O princípio da intervenção mínima, também conhecido como *ultima ratio*, orienta e limita o poder incriminador do Estado, preconizando que a criminalização de uma conduta só se legitima se constituir meio necessário para a proteção de determinado bem jurídico. Se outras formas de sanção ou outros meios de controle social revelarem-se suficientes para a tutela desse bem, a sua criminalização é inadequada e não recomendável. (BITENCOURT, Lições de direito penal – parte geral. pg. 32.)

Embora o discurso de ódio possa gerar conseqüências gravíssimas à dignidade humana e ao convívio democrático, a resposta estatal não deve se limitar exclusivamente ao Direito Penal, visto que medidas de natureza civil, administrativa, educacional e regulatória também desempenham um papel essencial na prevenção e repressão dessas condutas, especialmente no ambiente digital. O próprio entendimento firmado pelo Supremo Tribunal Federal no Tema

533 de Repercussão Geral demonstra essa lógica ao estabelecer deveres de cuidado e parâmetros para a remoção de conteúdos ilícitos pelas plataformas digitais, demonstrando que a ordenação desse ecossistema pode ser alcançada sem necessariamente recorrer à punição criminal imediata do usuário.

Portanto, a aplicação da teoria garantista de Luigi Ferrajoli ao debate sobre discurso de ódio revela a necessidade de equilíbrio entre dois valores constitucionais fundamentais: de um lado, a proteção da dignidade da pessoa humana e dos grupos vulneráveis; de outro, a preservação da liberdade de expressão e da segurança jurídica. Assim, o enfrentamento do discurso de ódio em um Estado Democrático de Direito exige critérios legais claros, respeito às garantias constitucionais e atuação estatal proporcional, evitando tanto a impunidade quanto o risco de censura ou expansão arbitrária do poder punitivo.

## 6 CONCLUSÃO

A presente pesquisa demonstrou que a liberdade de expressão constitui um dos pilares estruturantes do Estado Democrático de Direito, sendo indispensável à formação do pensamento crítico, participação política e preservação do pluralismo de ideias. Contudo, verificou-se que tal garantia constitucional não possui caráter absoluto, encontrando limites na proteção da dignidade da pessoa humana, da igualdade e dos direitos da personalidade, especialmente quando a manifestação ultrapassa o campo legítimo do debate e passa a incentivar discriminação, violência ou exclusão de grupos vulneráveis.

No contexto das redes sociais, observou-se que a propagação do discurso de ódio é potencializada pela dinâmica própria do ambiente digital, pois fenômenos como a “desinibição tóxica”, descrita por John Suler, e a “arquitetura da escolha”, desenvolvida por Cass Sunstein, contribuem para a formação de bolhas ideológicas, para a radicalização do discurso e para a desumanização do outro, criando um cenário favorável à banalização da violência virtual.

Diante desse cenário, verificou-se que o Supremo Tribunal Federal vem exercendo papel fundamental na construção de parâmetros jurídicos destinados a harmonizar a proteção da liberdade de expressão com a tutela da dignidade humana. O julgamento da ADPF 187 consolidou o entendimento de que opiniões minoritárias e manifestações impopulares permanecem protegidas constitucionalmente, desde que não configurem incitação à prática criminosa ou discurso discriminatório. Já o Tema 533 representou importante avanço ao

estabelecer deveres de cuidado às plataformas digitais e admitir a remoção célere de conteúdos manifestamente ilícitos, especialmente em hipóteses envolvendo racismo, terrorismo, violência contra grupos vulneráveis e ataques à ordem democrática.

Entretanto, a pesquisa evidenciou que a ausência de um tipo penal autônomo e de critérios legislativos objetivos acerca do discurso de ódio ainda fomenta um cenário de insegurança jurídica, cuja indeterminação amplia a margem para subjetivismos judiciais que oscilam perigosamente entre a impunidade e o risco de censura. Esse cenário de indefinição jurídica permite que o poder punitivo estatal extrapole as balizas constitucionais, tornando imprescindível a atuação do Poder Legislativo na formulação de normas rigorosamente taxativas que blindem o ordenamento contra o arbítrio e resguardem o princípio da legalidade estrita.

Dessa forma, a teoria do Garantismo Penal de Luigi Ferrajoli revelou-se essencial para a compreensão dos limites da intervenção estatal. Ao defender que o Direito Penal deve atuar como *última ratio*, o garantismo impõe que a criminalização somente seja admitida quando indispensável à proteção de bens jurídicos fundamentais e quando os demais mecanismos de controle social se mostrarem insuficientes. Assim, o enfrentamento do discurso de ódio não pode se restringir exclusivamente à repressão penal, devendo envolver medidas preventivas de natureza civil, administrativa, regulatória e educacional.

Além disso, constatou-se que o combate à violência digital exige uma atuação conjunta entre Estado, plataformas digitais e sociedade civil, através de mecanismos de autorregulamentação, políticas de transparência algorítmica, educação digital e conscientização pública para que se possa alcançar um ambiente virtual mais ético e democrático. Conforme sustentado pela Comissão Europeia contra o Racismo e a Intolerância, o enfrentamento do discurso de ódio depende não apenas da repressão jurídica, mas também da promoção de uma cultura de pluralismo, tolerância e respeito mútuo.

Conclui-se, portanto, que a proteção da liberdade de expressão e o combate ao discurso de ódio não representam valores antagônicos, mas direitos fundamentais que devem coexistir de forma harmônica no Estado Democrático de Direito. O desafio contemporâneo de impedir que as redes sociais sejam instrumentalizadas como espaços de desumanização, sem que a reação estatal gere censura ou retrocesso democrático, exige uma atuação proporcional, pautada na previsibilidade das sanções e na governança digital responsável, como meios necessários para

assegurar um ambiente virtual que preserve, simultaneamente, a livre circulação das ideias e a dignidade da pessoa humana.

## 7 REFERÊNCIAS

ALTMAN, Andrew. *Critical legal studies: a liberal critique*. Princeton: Princeton University Press, 1993.

ARAÚJO, Natália Ramos Nabuco de. *Liberdade de expressão e o discurso de ódio*. 1ª ed. Curitiba, PR: Editora Juruá, 2018

ARENDDT, Hannah. *Eichmann em Jerusalém: um relato sobre a banalidade do mal*. São Paulo: Companhia das Letras, 1999.

BARBOSA, Rui. *Oração aos moços: discurso aos bacharelados da Faculdade de Direito de S. Paulo em 1920*. São Paulo: Martinelli, Passos e Companhia, 1921.

BARROSO, Luís Roberto. *A dignidade da pessoa humana no direito constitucional contemporâneo: natureza jurídica, conteúdos mínimos e critérios de aplicação*. 5. ed. Rio de Janeiro: Renovar, 2013.

BITTENCOURT, Cezar Roberto. *Tratado de Direito Penal: Parte Geral 1*. 16. Ed. São Paulo: Editora Saraiva, 2011.

BOBBIO, Norberto. *O Futuro da Democracia: uma defesa das regras do jogo*. Tradução de Marco Aurélio Nogueira. Rio de Janeiro: Paz e Terra, 1986.

BRASIL. Constituição (1988). *Constituição da República Federativa do Brasil de 1988*. Brasília, DF: Senado Federal, 1988. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm)> Acesso em: 14 out. 2025.

BRASIL. Código Penal. Decreto-Lei nº 2.848, de 7 de dezembro de 1940. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/decreto-lei/del2848.htm](https://www.planalto.gov.br/ccivil_03/decreto-lei/del2848.htm)> Acesso em: 14 out. 2025.

BRASIL. Lei nº 7.716, de 5 de janeiro de 1989. Define os crimes resultantes de preconceito de raça ou de cor. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/leis/l7716.htm](https://www.planalto.gov.br/ccivil_03/leis/l7716.htm)>. Acesso em: 14 out. 2025.

BRASIL. Lei nº 14.197, de 1º de setembro de 2021. Altera o Decreto-Lei nº 2.848/1940 (Código Penal) para dispor sobre os crimes contra o Estado Democrático de Direito. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/\\_ato2019-2022/2021/lei/L14197.htm](https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14197.htm)> Acesso em: 14 out. 2025.

COMISSÃO EUROPEIA CONTRA O RACISMO E A INTOLERÂNCIA (ECRI). Hate speech and violence. Conselho da Europa. Disponível em:

<<https://www.coe.int/en/web/european-commission-against-racism-and-intolerance/hate-speech-and-violence>>. Acesso em: 08 maio 2026.

ESTADOS Unidos e Brasil: Entenda a diferença entre a liberdade de expressão no Brasil e nos EUA - Disponível em: <<https://noticias.uol.com.br/comprova/ultimas-noticias/2024/04/18/entenda-a-diferenca-entre-a-liberdade-de-expressao-no-brasil-e-nos-eua.htm?cmpid=copiaecola>> Acesso em: 14 maio 2026.

FERRAJOLI, Luigi. Direito e razão: teoria do garantismo penal. 5. ed. São Paulo: Revista dos Tribunais, 2002.

FGV DIREITO SP; CONIB. Guia para análise de discurso de ódio. São Paulo: FGV Direito SP/CONIB, 2017.

GRECO, Rogério. Curso de Direito Penal: parte geral. 26. ed. Rio de Janeiro: Impetus, 2022.

LUISI, Luiz. Os princípios constitucionais penais. 2. ed. rev. e aum. Porto Alegre: Sergio Antonio Fabris Editor, 2003. p. 24.

MATSUDA, Mari J. Public response to racist speech: considering the victim's story. Michigan Law Review, v. 87, n. 8, p. 2320-2381, 1989.

MILL, John Stuart. Sobre a Liberdade\*. Tradução de Alberto da Rocha Barros. São Paulo: Martins Fontes, 2000.

NUCCI, Guilherme de Souza. Código Penal Comentado. 20. ed. São Paulo: Forense, 2023.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Estratégia e Plano de Ação das Nações Unidas sobre Discurso de Ódio. Nova York: ONU, 2019. Disponível em: <<https://www.un.org/en/hate-speech/un-strategy-and-plan-of-action-on-hate-speech>> Acesso em: 22 maio 2026.

SARMENTO, Daniel. A Liberdade de Expressão e o Problema do "Discurso de Ódio". Rio de Janeiro: Editora FGV, 2015.

SARLET, Ingo Wolfgang. A eficácia dos direitos fundamentais: uma teoria geral dos direitos fundamentais na perspectiva constitucional. 14. ed. Porto Alegre: Livraria do Advogado, 2021.

SILVA, José Afonso da. Curso de direito constitucional positivo. 46. ed. São Paulo: Malheiros, 2023.

SULER, John. The online disinhibition effect. CyberPsychology & Behavior, v. 7, n. 3, p. 321-326, 2004. Disponível em: < <https://pubmed.ncbi.nlm.nih.gov/35438520/>> Acesso em: 27 de Abril de 2026

SUPREMO TRIBUNAL FEDERAL (Brasil). Recurso Extraordinário n. 1.057.258/RJ. Tema 533 da repercussão geral. Relator: Min. Luiz Fux. Supremo Tribunal Federal, Brasília, DF. Disponível em: <<https://portal.stf.jus.br/processos/detalhe.asp?incidente=5217273>> Acesso em: 20 nov. 2025.

SUPREMO TRIBUNAL FEDERAL (Brasil). Arguição de Descumprimento de Preceito Fundamental n. 187/DF. Relator: Min. Celso de Mello. Julgado em 15 jun. 2011. Disponível em: <<https://www.stf.jus.br>> Acesso em: 14 out. 2025.

SUPREMO TRIBUNAL FEDERAL (STF). Jurisprudência sobre liberdade de expressão e discurso de ódio. Brasília, DF, 2020. Disponível em: <<https://jurisprudencia.stf.jus.br>>. Acesso em: 14 out. 2025.

SUNSTEIN, Cass R. #Republic: divided democracy in the age of social media. Princeton: Princeton University Press, 2017.