



B1

ISSN: 2595-1661

ARTIGO

Listas de conteúdos disponíveis em [Portal de Periódicos CAPES](#)

## Revista JRG de Estudos Acadêmicos

Página da revista:

<https://revistajrg.com/index.php/jrg>



### O papel dos algoritmos na regulação da liberdade de expressão nas plataformas digitais

The role of algorithms in regulating freedom of expression on digital platforms

DOI: 10.55892/jrg.v8i18.2122

ARK: 57118/JRG.v8i18.2122

Recebido: 14/05/2025 | Aceito: 21/05/2025 | Publicado *on-line*: 22/05/2025

**Mírio Jean Araújo dos Santos<sup>1</sup>**

<https://orcid.org/0009-0004-4435-8570>

<https://lattes.cnpq.br/3755604433126616>

Universidade Estadual do Tocantins -Unitins, TO, Brasil

E-mail: mirioje4n@gmail.com

**Sara Brigida Farias Ferreira<sup>2</sup>**

<https://orcid.org/0000-0001-6588-2305>

<http://lattes.cnpq.br/9477160915420773>

Universidade Estadual do Tocantins -Unitins, TO, Brasil

E-mail: sara\_farias@hotmail.com



#### Resumo

O avanço das tecnologias digitais e o uso de algoritmos nas redes sociais têm impactado significativamente a liberdade de expressão. Embora essencial à democracia, esse direito enfrenta novos desafios no ambiente digital, especialmente diante da atuação de algoritmos que filtram, promovem ou removem conteúdos com base em critérios opacos de engajamento. Essa dinâmica pode reproduzir desigualdades e comprometer direitos fundamentais, como o contraditório e a ampla defesa. O presente estudo analisa a ausência de uma regulação específica no Brasil sobre moderação algorítmica, evidenciando a limitação do Marco Civil da Internet (Lei nº 12.965/2014) frente aos desafios atuais. Inspirando-se no Digital Services Act da União Europeia, a pesquisa tem como objetivo propor caminhos normativos e institucionais para garantir maior transparência e responsabilidade das plataformas digitais. Utiliza-se metodologia qualitativa, com revisão bibliográfica, análise documental e estudo de jurisprudência do Supremo Tribunal Federal. Justifica-se a investigação pela necessidade de atualizar o marco regulatório brasileiro diante do papel central dos algoritmos na organização do espaço informacional. Conclui-se que a ausência de regras claras sobre moderação automatizada representa risco à liberdade de expressão e à democracia. Recomenda-se, portanto, a adoção de auditorias independentes, revisão humana em decisões automatizadas e mecanismos de recurso para usuários, como medidas capazes de equilibrar inovação tecnológica e proteção de direitos fundamentais.

<sup>1</sup> Graduando em Direito pela Universidade Estadual do Tocantins -Unitins, TO, Brasil.

<sup>2</sup> Mestre em Planejamento e Desenvolvimento Regional e Urbano na Amazônia (PPGPAM), e mestra em Propriedade Intelectual e Transferência de Tecnologia para a Inovação (PROFNIT), pela Universidade Federal do Sul e Sudeste do Pará (UNIFESSPA). Bacharela em Direito, com habilitação em Relações Sociais, pela Universidade Federal do Paraná - UFPR.

**Palavras-chave:** Liberdade de expressão. Algoritmos. Regulação digital. Direitos fundamentais.

### **Abstract**

*The advancement of digital technologies and the use of algorithms on social media platforms have significantly impacted freedom of expression. Although essential to democracy, this right faces new challenges in the digital environment, particularly due to algorithms that filter, promote, or remove content based on opaque engagement criteria. This dynamic can reproduce inequalities and undermine fundamental rights such as the right to adversarial proceedings and full defense. This study analyzes the lack of specific regulation in Brazil regarding algorithmic moderation, highlighting the limitations of the Brazilian Civil Rights Framework for the Internet (Law No. 12.965/2014) in addressing current challenges. Inspired by the European Union's Digital Services Act, the research aims to propose normative and institutional pathways to ensure greater transparency and accountability from digital platforms. A qualitative methodology is used, based on literature review, document analysis, and examination of case law from the Brazilian Supreme Federal Court. The investigation is justified by the urgent need to update the national regulatory framework, given the central role algorithms play in shaping the informational space. It concludes that the absence of clear rules on automated moderation poses a concrete risk to freedom of expression and democratic functioning. Therefore, it recommends the adoption of independent audits, human review of automated decisions, and effective appeal mechanisms for users, as measures capable of balancing technological innovation with the protection of fundamental rights.*

**Keywords:** Freedom of expression. Algorithms. Digital regulation. Fundamental rights.

## **1. Introdução**

O avanço das tecnologias digitais e a crescente atuação das plataformas online na mediação da comunicação social têm transformado profundamente a forma como se exerce a liberdade de expressão. Se, por um lado, a internet ampliou o acesso à informação e democratizou o debate público, por outro, introduziu novos riscos relacionados à manipulação algorítmica, à disseminação de desinformação e à violação de direitos fundamentais (Bioni, 2020; Doneda, 2021). Nesse contexto, torna-se necessário refletir sobre os limites e as responsabilidades desse novo ecossistema comunicacional, especialmente no que tange ao uso de algoritmos para a moderação de conteúdo nas redes sociais.

A liberdade de expressão, consagrada no artigo 19 do Pacto Internacional dos Direitos Civis e Políticos (ONU, 1966) e no artigo 5º, inciso IV, da Constituição Federal de 1988, é um pilar essencial da democracia. No entanto, a atuação de sistemas automatizados, muitas vezes opacos e baseados em dados enviesados, levanta preocupações quanto à transparência, à proteção da privacidade, à dignidade da pessoa humana e ao devido processo legal (Piovesan, 2016; O'neil, 2016; Eubanks, 2018). A ausência de regulação específica sobre algoritmos de moderação em plataformas digitais agrava esse cenário, comprometendo a efetivação dos direitos fundamentais no ambiente virtual.

O problema central que este estudo busca abordar é a lacuna normativa e institucional existente no Brasil quanto à regulação da moderação algorítmica, especialmente diante dos riscos que ela representa para a liberdade de expressão e outros direitos fundamentais. O debate torna-se ainda mais relevante diante da

atuação crescente das plataformas digitais como agentes que definem o que pode ou não circular no espaço público online, influenciando diretamente a formação da opinião pública e a qualidade democrática (Pariser, 2011; Mendes, 2019).

Diante desse cenário, o objetivo geral desta pesquisa é analisar os impactos da moderação algorítmica sobre a liberdade de expressão no Brasil, à luz dos princípios constitucionais e das experiências legislativas internacionais, com o intuito de propor caminhos para uma regulação mais transparente, democrática e garantidora de direitos.

A justificativa desta pesquisa reside na necessidade urgente de atualização do marco jurídico nacional frente à complexidade dos desafios digitais contemporâneos. Diante do poder crescente das plataformas e da centralidade dos algoritmos na filtragem de informações, é imprescindível discutir limites normativos e institucionais que assegurem o equilíbrio entre inovação tecnológica e proteção dos direitos fundamentais. Ao tratar desse tema, o presente estudo busca contribuir com a construção de uma governança digital mais justa, ética e democrática.

## 2. Metodologia

A metodologia adotada é qualitativa, de cunho teórico e documental, baseada na análise da legislação brasileira — especialmente do Marco Civil da Internet (Lei nº 12.965/2014), da jurisprudência do Supremo Tribunal Federal (como a ADPF 403/2017 e a ADI 5527/2016), de projetos legislativos em tramitação, como o Projeto de Lei nº 2630/2020, bem como de marcos normativos internacionais, a exemplo do Digital Services Act (União Europeia, 2022). Também são examinados estudos doutrinários e pesquisas interdisciplinares que discutem os efeitos dos algoritmos no direito, na comunicação e na sociedade (Reis; Damasceno, 2019; Amaral, 2023).

## 3. Algoritmos, plataformas digitais e a necessidade de regulação

A regulação dos algoritmos tornou-se fundamental para conciliar a liberdade de expressão com a segurança digital e a proteção dos direitos fundamentais. No Brasil, o Marco Civil da Internet (Lei nº 12.965/2014) estabelece parâmetros importantes de responsabilidade para provedores de aplicações e serviços. Em especial, o artigo 19 dispõe que os provedores somente podem ser responsabilizados por conteúdos gerados por terceiros mediante ordem judicial — com exceção dos casos envolvendo pornografia infantil. Já o artigo 21 trata da proteção de dados pessoais, exigindo notificação do usuário em situações de vazamento de informações. No entanto, o Marco Civil não contempla de forma direta a moderação algorítmica, o que limita sua eficácia diante das complexas dinâmicas atuais do ambiente digital. Para Bioni (2020), é imprescindível que a legislação brasileira avance, incorporando dispositivos específicos sobre transparência e responsabilidade no uso de algoritmos.

Nesse contexto, a União Europeia desponta como referência com o Digital Services Act (DSA), aprovado em 2022, que institui um modelo robusto de regulação das plataformas digitais. O DSA impõe obrigações de transparência, como auditorias periódicas dos algoritmos, relatórios públicos sobre atividades de moderação de conteúdo e a criação de canais de recurso para usuários impactados por decisões automatizadas. Além disso, proíbe práticas discriminatórias em sistemas automatizados, estabelecendo diretrizes que resguardam os direitos fundamentais. Doneda (2021) sustenta que o Brasil pode adotar medidas inspiradas nesse modelo europeu, adaptando-as ao seu ordenamento jurídico e à realidade sociopolítica nacional, com vistas a garantir a observância de princípios constitucionais como a dignidade da pessoa humana e a igualdade.

Em território nacional, o Projeto de Lei nº 2.630/2020 — conhecido como “PL das Fake News” — encontra-se em tramitação na Câmara dos Deputados e visa enfrentar o fenômeno da desinformação. Entre as medidas propostas estão a rotulagem de conteúdos enganosos, a identificação de bots e a exigência de transparência quanto à veiculação de anúncios políticos. Contudo, o projeto ainda falha ao não regulamentar diretamente o funcionamento dos algoritmos de moderação, o que representa uma lacuna relevante. Diversos parlamentares e especialistas têm apontado, nos debates legislativos de 2023, que o foco excessivo na desinformação pode, sem salvaguardas adequadas, comprometer a liberdade de expressão. O Supremo Tribunal Federal, por sua vez, tem se manifestado sobre a temática: na ADI 4451, sob relatoria do Ministro Luiz Fux, a Corte destacou a importância de equilibrar a liberdade de expressão com a responsabilidade das plataformas, defendendo critérios transparentes de moderação, especialmente por empresas como o X.

Sarmiento (2018) propõe que futuras regulamentações contemplem medidas como: (1) revisão humana obrigatória em decisões automatizadas de remoção; (2) garantia do contraditório e da ampla defesa para os usuários impactados; e (3) participação ativa da sociedade civil na formulação das políticas de moderação. Entre as propostas concretas para o cenário brasileiro, destacam-se: (1) a implementação de auditorias independentes sobre algoritmos, nos moldes do DSA; (2) a criação de um órgão regulador específico para o ambiente digital — à semelhança da Anatel no setor de telecomunicações; e (3) a revisão do Marco Civil da Internet, com a inclusão de dispositivos sobre moderação algorítmica.

Tais medidas têm o potencial de transformar os algoritmos de simples mecanismos operacionais em instrumentos de promoção da inclusão e da democracia, alinhando o desenvolvimento tecnológico aos valores fundamentais da Constituição Federal, como defendem Mendes (2019) e Bioni (2020).

Alguns autores abordam o impacto dos algoritmos e da inteligência artificial (IA) no Direito e no sistema judiciário brasileiro, enfatizando os desafios éticos e práticos dessa tecnologia. Cambi, Alves e Alves (2019) analisam o uso de algoritmos e da metodologia fuzzy na aplicação do Direito, ressaltando que o objetivo é tornar o sistema judicial mais objetivo, mas alertam para o risco de injustiças caso esses mecanismos comprometam a interpretação jurídica. A pesquisa sugere que o uso de IA deve sempre respeitar a dignidade humana e os princípios constitucionais, assegurando que decisões automatizadas não prejudiquem os resultados legais, mas atuem como ferramentas de apoio.

Warren (2022) foca na necessidade de neutralizar os vieses nos processos de coleta e análise de dados para IA, propondo que a ideia de “justiça” na IA passa por reconhecer a heterogeneidade dos dados e ajustar algoritmos para refletir contextos diversos. Ele explora o aprendizado federado como uma técnica que possibilita a combinação de dados de diferentes fontes, reduzindo vieses que poderiam ser reproduzidos em um conjunto de dados isolado. Warren enfatiza que um desenvolvimento ético da IA exige constantes ajustes para garantir resultados mais equitativos e precisos.

Faria e Damasceno (2019) discutem o contexto da Indústria 4.0 e suas implicações para o trabalho jurídico, especialmente o impacto da automação e da IA na prática do Direito. Para eles, a evolução tecnológica exige um novo perfil de profissional, preparado para lidar com sistemas automatizados e algoritmos de aprendizado que, por sua capacidade de gerar soluções autônomas, transformam a prática jurídica. Esses autores apontam que a automação é inevitável e traz avanços,

mas também demanda qualificações adicionais para os operadores do Direito, a fim de evitar a degradação do trabalho e garantir a eficácia da tecnologia aplicada ao setor.

Amaral (2023) apresenta uma análise dialética sobre o avanço tecnológico, explorando as visões utópicas e distópicas da tecnologia no sistema jurídico. A autora discute como a IA e os algoritmos moldam os comportamentos e decisões no Judiciário, e alerta para os riscos de vieses humanos que podem ser reproduzidos pelos sistemas. A visão distópica da tecnologia, para Amaral, sugere que a IA, apesar de seu potencial de inovação, precisa ser rigidamente regulamentada para prevenir desigualdades e garantir que o progresso digital se traduza em uma justiça acessível e imparcial.

No estudo de Faustino e Silva (2023), analisa-se a ocorrência de discursos de ódio direcionados ao Supremo Tribunal Federal (STF) nas redes sociais. Os autores apontam que esses discursos, baseados na teoria do amigo-inimigo de Carl Schmitt, se propagam com a finalidade de manipular a opinião pública, criando um inimigo comum para gerar engajamento.

Através dos algoritmos das redes sociais, o conteúdo que se alinha com as preferências dos usuários é amplamente disseminado, reforçando as bolhas de opinião e facilitando a circulação de desinformação e ódio. Os algoritmos são cruciais nesse contexto, pois operam para amplificar as postagens que causam reações, promovendo a polarização e intensificando o alcance de discursos extremistas (Faustino; Silva, 2023).

Já Rodrigo da Guia Silva e Marcela Guimarães Barbosa Silva (2023) discutem a aplicação de direitos fundamentais no contexto digital, com foco nas garantias processuais no âmbito da exclusão de perfis e conteúdos em redes sociais. Através de algoritmos de monitoramento, as plataformas muitas vezes utilizam sistemas automatizados para identificar e excluir conteúdos que supostamente violam as diretrizes de comunidade, deixando o usuário sem direito ao contraditório. Esse uso de algoritmos levanta questões sobre a legitimidade dos critérios empregados e a necessidade de garantir o contraditório antes de uma exclusão definitiva, destacando a importância de uma regulação que contemple a revisão humana em casos de exclusão (Silva; Silva, 2023).

Paralelamente, Franco (2019) analisa a utilização de bots sociais para a disseminação de conteúdo falso, destacando o papel dos algoritmos de inteligência artificial no funcionamento desses bots. Esses programas são projetados para simular a atividade humana, interagindo com usuários reais e ampliando a disseminação de desinformação.

Algoritmos de aprendizado de máquina permitem que os bots adaptem suas interações, o que dificulta a identificação dos mesmos, além de potencializar seu impacto nas redes ao replicarem rapidamente informações enganosas e promoverem narrativas fabricadas. A autora salienta que, para o combate eficaz à disseminação de desinformação, é essencial compreender as dinâmicas dos algoritmos e suas implicações na manipulação da opinião pública (Franco, 2019).

Nesse mesmo contexto, Eduardo Mingorance de Freitas Gouvêa (2019) explora os desafios da privacidade no contexto das redes sociais, onde a atuação de algoritmos para o monitoramento de conteúdos expostos publicamente gera preocupações. O autor argumenta que a lógica algorítmica, ao priorizar conteúdos mais visualizados ou engajantes, acaba incentivando os usuários a compartilhar cada vez mais informações pessoais, aumentando a vulnerabilidade desses dados. Ele questiona se o uso desses algoritmos, que favorecem a exposição da vida privada,

está alinhado com os direitos de privacidade, observando a necessidade de uma regulamentação mais robusta para proteger os dados dos usuários contra o uso excessivo (Gouvêa, 2019).

#### **4. Liberdade de expressão na era digital: desafios constitucionais e jurisprudência brasileira**

A liberdade de expressão, consagrada no artigo 19 do Pacto Internacional dos Direitos Civis e Políticos (ONU, 1966) e no artigo 5º, inciso IV, da Constituição Federal de 1988, constitui um dos pilares do Estado Democrático de Direito. O referido artigo do Pacto assegura a todos o direito de expressar livremente suas opiniões, sem interferências, admitindo, no entanto, restrições legítimas voltadas à proteção da segurança pública ou dos direitos de terceiros. No contexto brasileiro, a Constituição garante a livre manifestação do pensamento, vedando o anonimato e estabelecendo limites sempre que esse direito colide com a dignidade da pessoa humana.

Com o avanço das plataformas digitais — como X, YouTube e WhatsApp —, a liberdade de expressão passou a assumir proporções globais, o que, embora positivo sob o aspecto da democratização das vozes, trouxe desafios complexos. Mendes (2019) observa que a internet potencializa a circulação de ideias diversas, mas também amplia a disseminação de discursos de ódio, como o racismo e a xenofobia, além de facilitar a prática de crimes virtuais, como a calúnia e a difamação. Reale Júnior (2009), por sua vez, sustenta que a liberdade de expressão não é um direito absoluto, devendo ser restringida quando incita à violência ou à discriminação.

No Brasil, o Supremo Tribunal Federal tem desempenhado papel decisivo na mediação desses conflitos. Na ADPF 403 (2017), sob relatoria do Ministro Alexandre de Moraes, a Corte decidiu que a remoção de conteúdos ofensivos nas redes sociais deve seguir o devido processo legal, rechaçando atitudes unilaterais de plataformas como o X.

A decisão deixou claro que a liberdade de expressão não pode ser utilizada para justificar manifestações de ódio que atentem contra a dignidade humana. Já na ADI 5527 (2016), o STF, por meio do voto do Ministro Edson Fachin, reconheceu a constitucionalidade do Marco Civil da Internet, fortalecendo a proteção à liberdade de expressão na rede ao condicionar a retirada de conteúdos à determinação judicial. Outro marco relevante foi a ADPF 548 (2018), que assegurou a liberdade de imprensa frente a tentativas de censura judicial, com repercussões diretas sobre o ambiente digital. Mais recentemente, a ADI 6996 (2023), ainda em tramitação, discute a responsabilidade das plataformas pela propagação de desinformação, evidenciando a constante evolução da jurisprudência nesse campo.

A desinformação, especialmente no contexto eleitoral, é outro desafio premente. O escândalo envolvendo a Cambridge Analytica nas eleições de 2018, revelando o uso indevido de dados do Facebook para manipulação de conteúdos, e a atuação de bots no WhatsApp e no X durante as eleições de 2022, conforme reportado pela Folha de S.Paulo (2022), demonstram os riscos à integridade do debate público. Bioni (2020) destaca que a proliferação de fake news compromete a formação de uma opinião pública consciente, exigindo a adoção de limites jurídicos e éticos que assegurem, ao mesmo tempo, a liberdade de expressão e a proteção contra conteúdos lesivos.

Diante desses desafios, impõe-se a tarefa de equilibrar o direito à livre manifestação com a necessidade de conter práticas que atentem contra a democracia, sobretudo nas plataformas digitais que hoje moldam o espaço público contemporâneo.

De acordo com Pedro Henrique Scott de Senna e Letícia Thomasi Jahnke Botton (2020), a liberdade de expressão é um direito fundamental historicamente evoluído e essencial para a comunicação em massa, particularmente na internet, onde se torna central para a sociedade em rede.

No entanto, esses autores destacam que, com o avanço da tecnologia, novos conflitos emergem entre a liberdade de expressão e outros direitos, como o direito à honra, especialmente quando se trata de crimes cibernéticos e manifestações que podem ser penalizadas (Senna; Botton, 2020)

Miguel Reale Júnior (2009) explora os limites da liberdade de expressão à luz de casos julgados pelo Supremo Tribunal Federal (STF), enfatizando que, embora a liberdade de expressão seja essencial ao Estado Democrático de Direito, ela não pode servir de pretexto para práticas ilícitas, como a incitação ao ódio racial. Reale ressalta que o STF, em casos de discursos racistas, opta por restringir a liberdade de expressão quando ela colide com outros valores constitucionais, como a dignidade humana (Reale, 2009).

Ana Claudia Sousa de Campos e Ila Barbosa Bittencourt (2020) também abordam a liberdade de expressão na internet, defendendo que as plataformas digitais têm o dever de respeitar esse direito ao moderar o conteúdo dos usuários. Os autores ressaltam que qualquer medida regulatória excessiva pode desvirtuar a função democrática da internet, promovendo uma censura privada que prejudica o diálogo público e ameaça a criatividade e a troca de ideias (Sousa de Campos; Bittencourt, 2020).

José Sérgio Miranda e Ivelise Fonseca De Matteu (2022) destacam a questão das fake news como uma ameaça à liberdade de expressão. Segundo eles, a disseminação intencional de desinformação não só ataca o direito à informação correta, mas também distorce a liberdade de expressão ao ser usada para manipular e induzir o público ao erro. Dessa forma, embora algumas correntes defendam o direito de difundir qualquer conteúdo como parte da liberdade de expressão, os autores argumentam que o caráter manipulador e lesivo das fake news as coloca fora desse direito (Miranda; De Matteu, 2022).

Por fim, Rocco Antonio Rangel Rosso Nelson (2023) analisa a liberdade de expressão como um indicador da saúde democrática, afirmando que tentativas institucionais de limitar esse direito, como ocorreu no Brasil entre 2018 e 2022, podem comprometer a integridade do regime democrático. Para ele, a liberdade de expressão é crucial para sustentar uma democracia vibrante, e qualquer restrição a esse direito deve ser cuidadosamente ponderada para evitar o esfacelamento dos princípios democráticos (Nelson, 2023).

## **5. Direito digital, algoritmos e redes sociais: impactos sobre liberdade de expressão e direitos fundamentais**

Estudos analisados abordam temas variados dentro do direito digital e das redes sociais, com ênfase na proteção de dados, liberdade de expressão, responsabilidade civil e os algoritmos que orientam o comportamento das plataformas digitais. O trabalho de Godoi (2024) explora como algoritmos de recomendação utilizam grandes volumes de dados para oferecer conteúdo personalizado aos usuários de redes sociais, visando mantê-los conectados por mais tempo. Esses algoritmos, programados por meio de inteligência artificial e aprendizado de máquina, identificam padrões de comportamento para recomendar conteúdos que são potencialmente lesivos, especialmente para crianças e adolescentes. Godoi enfatiza que a presença de menores em ambientes digitais traz desafios de proteção, pois o

tratamento de dados nesses casos muitas vezes infringe normas, propondo que haja transparência e regulamentação mais rigorosa na programação e gestão desses algoritmos para proteger os jovens usuários (Godoi, 2024)

Casabona (2006) discute as características dos delitos cibernéticos, enfatizando a capacidade dos sistemas telemáticos de armazenar e processar grandes volumes de informação e a ausência de uma hierarquia de controle efetiva nas redes. Ele alerta que os crimes cibernéticos aumentaram com a facilidade de acesso não autorizado a dados, espionagem e manipulação de informações, ampliados pelo uso dos algoritmos que ajudam criminosos a mascarar suas identidades e dificultam a atuação da justiça. A ausência de regulamentação clara aumenta a vulnerabilidade das redes e amplia os danos potenciais às vítimas (Casabona, 2006)

No estudo de Frumi (2022), é discutido o impacto do Marco Civil da Internet na definição da responsabilidade civil dos provedores diante de atos ilícitos, como o cyberbullying. A autora explica que o artigo 19 do Marco Civil exige notificação judicial específica para a remoção de conteúdos ofensivos, o que muitas vezes causa atraso na resolução dos conflitos e gera uma colisão entre a liberdade de expressão e o direito à privacidade. Frumi sugere que a regulamentação falha em proteger completamente as vítimas de cyberbullying, demonstrando que a aplicação dos algoritmos na moderação de conteúdos deve ser aperfeiçoada (Frumi, 2022)

Sala (2021) e Matteu et al. (2023) complementam a discussão abordando o papel das redes sociais na regulação de discursos ofensivos e de ódio. Sala propõe que as plataformas assumam um papel ativo na promoção da liberdade de expressão de forma equilibrada, pois os algoritmos podem ampliar a polarização social ao permitir a disseminação de discursos de ódio, tornando as redes um espaço hostil para as minorias.

Para Matteu et al., a liberdade de expressão é um direito fundamental, mas deve ser limitada em casos de discurso de ódio, onde algoritmos de redes sociais têm contribuído para espalhar tais discursos, impondo uma necessidade de políticas internas rigorosas que contenham o avanço de conteúdos que incitam violência e discriminação (Sala, 2021; Matteu et al., 2023)

A liberdade de expressão é um direito fundamental assegurado pela Constituição Federal de 1988, essencial para a manutenção de uma sociedade democrática, pois garante às pessoas a possibilidade de expressarem suas opiniões e convicções sem restrições indevidas. Contudo, esse direito não é ilimitado e pode entrar em conflito com outros direitos fundamentais, como a dignidade humana e a igualdade, exigindo, portanto, regulamentações adequadas para evitar abusos.

Com o avanço das redes sociais, essa discussão se tornou ainda mais urgente, já que essas plataformas frequentemente servem de canal para a propagação de discursos de ódio e discriminação. Como observa Sarmiento (2018), o desafio contemporâneo reside em traçar limites claros entre a liberdade de expressão e o discurso de ódio. No contexto brasileiro, o crescimento de manifestações racistas e preconceituosas online ressalta a necessidade de uma reflexão aprofundada sobre o papel do Código Penal e do Marco Civil da Internet para lidar com essas situações.

O Supremo Tribunal Federal (STF) também tem abordado esse tema, ampliando a interpretação dos limites entre a liberdade de expressão e o discurso de ódio. Piovesan (2016) destaca que o discurso de ódio, ao incitar violência, fere os princípios de convivência em uma sociedade democrática. Nessas discussões, o STF busca equilibrar a garantia da liberdade de expressão com a proteção contra

discriminação e violência, de modo a promover um ambiente de respeito e segurança para todos.

Plataformas digitais como X, Instagram e YouTube utilizam algoritmos baseados em inteligência artificial (IA) para moderar conteúdos, definindo o que será promovido, ocultado ou removido a partir de critérios como engajamento — curtidas, compartilhamentos e visualizações. Embora eficientes do ponto de vista técnico, esses sistemas, geralmente treinados com dados históricos, levantam sérias preocupações quanto à preservação de direitos fundamentais. Piovesan (2016) alerta que a exclusão de conteúdos sem transparência viola o princípio do contraditório, enquanto Doneda (2021) critica a remoção automatizada de publicações sem notificação prévia ou possibilidade de recurso, o que afronta o devido processo legal previsto no artigo 5º, inciso LV, da Constituição Federal.

O’Neil (2016) e Eubanks (2018) demonstram que algoritmos podem reproduzir e reforçar preconceitos históricos, como discriminações de cunho racial ou político, levando à supressão desproporcional de vozes pertencentes a grupos minoritários. Em 2022, a plataforma X foi alvo de críticas após excluir publicações de ativistas negros que denunciavam episódios de racismo, sob a justificativa de violação das políticas comunitárias, sem apresentar justificativas claras ou oferecer mecanismos de contestação (Electronic Frontier Foundation, 2022). Relatório da organização revelou que aproximadamente 60% das remoções atingiam conteúdos produzidos por minorias, indicando a presença de vieses algorítmicos enraizados nos dados utilizados para o treinamento dos sistemas.

Pariser (2011) adverte que os sistemas de recomendação também contribuem para o isolamento informacional, ao criar “bolhas” que priorizam conteúdos sensacionalistas e polarizadores com alto potencial de engajamento — como vídeos de teor extremista no YouTube —, intensificando a fragmentação do debate público. Durante as eleições de 2022, por exemplo, diversas postagens de cunho político foram removidas pelo X, incluindo denúncias de candidatos sobre possíveis irregularidades no processo eleitoral, o que gerou denúncias de censura, conforme noticiado pela Folha de S.Paulo (2022).

Sob o aspecto técnico, os algoritmos de moderação operam majoritariamente por meio de técnicas de machine learning, como redes neurais artificiais, treinadas com grandes volumes de dados para identificar conteúdos classificados como “inapropriados”. Entretanto, conforme observa O’Neil (2016), esses modelos não são neutros: eles refletem os vieses presentes nos dados utilizados ou mesmo na concepção dos programadores. Um algoritmo treinado com base nas interações dos usuários, por exemplo, pode acabar priorizando conteúdos polêmicos, justamente por gerarem mais engajamento. Eubanks (2018) ressalta que a opacidade desses sistemas — sua incapacidade de serem compreendidos ou auditados por usuários comuns — perpetua desigualdades estruturais e exige a adoção de mecanismos independentes de auditoria para identificação e correção de distorções.

No Brasil, a ausência de uma regulamentação específica voltada à moderação algorítmica agrava o cenário. Diante disso, torna-se imperativo exigir maior transparência e responsabilidade (accountability) por parte das plataformas digitais, assegurando que a moderação de conteúdo ocorra em conformidade com os princípios constitucionais e os direitos fundamentais dos usuários.

## 6. Considerações Finais

A crescente presença dos algoritmos e da inteligência artificial nas plataformas digitais tem gerado implicações relevantes para a proteção de direitos fundamentais, especialmente no que se refere à liberdade de expressão. Embora esse direito seja essencial à democracia e esteja amplamente reconhecido no ordenamento jurídico nacional e internacional, seu exercício no ambiente virtual exige a mediação de valores igualmente relevantes, como a dignidade da pessoa humana, a privacidade e a não discriminação.

A análise demonstrou que o Marco Civil da Internet representa um avanço importante na regulamentação do uso da rede no Brasil, mas se mostra insuficiente para lidar com as complexidades atuais da moderação algorítmica, que frequentemente opera de maneira opaca e sem garantias mínimas de contraditório e ampla defesa. A experiência europeia, com a aprovação do Digital Services Act, oferece um modelo interessante ao estabelecer critérios mais rigorosos de transparência e controle dos sistemas automatizados de decisão.

A jurisprudência do Supremo Tribunal Federal tem sinalizado a importância de equilibrar a liberdade de expressão com a responsabilização das plataformas digitais, reafirmando a necessidade de observância dos princípios constitucionais no ambiente digital. Ao mesmo tempo, estudos apontam que algoritmos podem reproduzir vieses históricos, acentuar desigualdades sociais e contribuir para a propagação de discursos de ódio, fake news e outras práticas nocivas à integridade do debate público.

Nesse cenário, torna-se urgente a atualização do marco regulatório brasileiro, com a inclusão de dispositivos específicos sobre a governança dos algoritmos, a exigência de auditorias independentes, a criação de instâncias reguladoras especializadas e a implementação de salvaguardas processuais que garantam a participação dos usuários. É fundamental que o desenvolvimento tecnológico caminhe ao lado de mecanismos jurídicos eficazes, capazes de assegurar que a inovação não se dê à custa da justiça, da inclusão e da democracia.

Portanto, o desafio contemporâneo consiste em construir um ecossistema digital regulado de forma transparente, ética e participativa, no qual a liberdade de expressão possa ser exercida em consonância com os demais direitos fundamentais, assegurando um espaço público plural, seguro e verdadeiramente democrático.

## Referências

AMARAL, Ana Luiza Lacerda. Entre a utopia e a distopia tecnológica no direito – Análises e predições do sistema judiciário brasileiro. **Revista dos Tribunais**, vol. 1050, p. 121-141, abr. 2023.

BIONI, Bruno Ricardo. **Proteção de dados pessoais: a função e os limites do consentimento**. São Paulo: Forense, 2020.

BRASIL. **Decreto no 592, de 6 de julho de 1992**. Atos Internacionais. Pacto Internacional sobre Direitos Civis e Políticos. Promulgação. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/decreto/1990-1994/d0592.htm](https://www.planalto.gov.br/ccivil_03/decreto/1990-1994/d0592.htm). Acesso em: 10 mai. 2025.

BRASIL. Supremo Tribunal Federal. **Ação Direta de Inconstitucionalidade nº 4451**. Relator: Min. Luiz Fux. Disponível em:

<https://portal.stf.jus.br/processos/detalhe.asp?incidente=3938343>. Acesso em: 10 mai. 2025.

BRASIL. Supremo Tribunal Federal. **Arguição de Descumprimento de Preceito Fundamental nº 548**. Relator: Min. Roberto Barroso, 25 out. 2018. Disponível em: <https://portal.stf.jus.br/processos/detalhe.asp?incidente=5576416>. Acesso em: 13 mai. 2025.

BRASIL. Supremo Tribunal Federal. **Ação Direta de Inconstitucionalidade nº 5527**. Relator: Min. Edson Fachin, 18 jun. 2020. Disponível em: <https://portal.stf.jus.br/processos/detalhe.asp?incidente=4983282>. Acesso em: 15 mai. 2025.

BRASIL. Supremo Tribunal Federal. **Arguição de Descumprimento de Preceito Fundamental nº 403**. Relator: Min. Alexandre de Moraes, 27 set. 2017. Disponível em: <https://portal.stf.jus.br/processos/detalhe.asp?incidente=4975500>. Acesso em: 14 mai. 2025.

BRASIL. Supremo Tribunal Federal. **Liberdade de Expressão**. Brasília, 2020. Disponível em: <https://www.stf.jus.br/arquivo/cms/publicacaoPublicacaoTematica/anexo/LiberdadeExpressao.pdf>. Acesso em: 30 out. 2024

BRASIL. **Lei nº 12.965, de 23 de abril de 2014**. Estabelece princípios, garantias, direitos e deveres para o uso da internet no Brasil. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/l12965.htm](https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm). Acesso em: 17 mai. 2025.

BRASIL. **Projeto de Lei nº 2630, de 2020**. Institui medidas de combate à desinformação e à disseminação de conteúdos falsos em plataformas digitais. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=8110634&disposition=inline#:~:text=1%C2%BA%20Esta%20lei%20estabelece%20normas,Liberdade%2C%20Responsabilidade%20e%20Transpar%C3%Aancia%20na>. Disponível em: 13 mai. 2025.

CAMBI, Eduardo; ALVES, Elidia Aparecida de Andrade; ALVES, Fernando de Brito. Interfaces artificiais e interpretação judicial: O problema do uso da inteligência artificial e da metodologia fuzzy na aplicação do direito. **Revista dos Tribunais**, vol. 1010, p. 245-273, dez. 2019.

CAMPOS, Ana Claudia Sousa de; BITTENCOURT, Ila Barbosa. O princípio da liberdade de expressão na internet. **Revista de Direito Constitucional e Internacional**, vol. 117, p. 177-194, Jan.-Fev., 2020.

CASABONA, Carlos María Romeo. Dos delitos informáticos ao crime cibernético: Uma aproximação conceitual e político-criminal. **Doutrinas Essenciais de Direito Penal Econômico e da Empresa**, vol. 6, p. 509-552, jul. 2011.

DONEDA, Danilo. Inteligência artificial e direitos fundamentais: desafios éticos e jurídicos. São Paulo: **Revista dos Tribunais**, 2021.

EUBANKS, Virginia. Automating inequality: how high-tech tools profile, police, and punish the poor. New York: St. **Martin's Press**, 2018.

FARIA, Edimur Ferreira de; DAMASCENO, Luiza Mascarenhas. A indústria 4.0 e o futuro da prática jurídica no século XXI. **Revista dos Tribunais**, vol. 1003, p. 239-261, maio 2019.

FAUSTINO, André; SILVA, Luis Delcides R. O discurso do ódio contra o STF nas redes sociais e a teoria do amigo-inimigo de Carl Schmitt. **Revista de Direito e as Novas Tecnologias**, v. 18, p. 1-17, 2023.

FRANCO, Sofia Lima. O meio é a mensagem: os bots sociais e o seu papel na disseminação de conteúdo inverídico nas redes sociais. **Revista de Direito e as Novas Tecnologias**, v. 2, p. 1-20, 2019.

FRUMI, Patrícia. Marco Civil da Internet, provedores de informação e responsabilidade civil por cyberbullying. **Revista dos Tribunais**, vol. 1044, p. 145-167, out. 2022.

GODOI, Luca Felipe Sousa. Inteligência Artificial na gerência de ambientes digitais: O impacto de recomendação de conteúdo lesivo as crianças e adolescentes realizado por algoritmo nas redes sociais. **Revista de Direito e as Novas Tecnologias**, vol. 24, jul.-set. 2024.

GOUVÊA, Eduardo Mingorance de Freitas. Privacidade e internet: o direito de não ser exposto na rede. **Revista de Direito Privado**, v. 97, p. 19-44, 2019.

MATTEU, Ivelise Fonseca; CARDOSO, Leonardo da Silva; CARDOSO, Rafael Pinheiro de Matos. A liberdade de expressão e o discurso de ódio nas redes sociais. **Revista de Direito e as Novas Tecnologias**, vol. 19, abr.-jun. 2023.

MENDES, Laura Guimarães. **A liberdade de expressão e o discurso de ódio**. São Paulo: Saraiva, 2019.

MIRANDA, José Sérgio; DE MATTEU, Ivelise Fonseca. Fake news: um atentado às liberdades de pensamento e consciência. **Revista de Direito Constitucional e Internacional**, vol. 132, p. 325-344, Jul.-Ago., 2022.

NELSON, Rocco Antonio Rangel Rosso. Direito à liberdade de expressão como elementar para manutenção do Estado Democrático. **Revista de Direito Constitucional e Internacional**, vol. 138, p. 237-261, Jul.-Ago., 2023.

O'NEIL, Cathy. Weapons of math destruction: how big data increases inequality and threatens democracy. New York: **Crown Publishing Group**, 2016.

PARISER, Eli. **The filter bubble**: what the internet is hiding from you. New York: Penguin Press, 2011.

PIOVESAN, Flávia. **Direitos Humanos e o Discurso de Ódio**. São Paulo: Saraiva, 2016.

REALE JÚNIOR, Miguel. Limites à liberdade de expressão. **Revista Brasileira de Ciências Criminais**, vol. 81, p. 61-91, Nov.-Dez., 2009.

SALA, Vinicius Guarnieri. Redes sociais e minorias: Uma proposta de protagonismo na busca de um efetivo equilíbrio entre liberdade de expressão e direitos individuais protegidos. **Revista de Direito e as Novas Tecnologias**, vol. 10, jan.-mar. 2021

SARMENTO, Daniel. **Liberdade de Expressão versus Discurso de Ódio: Desafios Constitucionais**. São Paulo: Revista dos Tribunais, 2018.

SENNA, Pedro Henrique Scott de; BOTTON, Letícia Thomasi Jahnke. O confronto entre o direito à liberdade de expressão e o direito à honra diante do Marco Civil da Internet. **Revista dos Tribunais**, vol. 1014, p. 127-143, Abr., 2020.

SILVA, Rodrigo da Guia; SILVA, Marcela Guimarães Barbosa da. Perspectivas da eficácia horizontal de direitos fundamentais no ambiente virtual: contraditório e ampla defesa na exclusão de perfis e publicações por provedores de redes sociais virtuais. **Revista dos Tribunais**, v. 1049, p. 151-174, 2023.

UNIÃO EUROPEIA. **Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services (Digital Services Act)**. **Official Journal of the European Union**, Luxemburgo, 27 out. 2022.

WARREN, Zach. Encontrando “justiça” na IA: Como combater os vieses no processo de coleta de dados. **Boletim Revista dos Tribunais Online**, vol. 33, nov. 2022.