


**ENTRE A LIBERDADE DE EXPRESSÃO E O DISCURSO DE ÓDIO: LIMITES E IMPLICAÇÕES JURÍDICAS**

**BETWEEN FREEDOM OF EXPRESSION AND HATE SPEECH: LIMITS AND LEGAL IMPLICATIONS**

**ENTRE LA LIBERTAD DE EXPRESIÓN Y EL DISCURSO DE ODIO: LÍMITES E IMPLICACIONES LEGALES**

 <https://doi.org/10.56238/arev7n12-145>

**Data de submissão:** 12/11/2025

**Data de publicação:** 12/12/2025

**Ana Paula de Souza e Silva Simões**  
Mestra em Educação

Lattes: <https://lattes.cnpq.br/0445160300536827>

**RESUMO**

A expansão das redes sociais transformou as formas de comunicação e exercício da cidadania, criando desafios jurídicos complexos relacionados à delimitação entre liberdade de expressão e discurso de ódio. Este estudo analisa as implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio nas redes sociais, considerando princípios constitucionais, normas internacionais de direitos humanos e especificidades do ambiente digital. A escolha do tema justifica-se pela necessidade de compreender os limites constitucionais da liberdade de expressão em contextos nos quais manifestações discriminatórias violam a dignidade humana de grupos vulneráveis. O objetivo principal consiste em examinar os fundamentos jurídicos da liberdade de expressão, os critérios de caracterização do discurso de ódio e a responsabilidade das plataformas digitais na moderação de conteúdo. A metodologia adota abordagem qualitativa, de natureza aplicada, por meio de revisão sistemática da literatura jurídica, análise documental de legislação e jurisprudência, e análise comparada de experiências internacionais. Os resultados revelam que a tensão entre liberdade de expressão e discurso de ódio exige ponderação de valores constitucionais, considerando contexto, intencionalidade e potencial de danos. Observa-se que a regulação jurídica efetiva depende da construção de marcos regulatórios transparentes, da supervisão pública sobre plataformas digitais e da promoção de educação para cidadania digital.

**Palavras-chave:** Liberdade de Expressão. Discurso de Ódio. Redes Sociais. Direitos Fundamentais.

**ABSTRACT**

The expansion of social media has transformed forms of communication and citizenship exercise, creating complex legal challenges related to the delimitation between freedom of expression and hate speech. This study analyzes the legal implications of the tension between freedom of expression and hate speech in the social media environment, considering constitutional principles, international human rights norms, and specificities of the digital environment. The choice of theme is justified by the need to understand the constitutional limits of freedom of expression in contexts where discriminatory manifestations violate the human dignity of vulnerable groups. The main objective consists of examining the legal foundations of freedom of expression, the criteria for characterizing hate speech, and the responsibility of digital platforms in content moderation. The methodology adopts a qualitative approach, of applied nature, through systematic review of legal literature, documentary analysis of legislation and jurisprudence, and comparative analysis of international

experiences. The results reveal that the tension between freedom of expression and hate speech requires balancing constitutional values, considering context, intentionality, and potential for harm. It is concluded that effective legal regulation depends on the construction of transparent regulatory frameworks, public supervision over digital platforms, and the promotion of education for digital citizenship.

**Keywords:** Freedom of Expression. Hate Speech. Social Media. Fundamental Rights.

## RESUMEN

La expansión de las redes sociales ha transformado las formas de comunicación y el ejercicio de la ciudadanía, creando complejos desafíos legales relacionados con la delimitación entre la libertad de expresión y el discurso de odio. Este estudio analiza las implicaciones legales de la tensión entre la libertad de expresión y el discurso de odio en las redes sociales, considerando los principios constitucionales, las normas internacionales de derechos humanos y las especificidades del entorno digital. La elección del tema se justifica por la necesidad de comprender los límites constitucionales de la libertad de expresión en contextos donde las manifestaciones discriminatorias violan la dignidad humana de grupos vulnerables. El objetivo principal es examinar los fundamentos jurídicos de la libertad de expresión, los criterios para caracterizar el discurso de odio y la responsabilidad de las plataformas digitales en la moderación de contenido. La metodología adopta un enfoque cualitativo aplicado, mediante una revisión sistemática de la literatura jurídica, el análisis documental de la legislación y la jurisprudencia, y el análisis comparativo de experiencias internacionales. Los resultados revelan que la tensión entre la libertad de expresión y el discurso de odio requiere un equilibrio de valores constitucionales, considerando el contexto, la intencionalidad y el potencial de daño. Se observa que una regulación jurídica eficaz depende de la construcción de marcos regulatorios transparentes, la supervisión pública de las plataformas digitales y la promoción de la educación para la ciudadanía digital.

**Palabras clave:** Libertad de Expresión. Incitación al Odio. Redes Sociales. Derechos Fundamentales.

## 1 INTRODUÇÃO

A expansão das redes sociais nas últimas duas décadas transformou radicalmente as formas de comunicação, interação social e exercício da cidadania. Plataformas digitais como Facebook, Twitter, Instagram e TikTok tornaram-se espaços privilegiados para a manifestação de opiniões, debates políticos e construção de identidades coletivas. No entanto, essa democratização do acesso à expressão pública trouxe consigo desafios jurídicos complexos, especialmente no que se refere à delimitação entre o exercício legítimo da liberdade de expressão e a propagação de discursos de ódio. Até que ponto a proteção constitucional da liberdade de expressão pode acomodar manifestações que incitam a violência, promovem a discriminação ou violam a dignidade humana de grupos vulneráveis? A resposta a essa questão exige uma análise jurídica que articule princípios constitucionais, normas internacionais de direitos humanos e as especificidades do ambiente digital, reconhecendo que as redes sociais operam em uma lógica de viralização e amplificação que potencializa os danos causados por discursos violentos.

A liberdade de expressão constitui um dos pilares fundamentais das democracias contemporâneas, garantindo aos cidadãos o direito de manifestar opiniões, críticas e convicções sem interferência estatal arbitrária. No entanto, esse direito não é absoluto e encontra limites na proteção de outros direitos fundamentais, especialmente a dignidade da pessoa humana, a honra, a intimidade e a igualdade. Araújo, Saraiva e Godinho (2019, p. 5) afirmam que "a liberdade de expressão e a ponderação de valores exigem a tutela da dignidade da pessoa humana em face do *hate speech*, reconhecendo que a proteção constitucional não se estende a manifestações que promovam a discriminação ou a violência". Essa perspectiva evidencia que a interpretação jurídica da liberdade de expressão deve considerar os valores constitucionais e a necessidade de proteger grupos historicamente marginalizados contra discursos que perpetuam a opressão e a exclusão social.

O ambiente das redes sociais apresenta características específicas que amplificam os desafios jurídicos relacionados à regulação da liberdade de expressão. A velocidade de disseminação de conteúdos, a possibilidade de anonimato, a fragmentação das audiências e os algoritmos de recomendação criam condições propícias para a propagação de discursos de ódio e desinformação. Anjos *et al.* (2025, p. 3) destacam que "os limites da liberdade de expressão nas redes sociais demandam uma análise jurídica das *fake news*, reconhecendo que a desinformação compromete a qualidade do debate público e a legitimidade dos processos democráticos". Essa constatação reforça a necessidade de compreender as especificidades do ambiente digital e de desenvolver instrumentos jurídicos que equilibrem a proteção da liberdade de expressão com a responsabilização por conteúdos

ilícitos, sem comprometer a pluralidade de vozes e a diversidade de opiniões que caracterizam as sociedades democráticas.

A responsabilidade das plataformas digitais na moderação de conteúdos constitui uma dimensão central do debate jurídico contemporâneo. As empresas que controlam as redes sociais exercem poder significativo sobre o que pode ou não ser publicado, implementando políticas de autorregulação que nem sempre são transparentes ou submetidas a controle democrático. Balbino (2024, p. 15) argumenta que "a responsabilidade social corporativa na limitação da liberdade de expressão em redes sociais suscita questionamentos sobre a legalidade dos atos de controle da autorregulação empresarial". Essa perspectiva evidencia a tensão entre a autonomia privada das plataformas e a necessidade de garantir que as decisões sobre moderação de conteúdos respeitem princípios constitucionais e direitos fundamentais. A ausência de marcos regulatórios claros e de mecanismos de supervisão pública sobre as práticas de moderação de conteúdos representa um desafio para a proteção da liberdade de expressão e para o combate ao discurso de ódio.

A distinção entre liberdade de expressão e discurso de ódio não é sempre evidente, especialmente em contextos de polarização política e conflitos sociais. Barbosa, Disconzi e Torres (2021, p. 125) observam que "o discurso de ódio na internet representa a linha tênue entre o crime e a liberdade de expressão, exigindo critérios jurídicos precisos para a identificação de manifestações que ultrapassam os limites constitucionais". Essa dificuldade de delimitação conceitual reflete a complexidade das manifestações discursivas no ambiente digital, onde ironia, sátira, crítica política e incitação à violência podem se confundir. A construção de critérios jurídicos para a identificação do discurso de ódio deve considerar não apenas o conteúdo das manifestações, mas também o contexto de sua produção, a intencionalidade do emissor e o potencial de causar danos concretos a indivíduos ou grupos vulneráveis.

A relevância deste estudo reside na necessidade de produzir conhecimentos jurídicos que subsidiem a formulação de políticas públicas, a elaboração de marcos regulatórios e a atuação do sistema de justiça no enfrentamento dos desafios colocados pelas redes sociais. Em um contexto de crescente polarização política, aumento de crimes de ódio e disseminação de desinformação, torna-se imperativo compreender as implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio. Além disso, a análise das experiências internacionais de regulação das redes sociais e das decisões judiciais sobre casos emblemáticos pode fornecer subsídios para o aprimoramento do ordenamento jurídico brasileiro e para a construção de respostas institucionais mais efetivas e equilibradas.

O objetivo geral deste estudo consiste em analisar as implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio no ambiente das redes sociais, considerando os princípios constitucionais, as normas internacionais de direitos humanos e as especificidades do ambiente digital. Como objetivos específicos, propõe-se: (a) examinar os fundamentos jurídicos da liberdade de expressão e seus limites constitucionais; (b) identificar os critérios jurídicos para a caracterização do discurso de ódio no ambiente digital; (c) analisar a responsabilidade das plataformas digitais na moderação de conteúdos e os limites da autorregulação empresarial; (d) avaliar as experiências internacionais de regulação das redes sociais e suas implicações para o ordenamento jurídico brasileiro.

Este artigo estrutura-se em cinco seções principais. Após esta introdução, apresenta-se a fundamentação teórica, que aborda os fundamentos constitucionais da liberdade de expressão, os conceitos jurídicos de discurso de ódio e as especificidades do ambiente digital. Na sequência, descreve-se a metodologia adotada para a análise jurídica. A quarta seção apresenta os resultados e a discussão, articulando os achados com o referencial teórico e com as experiências internacionais. Por fim, as considerações finais sintetizam as principais conclusões do estudo, tencionam suas limitações e problematizam direções para pesquisas futuras. Ao longo de todo o texto, busca-se manter o rigor analítico e a clareza expositiva necessários para contribuir com o debate jurídico sobre a regulação da liberdade de expressão no ambiente das redes sociais.

## **2 FUNDAMENTAÇÃO TEÓRICA**

A compreensão das implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio no ambiente das redes sociais exige uma análise que articule os fundamentos constitucionais dos direitos fundamentais, as teorias democráticas contemporâneas e as especificidades do ambiente digital. A liberdade de expressão constitui um dos pilares das democracias liberais, garantindo aos cidadãos o direito de manifestar opiniões, críticas e convicções sem interferência estatal arbitrária. No entanto, esse direito não opera de forma isolada no ordenamento jurídico, mas em constante tensão com outros direitos fundamentais, especialmente a dignidade da pessoa humana, a honra, a intimidade e a igualdade. A teoria constitucional contemporânea reconhece que os direitos fundamentais não são absolutos e que sua aplicação exige ponderação de valores, considerando as circunstâncias concretas de cada caso e a hierarquia axiológica estabelecida pela Constituição.

A relação entre liberdade de expressão e democracia constitui um dos temas centrais da teoria política e jurídica contemporânea. Iensue (2023, p. 118) afirma que "as *fake news* representam uma ameaça à liberdade de expressão e à democracia, uma vez que comprometem a qualidade do debate

público e a capacidade dos cidadãos de formar opiniões informadas". Essa perspectiva evidencia que a proteção da liberdade de expressão não se justifica apenas pela garantia da autonomia individual, mas também pela sua função instrumental para o funcionamento das instituições democráticas. O debate público livre e plural constitui condição essencial para a formação da opinião pública, para a legitimidade das decisões políticas e para o controle social sobre os atos do poder público. No entanto, a disseminação de desinformação e de discursos de ódio compromete a qualidade do debate democrático, criando um paradoxo no qual a proteção irrestrita da liberdade de expressão pode, em determinadas circunstâncias, ameaçar os próprios fundamentos da democracia.

A educação para a cidadania e para a cultura política democrática constitui um elemento fundamental para a compreensão dos limites da liberdade de expressão. Machado (2024, p. 8) argumenta que "o Estado Democrático de Direito e a educação representam uma aproximação necessária e urgente para a implementação de uma autêntica cultura política, capaz de promover o respeito aos direitos fundamentais e a convivência democrática". Essa constatação reforça a perspectiva de que a regulação jurídica da liberdade de expressão não pode ser dissociada de políticas públicas de educação que promovam a formação de cidadãos críticos, capazes de distinguir informação de desinformação e de exercer a liberdade de expressão de forma responsável. A construção de uma cultura política democrática exige não apenas a garantia formal dos direitos fundamentais, mas também a criação de condições materiais e simbólicas para que os cidadãos possam exercer esses direitos de forma plena e consciente.

A dignidade da pessoa humana constitui o fundamento axiológico central do ordenamento jurídico brasileiro e o principal limite à liberdade de expressão. Macêda (2025, p. 5) destaca que "a liberdade de expressão e a dignidade da pessoa humana exigem análise acerca do racismo nas redes sociais, reconhecendo que manifestações discriminatórias violam o núcleo essencial dos direitos fundamentais". Essa perspectiva evidencia que a proteção constitucional da liberdade de expressão não se estende a manifestações que promovam a discriminação racial, a incitação à violência ou a violação da dignidade humana de grupos vulneráveis. A jurisprudência constitucional brasileira tem reconhecido que o discurso de ódio não está protegido pela liberdade de expressão, uma vez que sua finalidade não é contribuir para o debate público, mas sim promover a exclusão, a opressão e a violência contra determinados grupos sociais. A caracterização do discurso de ódio exige, portanto, a análise do conteúdo das manifestações, do contexto de sua produção e do potencial de causar danos concretos a indivíduos ou grupos vulneráveis.

A análise do discurso de ódio contra grupos religiosos oferece insights relevantes para a compreensão das tensões entre liberdade de expressão e proteção da dignidade humana. Maia e

Paixão (2020, p. 850) examinam a situação dos muçulmanos no Brasil, afirmando que "entre os direitos humanos e o discurso de ódio, as manifestações islamofóbicas evidenciam a necessidade de mecanismos jurídicos que protejam as minorias religiosas sem comprometer a liberdade de expressão". Essa análise demonstra que o discurso de ódio não se manifesta apenas em formas explícitas de incitação à violência, mas também em estereótipos, preconceitos e representações discriminatórias que contribuem para a marginalização de grupos vulneráveis. A proteção jurídica contra o discurso de ódio deve, portanto, considerar não apenas as manifestações mais evidentes de discriminação, mas também as formas sutis de violência simbólica que perpetuam a exclusão social e comprometem a igualdade de oportunidades.

A teoria dos direitos fundamentais oferece instrumentos conceituais relevantes para a compreensão da tensão entre liberdade de expressão e discurso de ódio. A técnica da ponderação de valores, desenvolvida pela doutrina constitucional alemã e incorporada pela jurisprudência brasileira, propõe que, em situações de conflito entre direitos fundamentais, o intérprete deve buscar a solução que maximize a proteção de ambos os direitos, considerando as circunstâncias concretas do caso. A aplicação dessa técnica exige a identificação dos direitos em conflito, a análise da intensidade da restrição imposta a cada direito e a avaliação da proporcionalidade da medida restritiva. No contexto das redes sociais, a ponderação entre liberdade de expressão e dignidade humana deve considerar as especificidades do ambiente digital, especialmente a velocidade de disseminação de conteúdos, a possibilidade de anonimato e o potencial de amplificação de discursos violentos.

A literatura especializada reconhece que o ambiente das redes sociais apresenta características que amplificam os desafios jurídicos relacionados à regulação da liberdade de expressão. A arquitetura das plataformas digitais, baseada em algoritmos de recomendação que priorizam conteúdos que geram engajamento, tende a favorecer a disseminação de manifestações polarizadas e emocionalmente carregadas, incluindo discursos de ódio. Além disso, a possibilidade de anonimato e a fragmentação das audiências criam condições propícias para a propagação de manifestações discriminatórias sem que os responsáveis sejam identificados ou responsabilizados. A compreensão dessas especificidades é essencial para a formulação de respostas jurídicas que equilibrem a proteção da liberdade de expressão com a responsabilização por conteúdos ilícitos.

A análise crítica da literatura revela que a tensão entre liberdade de expressão e discurso de ódio não pode ser resolvida por meio de soluções simplistas que privilegiem um direito em detrimento do outro. A proteção da liberdade de expressão é essencial para o funcionamento das democracias, mas não pode servir de escudo para a propagação de manifestações que violem a dignidade humana e promovam a discriminação. Por outro lado, a regulação excessiva da liberdade de expressão pode



comprometer a pluralidade de vozes e a diversidade de opiniões que caracterizam as sociedades democráticas. A construção de respostas jurídicas equilibradas exige, portanto, a articulação entre princípios constitucionais, normas internacionais de direitos humanos e as especificidades do ambiente digital, reconhecendo que a efetividade da proteção dos direitos fundamentais depende não apenas da regulação jurídica, mas também de políticas públicas de educação, de mecanismos de transparência e de controle democrático sobre as plataformas digitais. Este referencial teórico fornece, portanto, a base conceitual necessária para a análise empírica que será desenvolvida nas seções subsequentes deste artigo.

### 3 METODOLOGIA

Este estudo caracteriza-se como uma pesquisa de abordagem qualitativa, de natureza aplicada e com objetivos exploratórios e descritivos. A escolha pela abordagem qualitativa justifica-se pela necessidade de compreender em profundidade as dimensões jurídicas, sociais e éticas que envolvem a tensão entre liberdade de expressão e discurso de ódio no ambiente das redes sociais. A natureza aplicada da pesquisa fundamenta-se na intenção de produzir conhecimentos que possam subsidiar a formulação de marcos regulatórios, a atuação do sistema de justiça e o desenvolvimento de políticas públicas voltadas à proteção dos direitos fundamentais no ambiente digital. Quanto aos objetivos, a pesquisa assume caráter exploratório ao investigar um fenômeno relativamente recente e em constante transformação, e caráter descritivo ao mapear as principais perspectivas teóricas, jurisprudenciais e normativas sobre o tema.

A complexidade dos conflitos jurídicos no ambiente das redes sociais exige abordagens metodológicas que considerem a multiplicidade de atores, normas e valores envolvidos. Montenegro e Filho (2025) analisam os conflitos e dilemas no tribunal das redes sociais, destacando que a liberdade de expressão, a cultura do cancelamento e as *fake news* constituem fenômenos interconectados que demandam análises multidimensionais. Essa perspectiva orienta a escolha metodológica deste estudo, que busca capturar a complexidade das implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio por meio de uma análise que articula doutrina jurídica, jurisprudência, legislação nacional e internacional, e literatura científica especializada. A articulação dessas diferentes fontes permite uma compreensão abrangente do estado atual do debate jurídico e das tendências de regulação das redes sociais.

O procedimento metodológico adotado consiste em uma revisão sistemática da literatura jurídica, complementada pela análise documental de legislação, jurisprudência e documentos normativos internacionais. A revisão sistemática constitui um método de pesquisa que permite



identificar, selecionar e avaliar criticamente estudos relevantes sobre determinado tema, seguindo critérios explícitos e replicáveis. Neto e Rodrigues (2021) destacam que a busca por um modelo brasileiro de regulação da liberdade de expressão e do discurso de ódio exige a análise comparada de experiências internacionais e a compreensão das especificidades do ordenamento jurídico nacional. A escolha por esse método justifica-se pela necessidade de mapear o estado atual do conhecimento jurídico sobre o tema e de identificar os principais desafios teóricos e práticos enfrentados pelos sistemas jurídicos contemporâneos na regulação das redes sociais.

A coleta de dados foi realizada em quatro etapas complementares. A primeira etapa consistiu na busca sistemática de artigos científicos em bases de dados jurídicas e multidisciplinares, incluindo SciELO, Google Acadêmico, Portal de Periódicos CAPES e repositórios institucionais de universidades brasileiras. Os descritores utilizados foram: "liberdade de expressão", "discurso de ódio", "*hate speech*", "redes sociais", "dignidade da pessoa humana", "direitos fundamentais" e "regulação digital". Foram incluídos artigos publicados entre 2015 e 2025, em português, inglês ou espanhol, que abordassem as implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio. A segunda etapa envolveu a análise documental de legislação brasileira, incluindo a Constituição Federal de 1988, o Código Penal, a Lei nº 7.716/1989 (Lei do Racismo), o Marco Civil da Internet (Lei nº 12.965/2014) e projetos de lei em tramitação no Congresso Nacional. A terceira etapa consistiu na análise de jurisprudência do Supremo Tribunal Federal e do Superior Tribunal de Justiça, selecionando decisões emblemáticas sobre liberdade de expressão e discurso de ódio. A quarta etapa envolveu a análise de documentos normativos internacionais, incluindo tratados de direitos humanos, recomendações de organismos internacionais e legislações de países que desenvolveram marcos regulatórios para as redes sociais.

A análise dos dados foi conduzida por meio da técnica de análise de conteúdo jurídica, que permite identificar, categorizar e interpretar os principais argumentos, princípios e critérios presentes nos textos analisados. Novaes, Ventura e Rodrigues (2025) demonstram que a análise de discursos de ódio relacionados a questões de gênero exige sensibilidade para as dimensões simbólicas e estruturais da discriminação, reconhecendo que o discurso de ódio não se manifesta apenas em formas explícitas de violência verbal. Essa abordagem analítica foi incorporada neste estudo, que busca identificar não apenas os critérios jurídicos formais para a caracterização do discurso de ódio, mas também as dimensões contextuais, intencionais e consequenciais das manifestações discriminatórias. Os textos jurídicos e científicos foram lidos integralmente, e as informações relevantes foram organizadas em categorias temáticas relacionadas aos objetivos da pesquisa, incluindo: fundamentos constitucionais da liberdade de expressão, critérios de caracterização do discurso de ódio, responsabilidade das

plataformas digitais, experiências internacionais de regulação e desafios para o ordenamento jurídico brasileiro.

Os aspectos éticos constituem uma dimensão fundamental na análise jurídica da tensão entre liberdade de expressão e discurso de ódio. Petiz (2023, p. 8) argumenta que o direito à tolerância, ao meio ambiente de trabalho adequado e às restrições ao discurso de ódio como difamação de grupo exigem a consideração de princípios éticos que transcendem a mera legalidade formal. A consideração dos aspectos éticos neste estudo envolve o reconhecimento de que a regulação jurídica da liberdade de expressão não pode ser dissociada de valores fundamentais como a dignidade humana, a igualdade, a não discriminação e o respeito à diversidade. Embora esta pesquisa não envolva coleta de dados primários com seres humanos, a análise crítica das normas jurídicas e das decisões judiciais considera as implicações éticas das diferentes abordagens regulatórias, especialmente no que se refere à proteção de grupos vulneráveis e à garantia de condições equitativas de participação no debate público.

A análise comparada de diferentes perspectivas teóricas sobre a relação entre discurso de ódio e liberdade de expressão constitui um elemento metodológico relevante para a compreensão da complexidade do tema. Rocha e Lopes (2023) examinam as tensões conceituais e normativas entre discurso de ódio e liberdade de expressão, identificando que diferentes tradições jurídicas adotam critérios distintos para a delimitação dos limites da liberdade de expressão. Essa perspectiva comparada foi incorporada na metodologia deste estudo, que analisa não apenas o ordenamento jurídico brasileiro, mas também as experiências de países como Estados Unidos, Alemanha, França e Reino Unido, que desenvolveram modelos distintos de regulação das manifestações discriminatórias. A análise comparada permite identificar as vantagens e desvantagens de diferentes abordagens regulatórias e avaliar sua adequação ao contexto constitucional e social brasileiro.

As limitações metodológicas deste estudo incluem a dependência de fontes secundárias, que podem não capturar a totalidade das práticas de moderação de conteúdos nas redes sociais, e a restrição temporal da revisão sistemática, que abrange publicações dos últimos dez anos. Além disso, a análise qualitativa está sujeita à interpretação do pesquisador, o que exige rigor metodológico e transparência nos procedimentos analíticos. A ausência de dados empíricos sobre a efetividade das diferentes abordagens regulatórias constitui outra limitação, que pode ser superada em estudos futuros por meio de pesquisas quantitativas que avaliem o impacto das normas jurídicas sobre a incidência de discursos de ódio e sobre a qualidade do debate público. Apesar dessas limitações, a triangulação de fontes de dados, incluindo doutrina jurídica, jurisprudência, legislação e literatura científica, contribui para ampliar a validade e a confiabilidade dos resultados.

A metodologia apresentada nesta seção reflete o compromisso com o rigor científico e com a produção de conhecimentos que possam contribuir para o aprimoramento do ordenamento jurídico brasileiro. A articulação entre revisão sistemática, análise documental e análise comparada permite uma compreensão multidimensional das implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio, considerando suas dimensões constitucionais, normativas, jurisprudenciais e éticas. Os procedimentos metodológicos descritos garantem a transparência e a replicabilidade da pesquisa, elementos essenciais para a validação científica dos resultados e para a contribuição ao debate jurídico sobre a regulação da liberdade de expressão no ambiente das redes sociais.

Quadro 1 –Referências Acadêmicas e Suas Contribuições para a Pesquisa

Autor	Título	Ano	Contribuições
Vencato, A.	Gênero e sexualidades em tempos instáveis: mídias digitais, identificações e conflitos	2017	Discute identidades digitais e tensões sociais relacionadas ao discurso online.
Araújo, J.; Saraiva, M.; Godinho, A.	Liberdade de expressão e ponderação de valores	2019	Analisa o equilíbrio entre dignidade humana e discursos de ódio.
Maia, G.; Paixão, M.	Entre os direitos humanos e o discurso de ódio: muçulmanos no Brasil	2020	Explora desafios enfrentados por minorias diante de ataques discursivos.
Sarlet, I.; Siqueira, A.	Liberdade de expressão e seus limites numa democracia	2020	Debate limites democráticos frente às fake news eleitorais.
Silva, B.	Discurso de ódio nas normativas transnacionais	2020	Avalia políticas de autorregulação em plataformas digitais.
Barbosa, K.; Disconzi, V.; Torres, L.	Discurso de ódio na internet	2021	Aborda a fronteira entre crime e liberdade expressiva.
Neto, E.; Rodrigues, M.	Liberdade de expressão e discurso de ódio	2021	Propõe modelos jurídicos para regular discursos extremistas.
Iensue, G.	"Fake news", liberdade de expressão e democracia	2023	Relaciona desinformação ao enfraquecimento democrático.
Petiz, M.	Direito à tolerância e restrições ao discurso de ódio	2023	Defende limites jurídicos considerando proteção coletiva.
Rocha, C.; Lopes, E.	Discurso de ódio e liberdade de expressão	2023	Debate tensões entre expressão individual e danos sociais.
Santos, V.; Disconzi, V.	Limites da liberdade de expressão nas redes sociais	2023	Examina fronteiras legais para manifestações digitais.
Silva, I.; Silveira, J.	Liberdade de expressão x discurso de ódio	2023	Discute fronteiras entre crítica e agressão.
Silveira, N.; Moraes, R.; Lustosa, T.; Santos, J.	Proselitismo religioso e discurso de ódio	2023	Analisa conflitos entre fé e discursos intolerantes.
Balbino, M.	Responsabilidade social corporativa na limitação da liberdade de expressão	2024	Examina a legalidade da autorregulação empresarial.
Machado, M.	O estado democrático de direito e a educação	2024	Conecta educação e fortalecimento democrático.
Sarlet, I.	Liberdade de expressão e regulação do discurso do ódio nas mídias sociais	2024	Propõe parâmetros para regulação responsável.
Soares, E.; Azevedo, F.	A era da exposição	2024	Discute impactos das redes sociais nos direitos da personalidade.
Anjos, M.; Lima, A.; Dias, L.; Câmara, I.; Lima, O.	Os limites da liberdade de expressão nas redes sociais	2025	Avalia responsabilidades jurídicas sobre fake news.
Macêda, J.	Liberdade de expressão e dignidade da pessoa humana	2025	Examina o racismo digital e implicações legais.

Montenegro, V.; Filho, B.	Conflitos e dilemas no tribunal das redes sociais	2025	Debate cultura do cancelamento e suas implicações sociais.
Novaes, E.; Ventura, G.; Rodrigues, S.	"Queimem a bruxa!": Butler e aqueles que têm medo de gênero	2025	Analisa ataques discursivos com base em teorias de gênero.

Fonte: Elaboração do próprio autor

Essa organização cronológica permite visualizar a evolução do debate acadêmico sobre liberdade de expressão, fake news e discurso de ódio ao longo dos anos, evidenciando como as discussões se intensificaram especialmente a partir de 2020, com o aumento das tensões nas redes sociais. A estrutura facilita identificar lacunas temporais, mudanças de enfoque teórico e a progressiva sofisticação das análises jurídicas e sociais sobre o tema, tornando-se ferramenta essencial para pesquisas que buscam compreender transformações conceituais e regulatórias no ambiente digital brasileiro.

#### 4 RESULTADOS E DISCUSSÃO

A análise da literatura jurídica e dos documentos normativos revelou que a tensão entre liberdade de expressão e discurso de ódio no ambiente das redes sociais constitui um dos desafios mais complexos enfrentados pelos sistemas jurídicos contemporâneos. Os resultados obtidos neste estudo foram organizados em cinco dimensões analíticas: os limites constitucionais da liberdade de expressão, os critérios de caracterização do discurso de ódio, a regulação das *fake news* em períodos eleitorais, a autorregulação das plataformas digitais e o impacto das mídias sociais sobre os direitos da personalidade. A interpretação desses resultados à luz do referencial teórico permite compreender as diferentes abordagens adotadas pela doutrina e pela jurisprudência brasileiras, bem como identificar as lacunas normativas que demandam intervenção legislativa.

A delimitação dos limites da liberdade de expressão nas redes sociais constitui uma questão central para a compreensão das implicações jurídicas do discurso de ódio. Santos e Disconzi (2023) identificaram que os limites da liberdade de expressão nas redes sociais devem ser estabelecidos a partir da ponderação entre a proteção da manifestação do pensamento e a tutela da dignidade humana, reconhecendo que o ambiente digital amplifica os danos causados por manifestações discriminatórias. Esses achados corroboram a perspectiva teórica de que a liberdade de expressão não constitui um direito absoluto e que sua proteção encontra limites na salvaguarda de outros direitos fundamentais. A jurisprudência brasileira tem adotado a técnica da ponderação de valores para resolver conflitos entre liberdade de expressão e dignidade humana, considerando as circunstâncias concretas de cada caso e a proporcionalidade das restrições impostas à manifestação do pensamento.

O desafio da regulação do discurso de ódio nas mídias sociais exige a construção de critérios jurídicos precisos que permitam distinguir manifestações protegidas pela liberdade de expressão de discursos que ultrapassam os limites constitucionais. Sarlet (2024) analisou o desafio da regulação do discurso de ódio nas mídias sociais, argumentando que a caracterização do *hate speech* deve considerar não apenas o conteúdo das manifestações, mas também o contexto de sua produção, a intencionalidade do emissor e o potencial de causar danos concretos a grupos vulneráveis. Essa perspectiva evidencia que a regulação jurídica do discurso de ódio não pode se limitar a critérios formais, mas deve incorporar uma análise substantiva que considere as dimensões simbólicas e estruturais da discriminação. A doutrina constitucional brasileira tem reconhecido que o discurso de ódio caracteriza-se pela finalidade de promover a exclusão, a opressão ou a violência contra determinados grupos sociais, não contribuindo para o debate público democrático.

A regulação das *fake news* em períodos eleitorais representa uma dimensão específica da tensão entre liberdade de expressão e proteção da democracia. Sarlet e Siqueira (2020) examinaram a liberdade de expressão e seus limites em uma democracia, focalizando o caso das *fake news* nas redes sociais em período eleitoral no Brasil, e identificaram que a disseminação de desinformação compromete a integridade do processo eleitoral e a capacidade dos cidadãos de formar opiniões informadas. Esses resultados sugerem que a proteção da liberdade de expressão deve ser compatibilizada com a garantia da autenticidade do debate público, especialmente em contextos nos quais a desinformação pode influenciar decisivamente os resultados eleitorais. A legislação eleitoral brasileira tem incorporado mecanismos de combate às *fake news*, incluindo a possibilidade de remoção de conteúdos falsos e a responsabilização de candidatos e partidos políticos pela disseminação de desinformação.

A autorregulação das plataformas digitais constitui uma dimensão controversa da regulação da liberdade de expressão nas redes sociais. Silva (2020) analisou o discurso de ódio nas normativas transnacionais de empresas de mídias sociais, explorando as possibilidades da autorregulação regulada como modelo de governança das plataformas digitais. Os achados evidenciam que as empresas que controlam as redes sociais exercem poder significativo sobre o que pode ou não ser publicado, implementando políticas de moderação de conteúdos que nem sempre são transparentes ou submetidas a controle democrático. A ausência de marcos regulatórios claros e de mecanismos de supervisão pública sobre as práticas de moderação de conteúdos representa um desafio para a proteção da liberdade de expressão e para o combate ao discurso de ódio, uma vez que as decisões sobre remoção de conteúdos são tomadas por atores privados sem garantias processuais adequadas.

A distinção entre liberdade de expressão e discurso de ódio não é sempre evidente, especialmente em contextos de polarização política e conflitos sociais. Silva e Silveira (2023) investigaram o debate inadiável sobre liberdade de expressão versus discurso de ódio, questionando até que ponto determinadas manifestações constituem agressão ou expressão legítima. Essa dificuldade de delimitação conceitual reflete a complexidade das manifestações discursivas no ambiente digital, onde ironia, sátira, crítica política e incitação à violência podem se confundir. A construção de critérios jurídicos para a identificação do discurso de ódio deve considerar não apenas o conteúdo das manifestações, mas também o contexto de sua produção, a intencionalidade do emissor e o potencial de causar danos concretos a indivíduos ou grupos vulneráveis.

A análise de casos específicos de discurso de ódio oferece insights relevantes para a compreensão das tensões entre liberdade de expressão e proteção da dignidade humana. Silveira *et al.* (2023) examinaram o proselitismo religioso, questionando se constitui liberdade de expressão ou uma via de propagação do discurso de ódio, e identificaram que manifestações religiosas podem, em determinadas circunstâncias, ultrapassar os limites da liberdade de expressão quando promovem a discriminação ou a violência contra grupos vulneráveis. Essa análise demonstra que o discurso de ódio não se manifesta apenas em formas explícitas de incitação à violência, mas também em estereótipos, preconceitos e representações discriminatórias que contribuem para a marginalização de grupos vulneráveis.

O impacto das mídias sociais sobre os direitos da personalidade constitui uma dimensão relevante para a compreensão das implicações jurídicas da liberdade de expressão no ambiente digital. Soares e Azevedo (2024) analisaram a era da exposição e o impacto das mídias sociais nos direitos da personalidade, identificando que a arquitetura das plataformas digitais favorece a disseminação de conteúdos que violam a honra, a imagem e a intimidade de indivíduos. Esses achados evidenciam que a proteção dos direitos da personalidade no ambiente digital exige não apenas a responsabilização por conteúdos ilícitos, mas também a implementação de mecanismos de prevenção e de reparação de danos que considerem as especificidades do ambiente digital.

A análise de discursos de ódio relacionados a questões de gênero e sexualidade oferece insights relevantes para a compreensão das dimensões simbólicas e estruturais da discriminação. Vencato (2017, p. 808) examinou gênero e sexualidades em tempos instáveis, analisando mídias digitais, identificações e conflitos, e identificou que as redes sociais constituem espaços de disputa simbólica nos quais discursos de ódio contra pessoas LGBTQIA+ são disseminados de forma sistemática. Essa análise demonstra que o discurso de ódio não se manifesta apenas em ataques



individuais, mas também em campanhas organizadas que visam promover a exclusão e a violência contra grupos vulneráveis.

As limitações deste estudo incluem a dependência de fontes secundárias e a restrição temporal da revisão sistemática. Apesar dessas limitações, os resultados obtidos fornecem subsídios importantes para a compreensão das implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio no ambiente das redes sociais. As implicações desses resultados incluem a necessidade de fortalecer os marcos regulatórios, ampliar a transparência das práticas de moderação de conteúdos, implementar mecanismos de supervisão pública sobre as plataformas digitais e promover políticas públicas de educação para a cidadania digital que capacitem os cidadãos a exercer a liberdade de expressão de forma responsável e a identificar manifestações discriminatórias.

## **5 CONSIDERAÇÕES FINAIS**

Este estudo teve como objetivo analisar as implicações jurídicas da tensão entre liberdade de expressão e discurso de ódio no ambiente das redes sociais, considerando os princípios constitucionais, as normas internacionais de direitos humanos e as especificidades do ambiente digital. A investigação partiu da compreensão de que a expansão das plataformas digitais transformou radicalmente as formas de comunicação e exercício da cidadania, criando desafios jurídicos complexos relacionados à delimitação entre o exercício legítimo da liberdade de expressão e a propagação de discursos que violam a dignidade humana. Os resultados obtidos demonstram que a tensão entre liberdade de expressão e discurso de ódio não pode ser resolvida por meio de soluções simplistas que privilegiem um direito em detrimento do outro, mas exige a construção de respostas jurídicas equilibradas que considerem a ponderação de valores constitucionais, as especificidades do ambiente digital e a proteção de grupos vulneráveis. A síntese dos achados revela que, embora a doutrina e a jurisprudência brasileiras tenham avançado na delimitação dos limites da liberdade de expressão, persistem lacunas normativas relacionadas à regulação das plataformas digitais, à transparência dos processos de moderação de conteúdos e à responsabilização por manifestações discriminatórias.

As contribuições deste estudo para a área do Direito Constitucional e dos Direitos Humanos residem na produção de uma análise crítica e multidimensional das implicações jurídicas da liberdade de expressão no ambiente das redes sociais, articulando perspectivas doutrinárias, jurisprudenciais e normativas. A identificação dos critérios jurídicos para a caracterização do discurso de ódio, considerando não apenas o conteúdo das manifestações, mas também o contexto de sua produção, a intencionalidade do emissor e o potencial de causar danos concretos, fornece subsídios importantes



para a atuação do sistema de justiça e para a formulação de marcos regulatórios. Além disso, a análise da autorregulação das plataformas digitais e dos limites da autonomia privada das empresas que controlam as redes sociais destaca a necessidade de mecanismos de supervisão pública e de garantias processuais que assegurem a transparência e a *accountability* das decisões sobre moderação de conteúdos. A compreensão das especificidades do ambiente digital, especialmente a velocidade de disseminação de conteúdos, a possibilidade de anonimato e os algoritmos de recomendação, contribui para a identificação de caminhos para a construção de respostas jurídicas que equilibrem a proteção da liberdade de expressão com a responsabilização por conteúdos ilícitos.

As limitações desta pesquisa incluem a dependência de fontes secundárias, que podem não capturar a totalidade das práticas de moderação de conteúdos nas redes sociais, e a restrição temporal da revisão sistemática, que abrange publicações dos últimos dez anos. A ausência de dados empíricos sobre a efetividade das diferentes abordagens regulatórias e sobre o impacto das decisões judiciais na redução do discurso de ódio constitui outra limitação que pode ser superada em estudos futuros. Sugere-se que pesquisas subsequentes adotem abordagens metodológicas que incluam a coleta de dados primários, por meio de entrevistas com magistrados, advogados, representantes de plataformas digitais e membros de grupos vulneráveis, que permitam capturar as percepções e experiências dos atores envolvidos na regulação da liberdade de expressão. Além disso, estudos quantitativos que avaliem a incidência de discursos de ódio nas redes sociais, a efetividade dos mecanismos de denúncia e remoção de conteúdos, e o impacto das políticas de moderação sobre a qualidade do debate público podem fornecer insights adicionais sobre os desafios e as oportunidades para o aprimoramento da regulação jurídica. A realização de estudos comparados que analisem as experiências de diferentes países na regulação das redes sociais também constitui uma agenda de pesquisa relevante para a compreensão das vantagens e desvantagens de diferentes modelos regulatórios.

A reflexão final sobre o impacto deste trabalho destaca a relevância de produzir conhecimentos jurídicos que subsidiem a construção de marcos regulatórios democráticos, transparentes e efetivos para a proteção dos direitos fundamentais no ambiente digital. A tensão entre liberdade de expressão e discurso de ódio representa um dos desafios mais complexos enfrentados pelas democracias contemporâneas, exigindo respostas que equilibrem a proteção da pluralidade de vozes com a garantia da dignidade humana e da igualdade. A efetividade da regulação jurídica depende não apenas da formulação de normas adequadas, mas também da capacidade do sistema de justiça em aplicá-las de forma consistente, da transparência das práticas de moderação de conteúdos pelas plataformas digitais, e da promoção de políticas públicas de educação para a cidadania digital que capacitem os cidadãos a exercer a liberdade de expressão de forma responsável. A contribuição

deste estudo para o debate jurídico e político sobre a regulação das redes sociais reside na produção de uma análise crítica que reconhece a complexidade do tema, identifica as lacunas normativas existentes e aponta caminhos para a construção de respostas jurídicas que protejam simultaneamente a liberdade de expressão e a dignidade humana, fortalecendo os fundamentos democráticos da sociedade brasileira.

## REFERÊNCIAS

- Anjos, M.; Lima, A.; Dias, L.; Câmara, I.; Lima, O. Os limites da liberdade de expressão nas redes sociais: uma análise jurídica das fake news. *Iosr Journal of Business and Management*, v. 27, n. 6, p. 01-13, 2025. DOI: <https://doi.org/10.9790/487x-2706050113>
- Araújo, J.; Saraiva, M.; Godinho, A. Liberdade de expressão e ponderação de valores: tutela da dignidade da pessoa humana versus hate speech. *Revista da Faculdade de Direito*, v. 1, n. 40, 2019. DOI: <https://doi.org/10.22456/0104-6594.84601>
- Balbino, M. A responsabilidade social corporativa (RSC) na limitação da liberdade de expressão em redes sociais: a legalidade dos atos de controle da autorregulação empresarial. *Revista Brasileira de Políticas Públicas*, v. 13, n. 3, 2024. DOI: <https://doi.org/10.5102/rbpp.v13i3.8760>
- Barbosa, K.; Disconzi, V.; Torres, L. Discurso de ódio na internet: a linha tênue entre o crime e a liberdade de expressão. *Revista Científica Multidisciplinar Núcleo do Conhecimento*, p. 122-136, 2021. DOI: <https://doi.org/10.32749/nucleodoconhecimento.com.br/lei/liberdade-de-expressao>
- Brasil. Constituição da República Federativa do Brasil de 1988. Brasília, DF: Senado Federal, 1988.
- Iensue, G. “Fake news”, liberdade de expressão e democracia. *Revista do Direito Público*, v. 18, n. 3, p. 115-136, 2023. DOI: <https://doi.org/10.5433/1980-511x.2023v18n3p115>
- Machado, M. O estado democrático de direito e a educação: uma aproximação necessária e urgente para implementação de uma autêntica cultura política. *Revista Observatório*, v. 10, n. 1, a48, 2024. DOI: <https://doi.org/10.20873/uft.2447-4266.2024v10n1a48pt>
- Macêda, J. Liberdade de expressão e dignidade da pessoa humana: análise acerca do racismo nas redes sociais. *Lumen et Virtus*, v. 16, n. 53, e9491, 2025. DOI: <https://doi.org/10.56238/levv16n53-158>
- Maia, G.; Paixão, M. Entre os direitos humanos e o discurso de ódio: muçulmanos no Brasil. *Revista Quaestio Iuris*, v. 13, n. 2, 2020. DOI: <https://doi.org/10.12957/rqi.2020.40577>
- Montenegro, V.; Filho, B. Conflitos e dilemas no tribunal das redes sociais: liberdade de expressão, cultura do cancelamento e as fake news. *Revista Foco*, v. 18, n. 5, e8401, 2025. DOI: <https://doi.org/10.54751/revistafoco.v18n5-080>
- Neto, E.; Rodrigues, M. Liberdade de expressão e discurso de ódio: o direito brasileiro à procura de um modelo. *Espaço Jurídico Journal of Law*, v. 22, n. 2, p. 481-516, 2021. DOI: <https://doi.org/10.18593/ejll.29220>
- Novaes, E.; Ventura, G.; Rodrigues, S. “Queimem a bruxa!”: Butler e aqueles que têm medo de gênero. *Eccos – Revista Científica*, n. 72, e27919, 2025. DOI: <https://doi.org/10.5585/2025.27919>
- Petiz, M. Direito à tolerância, ao meio ambiente de trabalho adequado e a restrições ao discurso de ódio como difamação de grupo. *Revista de Ciências do Estado*, v. 8, n. 2, p. 1-30, 2023. DOI: <https://doi.org/10.35699/2525-8036.2023.42552>

Rocha, C.; Lopes, E. Discurso de ódio e liberdade de expressão. Revista Multidisciplinar do Nordeste Mineiro, v. 13, n. 1, 2023. DOI: <https://doi.org/10.61164/rmm.v13i1.1959>

Santos, V.; Disconzi, V. Limites da liberdade de expressão nas redes sociais. Revista Ibero-Americana de Humanidades Ciências e Educação, v. 9, n. 9, p. 88-100, 2023. DOI: <https://doi.org/10.51891/rease.v9i9.11083>

Sarlet, I. Liberdade de expressão e o desafio da regulação do discurso do ódio nas mídias sociais. REJEF, n. 1, p. 149-185, 2024. DOI: <https://doi.org/10.70982/rejef.v1i1.17>

Sarlet, I.; Siqueira, A. Liberdade de expressão e seus limites numa democracia: o caso das assim chamadas “fake news” nas redes sociais em período eleitoral no Brasil. Revista Estudos Institucionais, v. 6, n. 2, p. 534-578, 2020. DOI: <https://doi.org/10.21783/rei.v6i2.522>

Silva, B. Discurso de ódio nas normativas transnacionais de empresas de mídias sociais: uma abordagem acerca das possibilidades da autorregulação regulada. Brazilian Journal of International Relations, v. 9, n. 2, p. 405-433, 2020. DOI: <https://doi.org/10.36311/2237-7743.2020.v9n2.p405-433>

Silva, I.; Silveira, J. Liberdade de expressão x discurso de ódio: o debate inadiável, até que ponto não é agressão e sim expressão? 2023. DOI: <https://doi.org/10.18066/inic0897.23>

Silveira, N.; Moraes, R.; Lustosa, T.; Santos, J. Proselitismo religioso: liberdade de expressão ou uma via de propagação do discurso de ódio. Revista Contemporânea, v. 3, n. 12, p. 27558-27577, 2023. DOI: <https://doi.org/10.56083/rcv3n12-145>

Soares, E.; Azevedo, F. A era da exposição: o impacto das mídias sociais nos direitos da personalidade. Revista Ibero-Americana de Humanidades Ciências e Educação, v. 10, n. 11, p. 5516-5537, 2024. DOI: <https://doi.org/10.51891/rease.v10i11.16965>

Vencato, A. Gênero e sexualidades em tempos instáveis: mídias digitais, identificações e conflitos. ETD – Educação Temática Digital, v. 19, n. 4, p. 808, 2017. DOI: <https://doi.org/10.20396/etd.v19i4.8646384>