

**XXXII CONGRESSO NACIONAL DO  
CONPEDI SÃO PAULO - SP**

**DIREITOS E GARANTIAS FUNDAMENTAIS III**

**EUDES VITOR BEZERRA**

**VIVIANNE RIGOLDI**

**WILLIAM PAIVA MARQUES JÚNIOR**

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

#### **Diretoria - CONPEDI**

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

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

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

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

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

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

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

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

#### **Conselho Fiscal:**

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

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

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

Prof. Dr. Fernando Passos - UNIARA - São Paulo

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

#### **Secretarias**

##### **Relações Institucionais:**

Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná

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

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

##### **Comunicação:**

Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina

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

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

##### **Relações Internacionais para o Continente Americano:**

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

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

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

##### **Relações Internacionais para os demais Continentes:**

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

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

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

##### **Educação Jurídica**

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

Prof. Dr. Rubens Beçak - USP - SP

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

##### **Eventos:**

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

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

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

##### **Comissão Especial**

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

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

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

Prof. Dr. Rogério Borba - UNIFACVEST - SC

D597

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

Coordenadores: Eudes Vitor Bezerra, Vivianne Rigoldi, William Paiva Marques Júnior – Florianópolis: CONPEDI, 2025.

Inclui bibliografia

ISBN: 978-65-5274-298-8

Modo de acesso: [www.conpedi.org.br](http://www.conpedi.org.br) em publicações

Tema: Os Caminhos Da Internacionalização E O Futuro Do Direito

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

CDU: 34

# XXXII CONGRESSO NACIONAL DO CONPEDI SÃO PAULO - SP

## DIREITOS E GARANTIAS FUNDAMENTAIS III

---

### **Apresentação**

O conjunto de pesquisas que são apresentadas neste livro faz parte do Grupo de Trabalho de “DIREITOS E GARANTIAS FUNDAMENTAIS III”, ocorrido no âmbito do XXXII Congresso Nacional, realizado entre os dias 26, 27 e 28 de novembro de 2025 na Universidade Presbiteriana Mackenzie na Cidade de São Paulo, promovido pelo Conselho Nacional de Pesquisa e Pós-Graduação em Direito – CONPEDI e que teve como temática central “Os caminhos da internacionalização e o futuro do Direito”.

Os trabalhos expostos e debatidos abordaram de forma geral distintas temáticas atinentes DIREITOS E GARANTIAS FUNDAMENTAIS, especialmente relacionadas aos principais desafios que permeiam os problemas emergentes, bem como propondo soluções advindas de pesquisas em nível de pós-graduação, especialmente, Mestrado e Doutorado.

A defesa dos direitos e garantias fundamentais perante a sociedades é cada vez mais complexas e com enormes desafios, sobrealçam a importância das reflexões feitas, em todas as dimensões e direções, a permitir um melhor entendimento do mundo contemporâneo.

A partir da realidade contemporânea campeiam as reflexões extraídas da necessidade de um conhecimento aberto à necessidade de proteção dos direitos fundamentais, especialmente no que concerne a uma proteção jurídica articulada nos aspectos conseqüenciais das complexidades oriundas das colisões de direitos que podem vir a implicar em abusos, plasmando um ponto de mutação de uma lógica racional-cartesiana para uma realidade essencialmente relativista e aberta, ou seja, os direitos fundamentais deixam de ser apenas direitos de defesa do indivíduo contra a intromissão estatal em sua esfera privada, exurgindo daí a necessidade de reflexão em torno dos limites e possibilidades de sua efetividade.

Por mais que o CONPEDI se firme como um dos mais importantes eventos da pós-graduação brasileira em Direito, as problemáticas trabalhadas neste Grupo de Trabalho possuem uma amplitude trans e interdisciplinar, a fazer com que o Direito dialogue com importantes outras áreas do conhecimento humano.

Considerando todas as relevantes temáticas tratadas no presente livro, não pode ser outro senão o sentimento de satisfação que nós coordenadores temos ao apresentar a presente obra. É necessário, igualmente, agradecer aos pesquisadores que estiveram envolvidos tanto na

confecção dos artigos quanto nos excelentes debates proporcionados neste Grupo de Trabalho. Por fim, fica o reconhecimento ao CONPEDI pela organização e realização de mais um relevante evento.

A expectativa é de que esta obra possa contribuir com a compreensão dos problemas do cenário contemporâneo, com o a esperança de que as leituras dessas pesquisas ajudem na reflexão do atual caminhar do DIREITOS E GARANTIAS FUNDAMENTAIS.

Foi uma tarde de exitosas discussões e de engrandecimento da pesquisa na área dos Direitos Fundamentais e que, agora, pretendemos compartilhar com a comunidade acadêmica.

As pesquisas ora apresentadas funcionam como canais indispensáveis nos debates e propostas das pendências existentes nos campos indicados e na busca de soluções efetivas para as problemáticas indicadas. Reiteramos a esperança que a obra ora apresentada sirva como parâmetro acadêmico para a compreensão dos problemas da complexa realidade mundial sob a óptica dos Direitos Fundamentais.

Desejamos leituras proveitosas na construção de uma nova perspectiva para os desafios impostos aos Direitos Fundamentais no contexto contemporâneo.

São Paulo, novembro de 2025.

Prof. Dr. Eudes Vitor Bezerra (PPGDIR – UFMA)

Prof. Dr. Vivianne Rigoldi (PPGD - Centro Universitário Eurípides de Marília)

Prof. Dr. William Paiva Marques Júnior (Universidade Federal do Ceará- UFC)

# MODERAÇÃO ALGORÍTMICA E LIBERDADE DE EXPRESSÃO: OS RISCOS DE SILENCIAMENTO DAS VOZES FEMININAS NA ERA DA INTELIGÊNCIA ARTIFICIAL

## ALGORITHMIC MODERATION AND FREEDOM OF EXPRESSION: THE RISKS OF SILENCING WOMEN'S VOICES IN THE AGE OF ARTIFICIAL INTELLIGENCE

Lucas Gonçalves da Silva <sup>1</sup>  
Isadora Inês Alves Correia <sup>2</sup>

### Resumo

O presente artigo analisa como o processo de moderação algorítmica está entrelaçado a liberdade de expressão, que é um direito fundamental elencado na Constituição da República Federativa do Brasil, destacando quais os cuidados devem ser tomados para que este direito seja respeitado, em face de um problema existente na estrutura social, que é o silenciamento das vozes femininas, como maneira de invalidar alguns assuntos importantes. Neste viés, o trabalho parte do pressuposto que existem riscos causados pela moderação algorítmica, já que esta molda e direciona comportamentos nos ambientes virtuais, onde muitas das vezes utiliza-se de critérios enviesados, que trazem consigo estruturas que acabam atingindo determinados grupos. Desta forma, é necessário que as plataformas invistam em uma revisão humana imparcial do conteúdo, especialmente quando se trata de temas sensíveis, como é o caso do direito das mulheres. Diante disso, o objetivo específico deste trabalho é demonstrar como é realizada a moderação algorítmica nas plataformas digitais, com foco na inteligência artificial, destacando os riscos que podem atingir os direitos fundamentais. A metodologia utilizada neste artigo foi a qualitativa, baseada em revisão bibliográfica, análise de legislações e estudos comparados, com abordagem crítica sobre as lacunas legislativas existentes. Através desta investigação, pretende-se demonstrar que os problemas sociais também existem no mundo digital, e que precisam de um olhar estruturante. O estudo reforça a importância do combate da perpetuação de narrativas preconceituosas; sendo necessário que se desenvolva meios para garantir os direitos fundamentais, inclusive, no meio digital.

**Palavras-chave:** Moderação algorítmica, Inteligência artificial, Direitos fundamentais, Liberdade de expressão, Silenciamento das vozes femininas

### Abstract/Resumen/Résumé

This article analyzes how the algorithmic moderation process is intertwined with freedom of expression, a fundamental right enshrined in the Constitution of the Federative Republic of

---

<sup>1</sup> Pós-doutor em Direito pela Università Degli Studi G. d'Annunzio (Italia) e pela Universidade Federal da Bahia. Doutor e Mestre em Direito do Estado pela Pontifícia Universidade Católica de São Paulo.

<sup>2</sup> Bacharela em Direito pela Universidade Tiradentes, Pós-Graduada em Direito Público pela Legale, Mestranda em Direito na Universidade Federal de Sergipe, Bolsista Capes.

Brazil. It highlights the precautions that must be taken to ensure this right is respected, given a problem inherent in social structure: the silencing of women's voices as a way to invalidate certain important issues. From this perspective, the work assumes that algorithmic moderation poses risks, as it shapes and directs behavior in virtual environments, where biased criteria are often used, resulting in structures that ultimately affect certain groups. Therefore, platforms must invest in impartial human review of content, especially when it concerns sensitive topics such as women's rights. Therefore, the specific objective of this work is to demonstrate how algorithmic moderation is carried out on digital platforms, with a focus on artificial intelligence, highlighting the risks that can affect fundamental rights. The methodology used in this article was qualitative, based on a literature review, legislative analysis, and comparative studies, with a critical approach to existing legislative gaps. Through this research, we aim to demonstrate that social problems also exist in the digital world and require a structured approach. The study reinforces the importance of combating the perpetuation of prejudiced narratives; it is necessary to develop means to guarantee fundamental rights, including in the digital environment.

**Keywords/Palabras-claves/Mots-clés:** Algorithmic moderation, Artificial intelligence, Fundamental rights, Freedom of expression, Silencing of women's voices

## 1 INTRODUÇÃO

O artigo explora sobre a relação entre a moderação algorítmica e a liberdade de expressão, que é um direito fundamental previsto na Constituição Federal, trazendo os riscos que podem existir no gerenciamento de conteúdo na plataforma digital, demonstrando que há um viés algorítmico, onde os algoritmos são treinados em dados com vieses, desenvolvendo decisões e ideias muita das vezes discriminatórias.

Através da metodologia qualitativa embasada na pesquisa bibliográfica, o objetivo principal do estudo é analisar os impactos da moderação algorítmica de conteúdo nas plataformas digitais, sobre a liberdade de expressão, destacando os riscos de silenciamento das vozes femininas e suas implicações para a proteção dos direitos fundamentais.

Assim, como objetivos específicos, pretende-se (i) Examinar o conceito de liberdade de expressão como direito fundamental previsto na Constituição Federal e sua relação com a igualdade de gênero, (ii) Investigar como funciona a moderação algorítmica de conteúdo, especialmente com o uso da inteligência artificial, e seus critérios, (iii) Identificar os riscos de viés algorítmico e de silenciamento das mulheres, (iv) Analisar as lacunas existentes na legislação brasileira; o Marco Civil da Internet (MCI), Lei Geral de Produção de Dados (LGPD) e projetos em tramitação, (v) Comparar experiências internacionais de regulação e governança digital que possam contribuir para soluções no contexto brasileiro, (vi) Propor medidas de governança digital e políticas públicas capazes de assegurar uma moderação de conteúdo mais transparente, imparcial e sensível a questão de gênero.

A partir de uma perspectiva crítica, pretende-se compreender de que maneira é realizada a moderação algorítmica, quais são as diretrizes utilizadas para as decisões e validações realizadas, e de que maneira tais pressupostos se apresentam como riscos em ações que acabam afetando os direitos fundamentais das mulheres.

Dessa forma, o artigo busca oferecer medidas para que esta moderação algorítmica não seja enviesada com padrões patriarcais e preconceituosos, perpetuando estereótipos sociais que atingem as mulheres em grande escala, destacando quais medidas podem ser tomadas para que haja uma maior imparcialidade nas plataformas digitais, com destaque para a inteligência artificial.

Com o passar do tempo, a sociedade fora inserida em um mundo cada vez mais tecnológico, com o desenvolvimento das novas tecnologias, e abrangência no uso da

Inteligência Artificial; o que gerou novos cenários e o aprofundamento de questões sociais, se tornando emergente focar nas nuances e nos problemas neste novo cenário, o que alerta para a necessidade de mudança de paradigmas e atualizações nas leis.

O que resultou na formulação da seguinte questão: Como garantir que a moderação algorítmica seja neutra de estereótipos para que haja a efetivação do direito de liberdade de expressão das mulheres?

Responder a esta pergunta, requer uma análise pragmática do funcionamento da moderação algorítmica e quais os pressupostos utilizados em sua criação e desenvolvimento, fazendo uma investigação de como o direito de liberdade de expressão das mulheres está sendo atingido através do uso da inteligência artificial desenvolvida por vieses discriminatórios.

## **2 DIREITOS FUNDAMENTAIS E LIBERDADE DE EXPRESSÃO NO ORDENAMENTO JURÍDICO BRASILEIRO**

A Constituição Federal é o ponto central quando se fala em um caminho para oferecer respostas nessa nova era digital, capazes de sustentar o Estado Democrático de Direito. Assim, em relação ao problema da moderação algorítmica e a perpetuação de discriminações, se torna importante propor mecanismos regulatórios que estejam dentro dos limites constitucionais, e que garantam ao mesmo tempo que os direitos fundamentais, como a liberdade de expressão, sejam garantidos. Existem limites constitucionais quando falamos da aplicabilidade dos direitos fundamentais, como exemplo, a liberdade de expressão encontra limites na dignidade da pessoa humana.

O que é aceitável e o que é repreensível enquanto expressão também têm estreita relação com a jurisprudência do país em análise. É inegável o papel dos tribunais, sobretudo dos tribunais constitucionais, na definição dos contornos dos direitos fundamentais, bem como em sua ponderação, em um Estado Democrático de Direito (Barroso, 2024, p. 206).

Nesse sentido, é imprescindível que o Direito Constitucional seja entrelaçado com as novas tecnologias, observando os limites constitucionais, no sentido de não prejudicar direitos, mas ao mesmo tempo, respeitar e assegurar as manifestações individuais, de maneira eficaz, fortalecendo as características de uma sociedade plena e democrática.

Os direitos fundamentais e as liberdades comunicativas são ameaçados pela atuação das companhias digitais e mecanismos de inteligência artificial, que acabam limitando a atuação de alguns grupos nos meios tecnológicos, como é o caso das mulheres.

Como aduz Silva (2020, p.320), o ser humano personifica suas próprias ações no mundo virtual:

Na vida on-line, o excesso de possível de vivências subjetivas está potencializado ao máximo. É possível realizar qualquer esfera do desejo, qualquer princípio volitivo humano, independentemente de seus fins, quer seja para o bem ou para o mal. A ação humana é compreendida no âmbito de uma reflexão sobre um poder-fazer que não conhece limites, circunscrita por uma infinitude de exigências, o que nos coloca diante de uma vida marcada pela finitude de uma existência e a infinitude de uma exigência.

Dessa forma, é importante repensar sobre a eficácia dos direitos fundamentais neste panorama, afirmando a incidência destes nos meios digitais, proporcionando que a liberdade de expressão seja garantida, mas que ao mesmo tempo seja amparada com limites constitucionais. O ambiente digital acaba estruturando uma nova dimensão destes direitos, já que muito dos seus cenários são mediados por algoritmos, o que requer mudanças.

### **3 OS PRESSUPOSTOS DA MODERAÇÃO ALGORÍTMICA E A INTELIGÊNCIA ARTIFICIAL**

A inteligência artificial (IA) tornou-se um dos temas mais vivenciados atualmente na sociedade, especialmente na ampla disponibilização e facilidade de acesso das suas ferramentas, que é estruturada pela moderação algorítmica.

É importante ressaltar que a IA serve como uma verdadeira extensão de panoramas e sistemas discriminatórios, que operam e se interconectam; vida real com a vida virtual. A moderação tem a ver com a seleção das informações que serão propagadas, onde a intervenção e a escola sobre o conteúdo ali perpetuado cabe as plataformas digitais.

Nesse sentido, uma das formas que a IA perpetua a violência é na repetição e reforço de estereótipos, o que acontece com as mulheres, que são vinculadas as características negativas, com a sobreposição do patriarcado, utilizando-se de vieses discriminatórios e sexistas.

Neste cenário, Tarleton Gillespie et al. fornecem um conceito técnico para o tipo de moderação nas plataformas digitais:

Moderação de conteúdo – a detecção, avaliação e intervenções realizadas sobre conteúdo ou comportamento considerados inaceitáveis por plataformas ou outros intermediários de informação, incluindo as regras que impõem, o trabalho humano e as tecnologias necessárias, bem como os mecanismos institucionais de adjudicação, execução e recurso que a sustentam (Gillespie et al., 2020, tradução nossa).

Assim, ideias e discursos discriminatórios e misóginos são desenvolvidos nas plataformas digitais, sendo regras recriadas do mundo real, onde no mundo digital, traz consigo seus pressupostos e características enraizadas.

A moderação, ainda segundo Gillespie, é a característica que define as plataformas digitais:

Moderação não é um aspecto auxiliar do que as plataformas fazem. É essencial, constitucional, definidor. As plataformas não só não podem sobreviver sem moderação, como não são plataformas sem ela. A moderação está lá desde o início e sempre; no entanto, deve ser amplamente rejeitada, escondida, em parte para manter a ilusão de uma plataforma aberta e em parte para evitar a responsabilidade legal e cultural. As plataformas enfrentam o que pode ser uma contradição irreconciliável: elas são representadas como meros condutores [de conteúdo] e têm como premissa fazer escolhas sobre o que os usuários veem e dizem (Gillespie, 2018, p. 21, tradução nossa).

Moderação de conteúdo – a detecção, avaliação e intervenções realizadas sobre conteúdo ou comportamento considerados inaceitáveis por plataformas ou outros intermediários de informação, incluindo as regras que impõem, o trabalho humano e as tecnologias necessárias, bem como os mecanismos institucionais de adjudicação, execução e recurso que a sustentam (Gillespie et al., 2020, tradução nossa).

Faz-se necessário entender como funciona a moderação algorítmica, e os mecanismos de desenvolvimento destes nos espaços virtuais, e de que maneira os algoritmos regulam o que pode ou não ser dito, se propagando de maneira muito rápida nos ambientes virtuais. Dessa forma, para compreender os novos cenários da era digital, se torna essencial conhecer suas principais técnicas, bem como seu desenvolvimento.

A moderação algorítmica passa por diversos ditames, tendo em vista que as regras e leis brasileiras eventualmente são criadas pelo Poder Legislativo, e sua aplicabilidade e interpretação, pelo Poder Judiciário. Assim, existem lacunas legislativas que não são preenchidas pela omissão do Poder Legislativo, que não coloca urgência nas pautas que versam sobre o ambiente digital, o que perpetua, muita das vezes, em uma ineficiência na fiscalização e nos julgamentos.

Mais do que avaliar o que será ou não transmitido e mantido nas plataformas digitais, a moderação de conteúdo pode ser compreendida como uma das espécies de seleção algorítmica, no sentido de orientar-se por um procedimento automatizado que atribui relevância a certos fragmentos de informações (Just; Latzer, 2016).

Embora a moderação de conteúdo não seja uma atividade recente, as suas diversas formas de operacionalização na modalidade digital têm se revelado cada vez mais inovadoras, considerando suas significativas implicações sociais, econômicas, concorrenciais, além de seus benefícios e riscos que demandam um tratamento jurídico adequado e desafiam as discussões regulatórias sobre o tema da moderação de conteúdo (Just; Lazer, 2016).

Nesta linha, Evelyn Douek, professora da Faculdade de Direito de Stanford, explica que a moderação não é mais apenas um agregado de decisões binárias (remover ou não remover um conteúdo), mas sim “um vasto sistema de administração que inclui um conjunto mais amplo de decisões e tomadores de decisões do que uma imagem padrão admite” (Douek, 2022, p. 531).

O mesmo pode ser verificado no caso da moderação de conteúdo digital, que apoia-se cada vez mais em operações algorítmicas, inteligência artificial, aprendizagem de máquina e sistemas de redes neurais (Sheng, 2022).

Em linhas gerais, a atividade moderadora pode ser executada de duas formas: a) matching (correspondência), em que o sistema realiza uma comparação entre a nova informação obtida e a sua base de dados já existente e b) prediction (predição) na qual o aprendizado anterior da máquina é responsável pela identificação e análise dos elementos da nova informação (Sheng, 2022).

A técnica do matching baseia-se no método conhecido como hashing, apto a transformar conteúdos como imagens e vídeos em uma cadeia única de dados (hash) que recebem um rótulo, uma espécie de “identidade digital”. Com isso, o sistema automatizado

consegue realizar rápidas comparações entre o valor atribuído ao hash e o valor atribuído à base de dados, por sua vez, composta por outros hashes de itens indesejados, como pornografia e violação aos direitos autorais (Sheng, 2022).

No caso do hashing é imprescindível ainda apresentar suas duas espécies, pois cada uma delas pode entregar resultados com eficácias diferentes, a depender do objetivo que se pretende alcançar. O hashing criptográfico utiliza um hash de alta sensibilidade a alterações, o que lhe permite autenticar uma informação de forma satisfatória. Por outro lado, ele não opera bem na moderação de conteúdo, sendo facilmente driblado, já que não consegue identificar pequenas alterações feitas em um conteúdo, como a inserção de marcas d'água, por exemplo (Sheng, 2022).

Já o hashing perceptivo atua em busca de encontrar similaridades entre a nova informação inserida no sistema automatizado e a base de dados existente. Portanto, ele foca em características específicas do conteúdo recebido para determinar o limiar de diferença permitido entre dois hashes (Sheng, 2022).

A título de exemplo, diante de uma mesma foto, colocada em duas imagens diferentes, com níveis diversos de saturação da cor, o hashing criptográfico tratará as duas imagens como completamente diferentes, enquanto o hashing perceptivo compreenderá melhor a relação entre as imagens e lhes atribuirá o mesmo rótulo (Sheng, 2022).

Deste modo, o hashing perceptivo é o método mais adequado para a atividade de moderação, pois consegue permanecer imune às eventuais alterações de um conteúdo, captando com maior acurácia as relações entre imagens e informações. Outra vantagem do hashing perceptivo é a sua maior transparência em relação a outras ferramentas moderadoras de inteligência artificial, o que contribui para uma maior auditabilidade dos conteúdos armazenados em qualquer base de dados de hash (Farid, 2021).

Dessa forma, é importante ressaltar que a moderação de conteúdo desenvolvida pela inteligência artificial perpetua as formas de discriminação, esta não possui uma força neutralizada, mas sim um programa estruturado por algoritmos e linhas de código, estes formados por humanos, reproduzindo assim, os vieses de seu idealizador.

Além disso, a IA em presta envergadura a novas dinâmicas, como a “Internet das Coisas” (Internet of Thing – IOT) e Smart Cities. Porém, sua utilização mais difundida se encontra no desenvolvimento de sistemas de análise e tomada de decisão baseados em

algoritmos, fornecendo subsídios para novas formas de monitoramento e pesquisa sobre as condições de vida e o controle do comportamento (Hoffman-Riem, 2021).

A racionalidade dos sistemas de IA também estão interligadas aos algoritmos de aprendizagem e como são direcionados, (Acioly, 2022, p. 172):

[...] desponta em celeuma ético-jurídica ainda mais profunda, notadamente quando se parte da noção de que é da própria essência dos algoritmos a utilização de correlações pautadas na estatística, bem como da noção de que tais mecanismos artificiais manifestam independência no próprio processo de aprendizagem – a deep learning.

A inteligência artificial se desenvolve no tempo, trazendo também novas complexidades, como as linhas de código utilizadas pela inteligência artificial, que decide quem é silenciado no mundo digital, e como esta ação ocorre. Sendo assim, é evidente os problemas éticos e estruturais nas bases da inteligência artificial, atingindo os direitos fundamentais constitucionais, tendo em vista que ocorre uma verdadeira seletividade virtual.

Nesse sentido, existe um entrelaçamento entre a liberdade de expressão e a moderação algorítmica, sendo necessário que a inteligência artificial utilize da transparência nas decisões tomadas por meio da moderação.

#### **4 RISCOS DA MODERAÇÃO ALGORÍTMICA PARA O DIREITO DAS MULHERES**

Atualmente, a violência contra as mulheres vem se configurando como um fenômeno global, envolvendo várias nuances, e com isso, novos cenários tecnológicos, o que reforça a permanência da desigualdade social que acompanha as mulheres historicamente, dentro de uma estrutura de desigualdade e discriminação. O desenvolvimento das tecnologias digitais, em especial a inteligência artificial, vem trazendo novos tipos de violência, como a violência digital, assim, impondo novos desafios para combatê-la.

Nesse sentido, tal violência deve ser compreendida de uma forma mais ampla, prestando atenção nas lacunas legislativas existentes, frente aos impactos da IA na atualidade, que se reverbera de diferentes maneiras.

Em uma sociedade da informação, a subjetivação algorítmica engendra-se por balizas reais de classificação e de validação de acessos, preconizadas na aplicação de tecnologia de Inteligência Artificial (Ferreira, 2021).

É inevitável perceber que a sociedade está cada vez mais interconectada, onde através das plataformas digitais são disseminados conteúdos que perpetuam discriminações, impactando diretamente na concretização dos direitos de forma plena.

Ao longo dos anos, as mulheres vêm vivenciando exclusões e cerceamento de seus direitos. Vive-se, hoje, uma Era de contingência das minorias, em que os grupos historicamente silenciados e subjugados a uma condição marginal pleiteiam o acesso e o usufruto de direitos sociais, políticos e econômicos (Carmo, 2016).

O desenvolvimento dos meios digitais necessita cada vez mais de um equilíbrio dos interesses sociais e econômicos, onde a moderação de conteúdo digital, se torna uma tarefa desafiadora, sendo necessário que sempre hajam adaptações em razão do surgimento de novos cenários.

Nesse ínterim, é perceptível a continuidade das estruturas e das relações sociais que se propagam ao longo dos anos; uma sociedade consolidada por vieses discriminatórios e bases patriarcais, onde formas de violência ainda se perpetuam, agora no meio digital.

Diante disso, torna-se medida a se tomar, o conhecimento amplo sobre a operação e a moderação de conteúdo; com seus elementos, riscos e vieses; tendo em vista que mesmo que certas ações ocorram dentro do cenário tecnológico e virtual, as consequências se tornam reais.

A moderação de conteúdo tem como uma de suas protagonistas principais a inteligência artificial, que são criadas e desenvolvidas por seres humanos, com predominância, de homens. Para que haja a propagação da inteligência artificial, são utilizadas linhas de código, onde estas são reproduzidas no meio digital, e assim os vieses discriminatórios, existentes no mundo offline, e que acabam se reverberando no mundo online.

No ambiente virtual reverberam diversas formas de violência de gênero, o que faz com que ofereça aos algoritmos, subsídios para o seu funcionamento, trazendo esta carga discriminatória, projetadas as formas de violência de forma ampla e exponencial. Nesse sentido, conforme (Cezarino e Contri, 2022, p.61):

A tecnologia não detecta algo que está dado no mundo, de forma natural e apriorística; pelo contrário: inevitavelmente, ela processa dados por meio da construção de uma suposição quanto ao gênero, a partir das dimensões fisiológicas das pessoas sobre as quais é aplicada. Para realizar esse tipo de inferência, a tecnologia é construída de forma a apreender se a pessoa pertence

ao gênero masculino ou feminino, partindo do pressuposto de que o gênero é meramente binário e uma continuidade da característica fisiológica.

Assim, perceber e reconhecer a violência de gênero no ambiente virtual, requer observá-la de maneira sistêmica; de como ela é construída, desenvolvida e perpetrada. A moderação dos algoritmos possui toda uma estrutura ao qual deve se observar, para assim se chegar na compreensão de todo o processo, fiscalizando os vieses que estruturam a violência de gênero em todo este sistema.

Se mostra necessário, portanto, reestruturar a noção de implementação social de algoritmos, para que passe a ser pautada nas diretrizes éticas de beneficência, não-maleficência e autonomia, atreladas aos princípios de justiça, acurácia e inteligibilidade (Mulholland, Frajhof, 2021).

A legitimação do uso de algoritmos, no contexto da proteção do livre desenvolvimento da personalidade em âmbito virtual, não prescinde da concretização de uma estrutura dialógica, na qual a inteligibilidade é apontada de forma ativa, em que informações claras e precisas ao usuário performam a própria estrutura, ressaltando a sua importância na construção da metodologia algorítmica (Acioly, 2022).

Diante deste cenário, é visto que discursos sobre direitos das mulheres, feminismo, reprodução, são atingidos em sua liberdade de expressão, onde muita das vezes são assuntos censurados por meio dos algoritmos. A lógica algorítmica, ao delimitar conteúdos como “sensíveis” ou “inapropriados”, acaba inviabilizando debates fundamentais, fragilizando o direito a liberdade de expressão, e diminuindo a pluralidade de vozes, restringindo o espaço democrático no ambiente digital.

Assim, a liberdade de expressão, assegurada pela Constituição Federal, deve ser efetivada, não sendo restringida de forma arbitrária por sistemas tecnológicos não transparentes, o que evidencia a responsabilidade pelas plataformas digitais, na garantia de uma de um ambiente digital inclusivo e sem discriminações. Tendo em vista que o impacto de tal ação afeta todo o processo coletivo de luta pela liberdade de expressão das mulheres, bem como a construção da conscientização social.

Dessa forma, é evidente que existem ineficiências na moderação de conteúdo interligada a inteligência artificial, especificamente no que tange à prática de discriminação algorítmica contra grupos vulnerabilizados, revelando uma certa seletividade de ideais e propagação de conteúdo.

Todo este panorama revela relações enraizadas na sociedade, e como a moderação algorítmica demonstra como é reproduzido os padrões sistêmicos sociais e discriminatórios no meio digital, silenciando grupos vulnerabilizados, como ocorre no caso das mulheres.

## **5 OS DESAFIOS JURÍDICOS E AS LACUNAS EXISTENTES FRENTE AOS NOVOS CENÁRIOS DIGITAIS**

Os novos cenários digitais aos quais vivenciamos, como a inteligência artificial e a moderação algorítmica, ainda precisam de novas vertentes e opções legislativas focadas em suas novas diretrizes, cenários e problematizações. Assim, é evidente a existência de desafios na moderação de conteúdo, principalmente referentes às discriminações algorítmicas. Os espaços digitais acabam por absorver as diretrizes do mundo real, e com isso, as atividades humanas e os estereótipos sociais, refletindo assim, em práticas discriminatórias.

É importante ressaltar que com a promulgação da Lei n. 12.965/2014, Marco Civil da Internet, (MCI), houve uma maior ênfase nos cenários vivenciados pela sociedade e os meios tecnológicos, trazendo contribuições para o direcionamento do desenvolvimento das plataformas digitais no Brasil, mas ainda não consegue abarcar todo o cenário vivenciado atualmente.

Nesse sentido, as tecnologias, sobretudo a inteligência artificial traz desafios, sendo necessário que estas se adequem as legislações do país, garantindo assim os direitos fundamentais. A União Europeia, por exemplo, já possui diretrizes mais robustas quanto a este novo panorama das tecnologias, mesmo já tendo influenciado na formulação legislativa do Brasil, ainda há muito o que ser feito.

A abordagem jurídico-regulatória brasileira ainda é primária, não abrangendo alguns pontos que fazem parte do novo cenário social. O Marco Civil da Internet busca um equilíbrio entre a liberdade de expressão e a moderação de conteúdo, mas ainda há muito o que ser feito, esta lei trouxe importantes contribuições para o uso da internet, bem como para as diretrizes do uso das plataformas digitais no Brasil, contudo, atualmente, existem novos panoramas que não são abarcados por esta lei, como é o caso da inteligência artificial.

Nesse diapasão, outra lei que tem importância no ordenamento jurídico brasileiro voltado as novas tecnologias é a Lei nº 13.709/18, Lei Geral de Proteção de Dados, trazendo diretrizes e regras para que os dados pessoais sejam respeitados e protegidos.

Nesse sentido, a Lei Geral de Proteção de Dados possui um instrumental principiológico e normativo relevante para a orientação da atividade moderadora no Brasil (Frazão, 2018).

Diante de um contexto de amplas e recorrentes transformações no que se refere ao tratamento de dados no mundo virtual, há de se propor uma programação algorítmica dialógica, que tão somente predisponha o dialogismo entre a discriminação negativa e a positiva, considerando a LGPD como ordenamento prioritário à garantia dos direitos fundamentais dos usuários, sem que sejam impactados negativamente na sua interação com os meios digitais (Dias, 2022).

Nesse ínterim, é perceptível que não há legislação focada em alguns pontos que emergem nessa nova Era Digital, o que permite que haja uma liberdade, tanto para os provedores, como para os usuários das plataformas digitais, que acabam ficando restritos apenas as políticas e termos de usos existentes.

Dessa forma, com o avanço das novas tecnologias e os novos cenários trazidos por esta, ascende-se a necessidade da análise das legislações já existentes, e as que devem ser aprovadas, bem como criadas, para que assim, possa haver uma democratização e segurança quando se fala dessa nova era dos algoritmos, definindo uma regulamentação robusta e eficaz, focando nas ações desenvolvidas, e oferecendo maior segurança jurídica a sociedade que está imersa neste meio digital.

Assim, para lidar com estas novas questões relacionadas a moderação algorítmica, é necessário que haja a adoção de uma perspectiva sistêmica, que possa integrar diversos atores, como cidadãos, os três poderes, empresas e agências reguladoras, fortalecendo assim, as diretrizes que devem ser desenvolvidas neste meio, focando nos princípios constitucionais e no desenvolvimento das legislações.

## **6 CONSIDERAÇÕES FINAIS**

A partir de tudo que foi exposto e pesquisado, verificou-se que existe um debate atual sobre as normas conjecturas digitais, sendo necessário a existência de normas jurídicas que possam acompanhá-las, estabelecendo parâmetros objetivos e claros sobre a moderação algorítmica, para assim se chegar a um equilíbrio sobre o tema.

Ao longo deste estudo, observou-se como a violência de gênero, em especial contra as mulheres, é reverberada no cenário digital, destacando a importância do direito através de legislações mais focadas ao tema, visando na proteção de pessoas vulnerabilizadas.

As interações realizadas no mundo virtual trazem as características da sociedade ao qual estão inseridas, por isso, faz-se necessário que se identifiquem os estereótipos que são desenvolvidos através da moderação algorítmica, identificando como se dão as discriminações dos direitos, e como isso se perpetua.

Dessa forma, foi possível concluir que, no espaço da inteligência artificial, são criados e desenvolvidos moldes algorítmicos discriminatórios, que ainda são perpetuados através de uma violência sistêmica e estruturada por raízes fincadas ao longo do tempo, o que anula a percepção de que existe uma neutralidade no desenvolvimento das tecnologias, em face desta realidade algoritmo-discriminatória, assim, é necessário produzir discursos não discriminatórios, combatendo a violência de gênero.

Assim, é indiscutível que existe uma complexidade entre a efetivação dos direitos fundamentais e a moderação de algorítmica utilizada na formação da Inteligência Artificial. Nesse ínterim, é necessário que novas medidas legislativas sejam tomadas, com iniciativas compromissadas entre as empresas e o Estado.

Os direitos fundamentais devem se apresentar como vetores de toda ordem jurídica, sendo primordiais na delimitação da moderação nos meios tecnológicos, interligando e equilibrando a liberdade de expressão e o direito à informação. Dessa forma, é necessário que haja uma destruição na perpetuação da violência das mulheres por meio da Inteligência Artificial, o que demonstra a necessidade da regulamentação neste campo.

Dessa forma, é necessário que haja a adoção de medidas eficazes e capazes de comprovar o cumprimento das normas, garantindo a efetivação dos direitos fundamentais. Sendo essencial a utilização de ferramentas que verifiquem eventuais vieses discriminatórios ilícitos ou abusivos nas plataformas digitais.

É primordial que haja o entrelaçamento de direitos como a liberdade de expressão, inovação, transparência de dados, acesso à informação; onde todos possam convergir na responsabilidade civil, diminuindo as desigualdades estruturais, e evitando o silenciamento de grupos historicamente marginalizados.

Diante de todo panorama, conclui-se que é necessário que as mulheres estejam presentes nos debates e no desenvolvimento de diretrizes nesse mundo digital, sendo representadas, mas também sendo sujeito de decisões, principalmente nas moderações algorítmicas, identificando as configurações algorítmicas em relação as questões de gênero e as discriminações ali embutidas, bem como a forma como os dados são tratados.

## REFERÊNCIAS

ACIOLY, L. H. M. A aplicação do Princípio da Função Social do Contrato em Contratos Eletrônicos de redes sociais e seu papel na democratização de Direitos Fundamentais. **Revista Conversas Civilísticas**, Salvador, v. 1, n. 2, p. 111-134, jul./dez. 2021. Disponível em: <https://periodicos.ufba.br/index.php/conversascivilisticas/article/view/47143>. Acesso em: 28 set. 2025.

ACIOLY, L. H. M. Reflexões sobre a Utilização da Inteligência Artificial e Algoritmos nas Relações Consumeristas à Luz da Lei Geral de Proteção de Dados. **Revista Brasileira de Direito Comercial – Concorrência, Empresarial e do Consumidor**. Porto Alegre: Lex Magister. v. 48, p. 165-187, ago./set. 2022.

AGÊNCIA NACIONAL DE PROTEÇÃO DE DADOS – ANPD. **Portal institucional da ANPD**. Disponível em: <https://www.gov.br/anpd/pt-br> . Acesso em: 27 set. 2025.

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

BARROSO, Luís Roberto. **Curso de direito constitucional contemporâneo**. 12. ed. São Paulo: Saraiva Jur, 2024.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm) . Acesso em: 11 ago. 2025.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Diário Oficial da União: seção 1, Brasília, DF, 24 abr. 2014. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/112965.htm](https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm) . Acesso em: 26 set. 2025.

BRASIL. **Lei nº 13.709**, de 14 de agosto de 2018. Dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, e dá outras providências. Diário Oficial da União: seção 1, Brasília,

DF, 15 ago. 2018. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/113709.htm](https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm) . Acesso em: 26 set. 2025.

CARMO, C. M. Grupos minoritários, grupos vulneráveis e o problema da (in)tolerância: uma relação linguístico-discursiva e ideológica entre o desrespeito e a manifestação do ódio no contexto brasileiro. **Revista do Instituto de Estudos Brasileiros**, Brasil, n. 64, p. 201-223, ago. 2016. DOI: <http://dx.doi.org/10.11606/issn.2316-901X.v0i64p201-223>.

CEZARINO, M. R.; CONTRI, C. L. As implicações da construção binária do gênero para a realização de decisões automatizadas que impactam diretamente as pessoas trans e não-binárias. In: BARBOSA, B.; TRESKA, L.; LAUSCHNER, T. (orgs.). **Governança da Internet e Gênero: tendências e desafios**. [s.l.]:CGI.Br. 2. ed., 2022. p. 47-68.

CHAMI, Amanda. Problematizando “ineficiências”: reflexões na fronteira entre inteligência artificial, moderação de conteúdo e discriminação algorítmica, em perspectiva decolonial. **Revista Brasileira de Direito Civil – RBDCivil**, Belo Horizonte, v. 33, n. 1, p. 281-296, jan./mar. 2024. DOI: 10.33242/rbdc.2024.01.012.

DA SILVA, Adriano Alves; JÚNIOR, Francisco Gilson Rebouças Pôrto. Algoritmos silenciadores: desinformação e espiral do silêncio na era da inteligência artificial. **Organicom**, v. 21, n. 44, p. 147-158, 2024.

DE GREGORIO, Giovanni. The Rise of Digital Constitutionalism in the European Union. **International Journal of Constitutional Law**, v. 19, n. 1, p. 41-70, 2020. Disponível em: <https://ssrn.com/abstract=3506692> . Acesso em: 21 set. 2025.

FARIA, Lucas Ribeiro de; SILVA, Lucas Gonçalves da; CARDOSO, Henrique Ribeiro. Deepfake pornográfico na sociedade de risco contemporânea: os desafios de regulamentação e controle da inteligência artificial. **Revista Interfaces Científicas – Direito**, Aracaju, v. 9, n. 3, p. 343-355, 2024. DOI: 10.17564/2316-381X.2024v9n3p343-355. Disponível em: <https://periodicos.set.edu.br/interfacesjuridicas/article/view/12029>. Acesso em: 27 abr. 2025.

DE JESUS PINHEIRO, Lucas Aparecido. Liberdade de expressão e a regulamentação da moderação em plataformas online: Considerações sobre o Projeto de Lei 2630/2020. **Virtuajus**, v. 9, n. 16, p. 270-282, 2024.

DOS SANTOS, Tathiane Graça. **O Poder Invisível do Algoritmo a Discriminação de Gênero no Mercado de Trabalho na era da Inteligência Artificial**. 2023. Dissertação de Mestrado. Universidade NOVA de Lisboa (Portugal).

DOUEK, Evelyn. Content Moderation as Systems Thinking. *Harvard Law Review*. Cambridge: **Harvard University Press**. v. 136, n. 2, p. 526-607, dez. 2022.

ÉVORA, S. L. Comunicação Política, inteligência artificial e ciberesfera. **Revista Internacional em Língua Portuguesa**, [S. l.], n. 43, p. 67–92, 2023. DOI: 10.31492/2184-2043.RILP2023.43/pp.67-92.

FARID, Hany. An Overview of Perceptual Hashing. **Journal of Online Trust and Safety**, v. 1, n. 1, 2021. Disponível em: <https://tsjournal.org/index.php/jots/article/view/24>. Acesso em: 26 set. 2025.

FRANCO, Fernanda Sathler Rocha. Regulação da moderação de conteúdo digital e principais estratégias jurídicas brasileiras: análise do Marco Civil da Internet, da Lei Geral de Proteção de Dados e do Projeto de Lei n. 2.630/20. **Revista de Direito, Inovação, Propriedade Intelectual e Concorrência**, [S.l.], v. 10, n. 2, 2024.

FREDES, Andrei Ferreira. **Liberdade de expressão, direito à informação e redes sociais: regulação constitucionalmente adequada sobre a moderação de conteúdo na construção de um espaço virtual democrático e plural**. 2022.

GONÇALVES DA SILVA, Lucas; CARDOZO OLIVEIRA, Francisco; DE MENDONÇA SIQUEIRA, Alessandra Cristina. A (HÁ) LIBERDADE DE EXPRESSÃO NA SOCIEDADE EM REDE (?): MANIPULAÇÃO NA ERA DIGITAL. **Brazilian Journal of Law & International Relations/Relações Internacionais no Mundo**, v. 2, n. 26, 2019.

GONÇALVES DA SILVA, L., & FELIX NASCIMENTO, R. (2023). Colonialismo de dados: o apagar das vozes no conhecimento social. **Revista De Direitos E Garantias Fundamentais**, 24(3),73–105. Disponível em: <https://sisbib.emnuvens.com.br/direitosegarantias/article/view/2296>. Acesso em 26 set.2025

HOGEMANN, Edna Raquel. **Sobre os impactos da inteligência artificial em sociedades caracterizadas pela desigualdade social**.

HOFFMANN-RIEM, W. **Teoria Geral do Direito Digital**. 2.ed. Trad. Italo Fuhrmann. Rio de Janeiro: Forense, 2021.

IRIS BH. **Inteligência artificial e discriminação contra mulheres: os dados e o sistema**. Disponível em: <https://irisbh.com.br/inteligencia-artificial-e-discriminacao-contra-mulheres-os-dados-e-o-sistema/> . Acesso em: 29 set. 2025.

JUST, Natascha; LATZER, Michael. Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet. **Media, Culture & Society**, 2016. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3871903](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3871903). Acesso em 26 set. 2025.

LEITE, George S.; LEMOS, Ronaldo (Coords.). **Marco Civil da Internet**. São Paulo: Atlas, 2014.

MARRAFON, Marco Aurélio; LUCENA, Marina Giovanetti Lili. Transparência e controle de conteúdo em redes sociais e o direito fundamental à liberdade de expressão. **Revista Direito & Paz**, v. 1, n. 46, p. 249-275, 2022.

NASCIMENTO, Reginaldo Felix; SANTOS, Helen Caroline Cardoso; SILVA, Lucas Gonçalves da. Responsabilidade civil por atos independentes de inteligências artificiais. **Revista Percorso Unicuritiba**, v. 2, n. 50, p. 136-149, abr./jun. 2025. Disponível em <https://revista.unicuritiba.edu.br/index.php/percurso/article/view/7774/371375347>. Acesso em: 26 set.2025.

TEFFÉ, Chiara; MORAES, Maria. Redes sociais virtuais: privacidade e responsabilidade civil - Análise a partir do Marco Civil da Internet. **Pensar**, n. 22, 2017. Disponível em: <https://ojs.unifor.br/rpen/article/view/6272>. Acesso em 27 set. 2025.

TIROCCHI, Simona et al. Violência artificial: violência contra mulheres e o lado obscuro da inteligência artificial. **Comunicação & Educação**, v. 30, n. 1, p. 181-197, 2025.