

**RESPONSABILIDADE  
PENAL DAS PLATAFORMAS  
DE IA POR DANOS À SAÚDE  
MENTAL DE JOVENS:  
CONTEÚDOS  
SUICIDOGÊNICOS  
RECOMENDADOS POR  
ALGORITMOS**

**CRIMINAL LIABILITY OF AI PLATFORMS FOR HARM TO THE MENTAL  
HEALTH OF YOUNG PEOPLE: SUICIDOGENIC CONTENT RECOMMENDED  
BY ALGORITHMS**

Ciências Sociais Aplicadas, Ciências da Saúde • 28/05/2026

REGISTRO DOI: [10.70773/revistatopicos/779820984](https://doi.org/10.70773/revistatopicos/779820984)

Cleiany Venâncio Gonçalves Figueira Cardoso

Tayana de Souza Bordalo

## RESUMO

A crescente utilização de plataformas digitais baseadas em inteligência artificial tem levantado debates relevantes acerca de sua responsabilidade jurídica, especialmente no que se refere à saúde mental de jovens usuários. Este artigo analisa a possibilidade de responsabilização penal das plataformas de IA diante da recomendação algorítmica de conteúdos suicidogênicos, considerando o papel dos algoritmos na amplificação e direcionamento de informações sensíveis. A pesquisa aborda o funcionamento dos sistemas de recomendação, que, ao priorizarem engajamento, podem expor usuários vulneráveis a conteúdos prejudiciais de forma reiterada. Discute-se, ainda, os limites da responsabilidade penal dessas plataformas, à luz da legislação vigente, bem como os desafios na comprovação do nexo causal entre a recomendação automatizada e os danos psíquicos causados. A metodologia adotada baseia-se em revisão bibliográfica e análise de casos recentes que evidenciam a influência das redes sociais no comportamento de jovens. Como resultado, observa-se a necessidade de atualização dos marcos regulatórios, visando maior controle e transparência dos algoritmos, além da implementação de medidas preventivas para proteção de usuários em situação de vulnerabilidade. Conclui-se que, embora complexa, a responsabilização penal das plataformas é um tema urgente no contexto da sociedade digital contemporânea.

**Palavras-chave:** Inteligência artificial; Responsabilidade penal; Saúde mental juvenil.

## ABSTRACT

The increasing use of digital platforms based on artificial intelligence has raised important debates regarding their legal responsibility, especially concerning the mental health of young users. This article

analyzes the possibility of criminal liability of AI platforms in cases involving the algorithmic recommendation of suicide-related content, considering the role of algorithms in amplifying and directing sensitive information. The study examines how recommendation systems, by prioritizing user engagement, may repeatedly expose vulnerable individuals to harmful content. It also discusses the limits of criminal liability of these platforms under current legislation, as well as the challenges in establishing a causal link between automated recommendations and the resulting psychological harm. The methodology is based on a bibliographic review and analysis of recent cases that highlight the influence of social media on youth behavior. The results indicate the need to update regulatory frameworks to ensure greater control and transparency of algorithms, as well as the implementation of preventive measures to protect vulnerable users. It is concluded that, although complex, the criminal liability of platforms is an urgent issue in today's digital society.

**Keywords:** Artificial intelligence; Criminal liability; Youth mental health.

## 1. INTRODUÇÃO

A expansão das tecnologias digitais nas últimas décadas transformou profundamente as formas de interação social, comunicação e acesso à informação, especialmente entre jovens. Nesse contexto, as plataformas baseadas em inteligência artificial passaram a desempenhar papel central na mediação de conteúdos, influenciando diretamente hábitos, comportamentos e percepções dos usuários.

Os sistemas de recomendação algorítmica, desenvolvidos para maximizar o engajamento, utilizam grandes volumes de dados para personalizar a experiência do usuário. Embora essa personalização possa trazer benefícios, também levanta preocupações quanto à exposição contínua a conteúdos potencialmente prejudiciais, sobretudo quando direcionados a indivíduos em situação de vulnerabilidade emocional.

Entre os conteúdos considerados sensíveis, destacam-se aqueles de natureza suicidogênica, que podem incluir desde discussões sobre automutilação até a romantização do suicídio. A disseminação e recomendação desses materiais têm sido objeto de crescente atenção por parte da comunidade científica, jurídica e da sociedade em geral.

A adolescência e a juventude constituem fases marcadas por intensas transformações psicológicas e sociais, o que torna os indivíduos mais suscetíveis a influências externas. Nesse cenário, a exposição repetida a conteúdos negativos pode agravar quadros de ansiedade, ideação suicida, especialmente quando não há mediação adequada.

Diante dessa realidade, surge a necessidade de discutir o papel das plataformas digitais na proteção da saúde mental dos usuários. A atuação dos algoritmos não é neutra, pois envolve decisões automatizadas que priorizam determinados conteúdos em detrimento de outros, impactando diretamente a experiência do usuário.

A problemática se intensifica quando se observa que muitos desses sistemas operam com baixa transparência, dificultando a

compreensão de seus critérios de funcionamento. Essa opacidade algorítmica representa um desafio significativo para a responsabilização jurídica das plataformas.

No campo do Direito Penal, a discussão acerca da responsabilização das plataformas digitais ainda é recente e complexa. Tradicionalmente, a responsabilidade penal está vinculada à conduta humana, o que levanta questionamentos sobre a possibilidade de atribuir culpa a sistemas automatizados ou às empresas que os desenvolvem e operam.

Além disso, a comprovação do nexo causal entre a recomendação de conteúdo e o dano à saúde mental constitui um dos principais entraves para a responsabilização penal. Trata-se de uma relação multifatorial, na qual diversos elementos externos podem influenciar o comportamento do indivíduo.

Outro aspecto relevante refere-se à liberdade de expressão e ao acesso à informação, direitos fundamentais que também devem ser considerados nesse debate. A imposição de limites às plataformas exige equilíbrio entre a proteção dos usuários e a preservação desses direitos.

No âmbito internacional, observa-se o surgimento de iniciativas regulatórias voltadas à responsabilização de plataformas digitais e à proteção de usuários vulneráveis. Essas experiências podem servir de referência para o desenvolvimento de políticas públicas mais eficazes em outros contextos.

A presente pesquisa justifica-se pela necessidade de aprofundar a análise sobre os impactos dos algoritmos na saúde mental de jovens, bem como pela urgência de se discutir mecanismos jurídicos

capazes de responsabilizar eventuais condutas lesivas promovidas por plataformas digitais.

Dessa forma, o objetivo deste estudo é analisar a possibilidade de responsabilização penal das plataformas de inteligência artificial em casos envolvendo a recomendação de conteúdos suicidogênicos, contribuindo para o debate acadêmico e jurídico sobre os desafios impostos pela sociedade digital contemporânea.

## **2. FUNDAMENTAÇÃO TEÓRICA**

A fundamentação teórica deste estudo parte da compreensão do funcionamento dos algoritmos de recomendação no contexto da inteligência artificial, especialmente aqueles utilizados por plataformas digitais para personalização de conteúdo. Tais sistemas operam com base na coleta e análise de dados comportamentais dos usuários, priorizando conteúdos com maior potencial de engajamento. Segundo a literatura contemporânea, essa lógica pode gerar efeitos colaterais, como a amplificação de conteúdos sensíveis, incluindo materiais de cunho suicidogênico, sobretudo quando direcionados a públicos vulneráveis, como jovens em desenvolvimento emocional.

No campo jurídico, a discussão acerca da responsabilidade penal das plataformas digitais envolve conceitos clássicos como dolo, culpa e nexo de causalidade, os quais enfrentam desafios quando aplicados a sistemas automatizados. A doutrina aponta que, embora os algoritmos não possuam vontade própria, as empresas que os desenvolvem e operam podem ser responsabilizadas caso se comprove negligência, imprudência ou omissão na prevenção de danos previsíveis. Assim, emerge o debate sobre a responsabilidade

penal por omissão imprópria, especialmente quando há dever jurídico de agir para evitar resultados lesivos.

Por fim, no âmbito da saúde mental, estudos indicam que a exposição contínua a conteúdos negativos pode impactar significativamente o bem-estar psicológico de jovens, contribuindo para o agravamento de quadros de ansiedade, depressão e ideação suicida. A teoria da vulnerabilidade reforça que indivíduos em determinadas fases da vida ou condições emocionais estão mais suscetíveis a influências externas, o que impõe maior responsabilidade às plataformas digitais. Dessa forma, a interseção entre tecnologia, direito e saúde evidencia a necessidade de abordagens multidisciplinares para enfrentar os riscos associados ao uso de inteligência artificial na mediação de conteúdo.

## **2.1. Funcionamento dos Algoritmos de Recomendação**

Os algoritmos de recomendação constituem o núcleo operacional das plataformas digitais contemporâneas, sendo responsáveis por selecionar, organizar e sugerir conteúdos aos usuários com base em seus comportamentos e preferências. Esses sistemas utilizam técnicas avançadas de inteligência artificial e aprendizado de máquina para analisar grandes volumes de dados, permitindo a personalização da experiência digital de forma contínua e dinâmica. A partir dessa lógica, cada usuário passa a receber conteúdos específicos, moldados conforme seu histórico de interações.

O funcionamento desses algoritmos baseia-se, principalmente, na coleta de dados como curtidas, compartilhamentos, tempo de visualização e padrões de navegação. Essas informações são processadas por modelos matemáticos que identificam padrões e

tendências comportamentais, possibilitando prever quais conteúdos têm maior probabilidade de engajar o usuário. Esse processo ocorre de maneira automatizada e em tempo real, o que amplia significativamente sua eficácia e alcance.

Segundo Silva (2021, p. 87):

*Os algoritmos de recomendação utilizados pelas plataformas digitais operam a partir da análise contínua de dados comportamentais dos usuários, priorizando conteúdos que geram maior engajamento, o que pode resultar na amplificação de materiais sensíveis e potencialmente prejudiciais, especialmente quando direcionados a indivíduos em situação de vulnerabilidade emocional.*

Uma das principais técnicas utilizadas nesses sistemas é a filtragem colaborativa, que considera o comportamento de usuários com perfis semelhantes para recomendar conteúdos. Além disso, há também a filtragem baseada em conteúdo, que analisa características específicas dos materiais consumidos anteriormente. A combinação dessas abordagens resulta em um sistema altamente eficiente na retenção da atenção dos usuários.

Entretanto, a lógica de funcionamento dos algoritmos está diretamente vinculada a objetivos comerciais das plataformas, sobretudo a maximização do engajamento e do tempo de permanência. Isso significa que conteúdos que despertam emoções intensas, como medo, tristeza ou curiosidade, tendem a ser

priorizados, independentemente de seus possíveis impactos negativos sobre o usuário.

Nesse contexto, surge o fenômeno das chamadas “bolhas informacionais”, nas quais o usuário é exposto repetidamente a conteúdos semelhantes, reforçando determinadas visões, comportamentos e emoções. Esse efeito pode limitar o acesso a informações diversificadas e contribuir para a intensificação de determinados estados psicológicos.

A repetição contínua de conteúdos com características específicas pode gerar um efeito cumulativo, especialmente quando se trata de materiais sensíveis. No caso de conteúdos de cunho negativo ou prejudicial, como aqueles relacionados à automutilação ou suicídio, a exposição frequente pode influenciar significativamente o estado emocional do usuário, principalmente entre jovens.

Outro aspecto relevante é a opacidade dos algoritmos, frequentemente descrita como uma “caixa-preta”. As plataformas, em geral, não divulgam de forma transparente os critérios utilizados em seus sistemas de recomendação, o que dificulta a compreensão de como e por que determinados conteúdos são sugeridos.

Essa falta de transparência representa um desafio tanto para pesquisadores quanto para legisladores, que enfrentam dificuldades na análise e regulamentação desses sistemas. Sem acesso aos mecanismos internos, torna-se mais complexo identificar falhas, vieses e possíveis negligências no funcionamento dos algoritmos.

Além disso, os algoritmos são constantemente atualizados e ajustados com base em novos dados, o que torna seu comportamento dinâmico e, por vezes, imprevisível. Essa

característica reforça a necessidade de monitoramento contínuo e de mecanismos de controle mais eficazes.

A interação entre usuário e algoritmo também deve ser considerada como um processo bidirecional, no qual o comportamento do indivíduo influencia o sistema, ao mesmo tempo em que é influenciado por ele. Essa relação cria um ciclo de retroalimentação que pode intensificar padrões de consumo de conteúdo.

No caso de usuários em situação de vulnerabilidade emocional, esse ciclo pode se tornar ainda mais preocupante, pois o algoritmo tende a reforçar conteúdos compatíveis com o estado psicológico do indivíduo. Assim, alguém que consome conteúdos relacionados à tristeza ou isolamento pode passar a receber cada vez mais materiais com esse perfil.

Dessa forma, o funcionamento dos algoritmos de recomendação, embora eficiente do ponto de vista tecnológico e comercial, apresenta implicações relevantes no campo social e psicológico. Compreender sua lógica é essencial para o desenvolvimento de estratégias que minimizem riscos e promovam um ambiente digital mais seguro, especialmente para públicos mais vulneráveis.

## **2.2. Inteligência Artificial e Curadoria de Conteúdo**

A inteligência artificial tem desempenhado um papel cada vez mais relevante na curadoria de conteúdo nas plataformas digitais, sendo responsável por selecionar, organizar e priorizar informações de acordo com o perfil de cada usuário. Diferentemente dos modelos tradicionais de curadoria, que envolviam intervenção humana, a curadoria automatizada baseia-se em algoritmos capazes de

processar grandes volumes de dados em alta velocidade, proporcionando uma experiência personalizada e contínua.

Esse processo de curadoria fundamenta-se na análise de dados comportamentais, como histórico de navegação, interações anteriores e preferências explícitas e implícitas dos usuários. A partir dessas informações, os sistemas de inteligência artificial conseguem identificar padrões e prever quais conteúdos possuem maior probabilidade de atrair a atenção do indivíduo, promovendo assim uma seleção automatizada que atende aos interesses específicos de cada perfil.

A utilização da inteligência artificial na curadoria de conteúdo também está diretamente associada à eficiência operacional das plataformas digitais. Com bilhões de usuários ativos e uma quantidade massiva de informações sendo produzidas diariamente, torna-se inviável realizar esse processo de forma manual. Dessa forma, os algoritmos assumem a função de filtrar o que será exibido, organizando o fluxo de informações de maneira escalável.

Segundo Pereira (2020, p. 142):

*A exposição recorrente a conteúdos relacionados à automutilação e ao suicídio no ambiente digital pode contribuir para a normalização desses comportamentos, sobretudo entre jovens, que se encontram em fase de desenvolvimento psicológico e apresentam maior suscetibilidade à influência de estímulos externos.*

No entanto, essa automatização não está isenta de críticas, especialmente no que se refere à neutralidade das decisões algorítmicas. Embora os sistemas sejam projetados para operar com base em dados, eles refletem objetivos definidos por seus desenvolvedores, como o aumento do engajamento e da retenção de usuários. Isso pode resultar na priorização de conteúdos mais apelativos ou emocionalmente intensos, independentemente de sua qualidade ou impacto social.

Outro ponto relevante diz respeito à formação de ambientes digitais altamente personalizados, nos quais os usuários são expostos predominantemente a conteúdos alinhados às suas preferências anteriores. Essa dinâmica pode limitar a diversidade informacional e contribuir para a formação de visões parciais da realidade, reforçando crenças e comportamentos já existentes.

A curadoria automatizada também levanta questões éticas relacionadas à responsabilidade das plataformas sobre os conteúdos recomendados. Ao selecionar ativamente o que será exibido, os sistemas de inteligência artificial deixam de ser meros intermediários e passam a exercer influência direta sobre a experiência do usuário, o que intensifica o debate sobre sua responsabilidade em casos de exposição a conteúdos prejudiciais.

Além disso, a ausência de transparência nos critérios utilizados pelos algoritmos dificulta a fiscalização e a responsabilização das plataformas. Os usuários, em geral, não possuem clareza sobre os motivos pelos quais determinados conteúdos são exibidos, o que reforça a assimetria de informação entre as empresas e o público.


Dessa forma, a inteligência artificial aplicada à curadoria de conteúdo representa um avanço significativo do ponto de vista tecnológico, mas também impõe desafios relevantes no campo ético, social e jurídico. A compreensão desse processo é fundamental para o desenvolvimento de mecanismos que garantam maior equilíbrio entre inovação, responsabilidade e proteção dos usuários no ambiente digital.

### **2.3. Conteúdos Suicidogênicos no Ambiente Digital**

Os conteúdos suicidogênicos no ambiente digital referem-se a materiais que abordam, incentivam ou romantizam práticas de automutilação e suicídio, podendo incluir textos, imagens, vídeos e interações em comunidades virtuais. Com o avanço das plataformas digitais, esse tipo de conteúdo passou a circular com maior facilidade, alcançando especialmente públicos jovens, que são mais ativos no consumo de mídias sociais e, ao mesmo tempo, mais vulneráveis a influências externas.

A presença desses conteúdos nas plataformas digitais tem sido associada a efeitos negativos sobre a saúde mental, sobretudo quando ocorre exposição frequente e não mediada. Estudos indicam que o contato contínuo com esse tipo de material pode contribuir para a normalização de comportamentos autodestrutivos, além de potencializar sentimentos de desesperança, isolamento e baixa autoestima em indivíduos já fragilizados emocionalmente.

Segundo Almeida (2022, p. 59):



*A responsabilidade das plataformas digitais não pode ser compreendida apenas sob a ótica da neutralidade tecnológica, uma vez que os sistemas de curadoria automatizada exercem papel ativo na seleção e priorização de conteúdos, influenciando diretamente a experiência do usuário e seus possíveis desdobramentos.*

No ambiente digital, os conteúdos suicidogênicos podem assumir diferentes formas, desde discussões aparentemente informativas até representações explícitas ou simbólicas do suicídio. Em muitos casos, esses materiais são apresentados de maneira estética ou emocionalmente envolvente, o que pode aumentar seu impacto psicológico, especialmente entre adolescentes em fase de desenvolvimento identitário.

Outro fator preocupante é a formação de comunidades virtuais que compartilham e reforçam esse tipo de conteúdo, criando espaços onde comportamentos autodestrutivos são validados ou até incentivados. Essas interações podem gerar um sentimento de pertencimento, o que, paradoxalmente, fortalece a permanência do indivíduo em um ciclo de sofrimento e exposição contínua a estímulos negativos.

A atuação dos algoritmos de recomendação intensifica esse problema ao sugerir conteúdos semelhantes com base no histórico de navegação do usuário. Dessa forma, um primeiro contato com materiais relacionados pode desencadear uma sequência de

recomendações que ampliam a exposição, criando um ambiente propício à repetição e ao aprofundamento desse tipo de consumo.

Além disso, a dificuldade de moderação eficaz desses conteúdos representa um desafio significativo para as plataformas digitais. Embora existam políticas de remoção e mecanismos de denúncia, a grande quantidade de informações produzidas diariamente e a complexidade na identificação de conteúdos sensíveis dificultam o controle total desse material.

No campo jurídico e ético, a circulação de conteúdos suicidogênicos levanta debates sobre os limites da liberdade de expressão e a responsabilidade das plataformas na proteção dos usuários. A ausência de regulamentação específica em muitos contextos contribui para a manutenção de lacunas que dificultam a responsabilização em casos de danos à saúde mental.

Dessa forma, os conteúdos suicidogênicos no ambiente digital configuram um problema complexo e multifacetado, que envolve aspectos tecnológicos, psicológicos e jurídicos. A compreensão de suas dinâmicas é essencial para o desenvolvimento de estratégias de prevenção, regulação e proteção, especialmente voltadas ao público jovem, que se encontra em maior situação de vulnerabilidade.

#### **2.4. Vulnerabilidade Psicológica de Jovens**

A vulnerabilidade psicológica de jovens constitui um dos principais fatores a serem considerados na análise dos impactos do ambiente digital sobre a saúde mental. A adolescência e o início da vida adulta são fases marcadas por intensas transformações emocionais, cognitivas e sociais, nas quais o indivíduo ainda está em processo de

formação de identidade, autoestima e habilidades de regulação emocional. Nesse contexto, a exposição a conteúdos digitais pode exercer influência significativa sobre pensamentos, sentimentos e comportamentos.

Durante esse período, os jovens tendem a buscar pertencimento, validação social e reconhecimento, muitas vezes por meio das interações em plataformas digitais. A necessidade de aceitação pode torná-los mais suscetíveis à internalização de padrões, discursos e conteúdos que circulam nesses ambientes, inclusive aqueles de natureza negativa ou prejudicial à saúde mental.

Além disso, fatores como insegurança, baixa autoestima, dificuldades familiares, pressões sociais e experiências de exclusão podem intensificar a vulnerabilidade emocional. Quando combinados com o uso frequente de redes sociais, esses elementos podem contribuir para o agravamento de quadros de ansiedade, depressão e sentimentos de inadequação.

Segundo Costa (2019, p. 203):

*A dificuldade de estabelecer o nexo de causalidade entre a atuação dos algoritmos e os danos à saúde mental dos usuários constitui um dos principais desafios para a responsabilização jurídica das plataformas, tendo em vista a natureza multifatorial dos transtornos psicológicos.*

A imaturidade no desenvolvimento cognitivo também desempenha papel relevante, uma vez que jovens podem apresentar maior dificuldade em avaliar criticamente informações, interpretar riscos e compreender as consequências de determinados comportamentos. Isso os torna mais propensos a influências externas, especialmente quando estas são apresentadas de forma repetitiva ou emocionalmente envolvente.

No ambiente digital, a exposição contínua a conteúdos negativos ou sensíveis pode atuar como um gatilho para o surgimento ou agravamento de transtornos psicológicos. Conteúdos relacionados à automutilação, sofrimento extremo ou ideação suicida, quando consumidos de forma recorrente, podem ser interpretados como formas de expressão legítimas ou até soluções para conflitos internos.

Outro aspecto relevante é o fenômeno da comparação social, amplamente intensificado pelas redes digitais. A constante visualização de padrões idealizados de vida, corpo e sucesso pode gerar frustração, sensação de inadequação e diminuição da autoestima, contribuindo para o sofrimento psíquico.

A interação com algoritmos de recomendação também influencia diretamente essa vulnerabilidade, uma vez que os sistemas tendem a reforçar conteúdos alinhados ao estado emocional do usuário. Assim, jovens que demonstram interesse por conteúdos tristes ou negativos podem ser expostos a uma quantidade crescente desse tipo de material, criando um ciclo de retroalimentação prejudicial.

Além disso, a falta de acompanhamento adequado por parte de adultos, como pais e educadores, pode agravar esse cenário. A

ausência de orientação e diálogo dificulta a identificação precoce de sinais de sofrimento emocional e limita a capacidade de intervenção em situações de risco.

Do ponto de vista psicológico, a vulnerabilidade não deve ser compreendida apenas como fragilidade, mas como uma condição que exige atenção, cuidado e proteção. Trata-se de um estado que pode ser agravado ou atenuado conforme o ambiente em que o jovem está inserido, incluindo o espaço digital.

Dessa forma, compreender a vulnerabilidade psicológica de jovens é fundamental para a construção de estratégias de prevenção e intervenção, bem como para o desenvolvimento de políticas públicas e mecanismos de regulação das plataformas digitais. Esse entendimento contribui para a promoção de um ambiente virtual mais seguro e responsável, capaz de proteger indivíduos em fases críticas do desenvolvimento humano.

## **2.5. Responsabilidade Penal das Plataformas Digitais**

A responsabilidade penal das plataformas digitais constitui um tema emergente no campo jurídico, especialmente diante do avanço das tecnologias baseadas em inteligência artificial e da crescente influência dessas empresas na mediação de conteúdos. Tradicionalmente, o Direito Penal é orientado pelo princípio da responsabilidade subjetiva, que exige a presença de dolo ou culpa para a imputação de um crime, o que torna complexa a aplicação desse modelo a estruturas automatizadas como algoritmos.

Nesse contexto, surge o debate acerca da possibilidade de responsabilização das próprias plataformas enquanto pessoas jurídicas. Embora o ordenamento jurídico admita, em situações

específicas, a responsabilização penal de empresas, ainda há limitações quanto à sua aplicação em casos envolvendo danos à saúde mental decorrentes de conteúdos digitais. A principal dificuldade reside na identificação de uma conduta típica diretamente atribuível à plataforma.

Outro ponto central refere-se à análise da omissão imprópria, na qual a responsabilidade pode ser atribuída quando há um dever jurídico de agir para evitar o resultado danoso. No caso das plataformas digitais, discute-se se essas empresas possuem o dever de impedir a circulação ou recomendação de conteúdos prejudiciais, especialmente quando envolvem usuários em situação de vulnerabilidade.

Segundo Rodrigues (2023, p. 118):

*A ausência de transparência nos sistemas algorítmicos das plataformas digitais impede uma fiscalização eficaz e compromete a construção de mecanismos regulatórios adequados, dificultando a proteção de usuários diante da exposição a conteúdos potencialmente nocivos.*

A comprovação do nexo de causalidade também representa um desafio significativo, uma vez que os danos à saúde mental são, em geral, multifatoriais. Estabelecer uma relação direta entre a atuação dos algoritmos e o resultado lesivo exige uma análise aprofundada, que considere aspectos individuais, sociais e tecnológicos.

Além disso, há a necessidade de equilibrar a responsabilização penal com direitos fundamentais, como a liberdade de expressão e o acesso à informação. A imposição de sanções às plataformas deve ser cuidadosamente analisada para evitar excessos que possam comprometer esses direitos, mantendo-se a proporcionalidade e a razoabilidade.

Dessa forma, a responsabilização penal das plataformas digitais demanda uma releitura dos institutos jurídicos tradicionais à luz das transformações tecnológicas. Trata-se de um campo em construção, que exige atualização normativa e reflexão crítica para garantir a proteção dos usuários sem inviabilizar a inovação no ambiente digital.

## **2.6. Desafios Regulatórios e Limites Jurídicos**

Os desafios regulatórios e os limites jurídicos relacionados às plataformas digitais e à inteligência artificial refletem a complexidade de um cenário em constante transformação tecnológica. A velocidade com que novas ferramentas e sistemas são desenvolvidos frequentemente supera a capacidade do Direito de acompanhá-los, gerando lacunas normativas que dificultam a atuação eficaz do Estado na proteção dos usuários, especialmente em situações que envolvem riscos à saúde mental.

Um dos principais desafios consiste na ausência de legislações específicas que tratem diretamente da responsabilidade das plataformas em relação aos conteúdos recomendados por algoritmos. Embora existam normas gerais aplicáveis ao ambiente digital, muitas delas não contemplam as particularidades da

inteligência artificial, como a autonomia relativa dos sistemas e a opacidade de seus processos decisórios.

A natureza global das plataformas digitais também representa um obstáculo significativo para a regulação. Empresas operam em múltiplos países, com diferentes ordenamentos jurídicos, o que dificulta a aplicação uniforme de normas e a responsabilização em casos concretos. Essa dimensão transnacional exige cooperação internacional e harmonização de entendimentos jurídicos.

Outro ponto relevante refere-se à transparência algorítmica, frequentemente limitada por questões comerciais e estratégicas das empresas. A falta de acesso aos critérios utilizados pelos sistemas de recomendação impede uma fiscalização mais efetiva por parte de autoridades reguladoras e da sociedade civil, comprometendo a identificação de possíveis abusos ou falhas.

Além disso, há o desafio de equilibrar a regulação com a preservação de direitos fundamentais, como a liberdade de expressão, a privacidade e o acesso à informação. Medidas excessivamente restritivas podem gerar efeitos colaterais indesejados, como censura ou limitação da inovação tecnológica, o que exige cautela na formulação de políticas públicas.

A definição de limites jurídicos também passa pela dificuldade de estabelecer critérios objetivos para caracterizar conteúdos prejudiciais, como os suicidogênicos, sem incorrer em subjetividades ou interpretações divergentes. Essa indefinição pode comprometer a segurança jurídica e dificultar a atuação das plataformas na moderação de conteúdos.

Outro aspecto importante é a responsabilização proporcional das empresas, considerando seu grau de controle sobre os conteúdos e os mecanismos disponíveis para prevenção de danos. A ausência de parâmetros claros pode levar tanto à impunidade quanto à responsabilização excessiva, prejudicando o equilíbrio do sistema jurídico.

Dessa forma, os desafios regulatórios e os limites jurídicos no contexto das plataformas digitais e da inteligência artificial evidenciam a necessidade de atualização legislativa, desenvolvimento de mecanismos de fiscalização mais eficazes e construção de um marco regulatório que concilie inovação tecnológica com proteção dos direitos fundamentais e da saúde mental dos usuários.

## **2.7. Impactos da Exposição Contínua a Conteúdos Sensíveis**

A exposição contínua a conteúdos sensíveis no ambiente digital tem se configurado como um fator de risco relevante para a saúde mental, especialmente entre jovens. Esses conteúdos, que podem envolver temas como sofrimento psicológico intenso, automutilação e suicídio, quando consumidos de forma repetitiva, tendem a influenciar diretamente o estado emocional dos indivíduos, contribuindo para a intensificação de sentimentos negativos.

Um dos principais impactos dessa exposição é a dessensibilização progressiva em relação a temas delicados. O contato frequente com conteúdos de natureza extrema pode reduzir a percepção de gravidade dessas situações, levando à naturalização de comportamentos autodestrutivos e à diminuição das barreiras psicológicas que normalmente impediriam tais práticas.

Além disso, a repetição de conteúdos sensíveis pode atuar como um reforço emocional negativo, criando um ciclo no qual o usuário passa a consumir materiais que refletem ou intensificam seu próprio sofrimento. Esse processo pode dificultar a busca por ajuda e agravar quadros de ansiedade, depressão e desesperança.

Outro efeito importante é o fenômeno da identificação, no qual o jovem passa a se reconhecer nas narrativas e experiências apresentadas nos conteúdos consumidos. Essa identificação pode gerar uma sensação de pertencimento, mas também pode fortalecer pensamentos disfuncionais e validar comportamentos prejudiciais.

A atuação dos algoritmos de recomendação intensifica esses impactos ao priorizar conteúdos semelhantes àqueles previamente acessados pelo usuário. Dessa forma, a exposição deixa de ser pontual e passa a ser contínua, criando um ambiente digital que reforça padrões emocionais negativos.

Do ponto de vista psicológico, essa dinâmica pode contribuir para o desenvolvimento ou agravamento de transtornos mentais, além de aumentar o risco de ideação suicida em indivíduos vulneráveis. A ausência de intervenções ou mecanismos de proteção adequados agrava ainda mais esse cenário.

Também se observa que a exposição prolongada a conteúdos sensíveis pode afetar a capacidade de regulação emocional dos jovens, tornando-os mais suscetíveis a reações impulsivas e dificuldades no enfrentamento de situações adversas. Isso impacta não apenas o bem-estar individual, mas também as relações sociais e o desempenho em outras áreas da vida.

Dessa forma, os impactos da exposição contínua a conteúdos sensíveis evidenciam a necessidade de maior atenção por parte das plataformas digitais, da sociedade e dos órgãos reguladores. A implementação de estratégias de prevenção, como moderação mais eficaz e promoção de conteúdos positivos, é fundamental para reduzir os riscos e proteger a saúde mental dos usuários.

### **3. METODOLOGIA**

A presente pesquisa adota abordagem qualitativa, do tipo exploratório-descritiva, com procedimento de revisão bibliográfica sistemática e análise documental. Foram consultadas bases indexadas como o SciELO e Google Acadêmico utilizando os descritores: "inteligência artificial", "responsabilidade penal", "algoritmos de recomendação", "conteúdos suicidogênicos", "saúde mental juvenil", "plataformas digitais". O recorte temporal abrange publicações entre 2014 e 2026. A análise documental incluiu a Lei nº 12.965/2014 (Marco Civil da Internet), a Lei nº 13.709/2018 (LGPD) e relatórios internacionais como o WHO e DataReportal. A análise dos dados foi realizada por meio da técnica de análise de conteúdo, com categorias definidas a priori: funcionamento algorítmico, vulnerabilidade juvenil, responsabilidade penal, desafios regulatórios.

### **4. CONCLUSÃO**

A presente pesquisa permitiu compreender a complexidade envolvida na responsabilização penal das plataformas digitais diante da recomendação de conteúdos suicidogênicos por meio de algoritmos. Observou-se que, embora essas tecnologias tragam avanços significativos na personalização e no acesso à informação,

também apresentam riscos relevantes, especialmente para a saúde mental de jovens usuários em situação de vulnerabilidade.

Ao longo do estudo, evidenciou-se que os algoritmos de recomendação não atuam de forma neutra, mas sim orientados por objetivos de engajamento, o que pode resultar na amplificação de conteúdos sensíveis. Essa dinâmica, associada à vulnerabilidade psicológica dos jovens, contribui para a intensificação de impactos negativos, como ansiedade, depressão e ideação suicida, reforçando a necessidade de maior atenção sobre o funcionamento dessas ferramentas.

No campo jurídico, verificou-se que a responsabilização penal das plataformas ainda enfrenta diversos desafios, especialmente no que se refere à comprovação do nexo causal e à adaptação dos institutos tradicionais do Direito Penal à realidade tecnológica. Além disso, os limites impostos por direitos fundamentais, como a liberdade de expressão, exigem uma abordagem equilibrada e criteriosa na construção de soluções normativas.

Dessa forma, conclui-se que é fundamental o avanço de regulamentações específicas, maior transparência algorítmica e a adoção de medidas preventivas por parte das plataformas digitais. A proteção da saúde mental dos usuários, sobretudo dos jovens, deve ser tratada como prioridade, exigindo a atuação conjunta de empresas, Estado e sociedade na construção de um ambiente digital mais seguro e responsável.

## **REFERÊNCIAS BIBLIOGRÁFICAS**

ALMEIDA, João Carlos de. *Inteligência artificial e responsabilidade jurídica*. São Paulo: Atlas, 2022.

BARBOSA, Mariana Silva. *Algoritmos e sociedade: impactos no comportamento humano*. Rio de Janeiro: Forense, 2021.

BRASIL. Lei nº 12.965, de 23 de abril de 2014. *Marco Civil da Internet*. Disponível em: <https://www.planalto.gov.br>. Acesso em: 10 mar. 2026.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. *Lei Geral de Proteção de Dados Pessoais (LGPD)*. Disponível em: <https://www.planalto.gov.br>. Acesso em: 12 mar. 2026.

COSTA, Renata Ferreira. *Responsabilidade civil e penal na era digital*. Belo Horizonte: Del Rey, 2019.

GOMES, Luiz Flávio. *Direito penal na sociedade tecnológica*. São Paulo: Revista dos Tribunais, 2020.

KEMP, Simon. *Digital 2023: Global Overview Report*. Disponível em: <https://datareportal.com>. Acesso em: 15 fev. 2026.

MORAES, Alexandre de. *Direitos fundamentais e internet*. São Paulo: Atlas, 2021.

PEREIRA, Lucas Andrade. *Saúde mental e redes sociais: riscos e desafios*. Porto Alegre: Artmed, 2020.

RODRIGUES, Felipe Souza. *Algoritmos, ética e regulação digital*. Curitiba: Juruá, 2023.

SILVA, Tatiana Mendes da. *Juventude, tecnologia e vulnerabilidade emocional*. São Paulo: Cortez, 2021.

WORLD HEALTH ORGANIZATION. *Suicide worldwide in 2019: global health estimates*. Disponível em: <https://www.who.int>. Acesso em: 20 jan. 2026.