

# DESINFORMAÇÃO E SEUS EFEITOS NO PÚBLICO INFANTIL: COMO A ANÁLISE AUTOMATIZADA PODE PROTEGER CRIANÇAS DA VIOLÊNCIA DIGITAL

Gisele Aparecida Lima de Oliveira, Julia Lima de Oliveira, Kleber Rocha de Oliveira  
30-11-2024

## Resumo

O ambiente digital tem se tornado cada vez mais presente na vida de crianças e adolescentes, o que torna urgente a adoção de medidas eficazes para proteger este público vulnerável. Este artigo aborda os desafios da implementação de análises automatizadas para combater desinformação, violência digital e outras ameaças. Utilizando tecnologias como inteligência artificial e aprendizado de máquina, as plataformas digitais têm buscado monitorar conteúdos prejudiciais, mas enfrentam obstáculos relacionados a questões éticas, como viés algorítmico, e à falta de transparência na coleta e uso de dados. A proteção infantil no digital não se limita a tecnologias automatizadas, sendo essencial a promoção de educação digital tanto para crianças quanto para pais e educadores, capacitando-os a identificar desinformação e proteger os jovens de riscos. Além disso, a resistência das plataformas à regulamentação e a falta de políticas públicas claras para a moderação de conteúdo são desafios que precisam ser superados. A responsabilidade de proteger crianças e adolescentes no mundo digital é compartilhada entre o poder público, as plataformas, as famílias e a sociedade. Este artigo também destaca a importância da regulação e da transparência nas práticas digitais, com o intuito de garantir a privacidade, a liberdade de expressão e a segurança dos jovens. O compromisso com a criação de um ambiente digital seguro exige uma abordagem colaborativa e multifacetada, onde a ética, a educação e a inovação tecnológica caminham juntas para a proteção integral das futuras gerações.

**Palavras-chave:** proteção infantil, educação digital, inteligência artificial, regulação, desinformação.

## INTRODUÇÃO

A crescente presença de crianças e adolescentes no ambiente digital trouxe consigo uma série de desafios relacionados à segurança, à privacidade e ao bem-estar desses indivíduos, especialmente quando expostos a conteúdos prejudiciais como desinformação, discursos de ódio, violência e abuso. A evolução das tecnologias de inteligência artificial (IA) e aprendizado de máquina (AM) tem oferecido novas oportunidades para combater esses riscos, através de análises automatizadas que podem detectar e filtrar conteúdos nocivos em tempo real. No entanto, essa inovação também impõe uma série de desafios éticos, técnicos e regulatórios que precisam ser enfrentados para garantir que tais ferramentas cumpram sua função sem violar direitos fundamentais, como a privacidade e a liberdade de expressão.

Além disso, a educação digital emerge como uma ferramenta essencial para a capacitação de pais, educadores e da própria comunidade escolar, permitindo que a sociedade como um todo participe ativamente na construção de um ambiente online mais seguro. A conscientização sobre os riscos da desinformação e a promoção do pensamento crítico são aspectos fundamentais para garantir que crianças e adolescentes possam navegar de forma segura, responsável e ética no mundo digital.

Este artigo explora os desafios e as possibilidades envolvidas na implementação de análises automatizadas para proteger crianças e adolescentes contra os riscos do ambiente digital. Destacando a necessidade de uma responsabilidade compartilhada entre o poder público, as famílias, as plataformas digitais e as comunidades.

## **1. METODOLOGIA UTILIZADA**

Este estudo adota uma abordagem qualitativa e exploratória, baseada em análise documental e revisão de literatura. Foram analisados artigos acadêmicos, relatórios institucionais e marcos legais sobre desinformação, proteção infantil no ambiente digital e o uso de inteligência artificial para mitigação de riscos. A literatura forneceu subsídios para identificar desafios técnicos e éticos na implementação de sistemas automatizados e lacunas nas políticas públicas.

O relatório do Comitê Gestor da Internet no Brasil (CGI.br), *Sistematização das Contribuições à Consulta sobre Regulação de Plataformas Digitais* (2023), aprofundou a compreensão dos riscos do uso de plataformas digitais por crianças e adolescentes. Os principais riscos mapeados foram a exploração infantil por meio de estratégias de perfilamento e algoritmos que priorizam o engajamento. As propostas de mitigação destacaram medidas regulatórias voltadas à proteção de grupos vulneráveis e a responsabilidade das plataformas na moderação de conteúdo prejudicial.

As contribuições do relatório foram articuladas com conceitos de ética digital, governança de plataformas e regulação, enfatizando a importância de práticas transparentes por parte das plataformas no tratamento de dados de crianças, bem como a aplicação de algoritmos para identificar conteúdos prejudiciais, equilibrando inovação tecnológica e proteção de direitos fundamentais. O estudo também considerou a legislação nacional, como o Marco Civil da Internet e a Lei Geral de Proteção de Dados (LGPD), além de comparações com marcos regulatórios internacionais, como a

Convenção sobre os Direitos da Criança, para propor recomendações adaptadas ao contexto brasileiro.

## **2. DESINFORMAÇÃO NO MUNDO DIGITAL: IMPACTOS E RISCOS PARA CRIANÇAS E ADOLESCENTES**

No contexto atual, a desinformação é um problema crescente, especialmente com a disseminação de informações na internet. Embora frequentemente usados de forma intercambiável, os termos "desinformação" e "fake news" têm significados distintos que precisam ser compreendidos para uma análise mais precisa do fenômeno.

Desinformação é o ato de compartilhar informações falsas, enganosas ou imprecisas, com o objetivo de manipular ou enganar o público. Ela pode ser propagada de forma deliberada ou acidental, mas, de maneira geral, sua principal característica é a intenção de distorcer a verdade. O termo está ligado a contextos políticos, sociais e econômicos, mas também pode afetar outras áreas, como saúde e segurança pública.

A desinformação digital abrange um conjunto de informações que são deliberadamente falsas ou manipuladas. Segundo Brizola e Bezerra (2018), a desinformação pode envolver conteúdos descontextualizados, fragmentados, tendenciosos, ou que simplesmente apagam a realidade para distorcer a percepção dos fatos.

A desinformação é um conceito antigo que nasce ligado a projetos militares de contrainformação e espionagem, mas extrapola para os meios de comunicação e para aparelhos privados e estatais. A desinformação pode estar presente em livros de história ou em discursos políticos, em histórias em quadrinhos ou em jornais de ampla circulação. Desinformação envolve informação descontextualizada, fragmentada, manipulada, retirada de sua historicidade, tendenciosa, que apaga a realidade, distorce, subtrai, rotula ou confunde. A desinformação não é necessariamente falsa; muitas vezes, trata-se de distorções ou partes de verdade (BRIZOLA, BEZERRA, 2018).

Fake News, por sua vez, é uma expressão popularmente associada à desinformação. Embora amplamente utilizada para descrever informações falsas, o termo se refere, especificamente, a notícias ou conteúdos criados para imitar um formato jornalístico legítimo, mas com o objetivo de enganar o leitor. A "fake news" pode ser um tipo específico de desinformação, mas a desinformação abrange um espectro mais amplo, incluindo qualquer tipo de conteúdo falso ou manipulado, não necessariamente no formato de notícias.

A desinformação é um sistema informacional que molda a opinião pública de acordo com seus interesses utilizando uma série de artifícios e mecanismos para manter a hegemonia, e isso não é nada novo. Contudo, com a velocidade da demanda de informações no meio digital e o espalhamento de informações com uma rapidez multiplicada, a desinformação ganha vulto com o fenômeno de circulação de fake News (PRADO, 2022).

Além disso, pode-se ter a Informação Desatualizada refere-se a conteúdos que, embora possam ter sido verdadeiros em algum momento, se tornaram obsoletos ou irrelevantes devido a mudanças nas circunstâncias, dados ou descobertas. Diferente da desinformação intencional, a informação desatualizada pode ser propagada de boa-fé, mas ainda assim pode levar à confusão ou interpretação errônea.

Informação Fora de Contexto é uma outra forma de distorção, onde uma informação verdadeira é retirada de seu contexto original e utilizada para gerar um novo significado, que pode ser enganoso ou manipulado.

Enquanto a desinformação e as fakes news geralmente têm uma intenção manipulativa clara, a informação desatualizada e fora de contexto podem surgir sem má intenção, mas igualmente causam distorções no entendimento do público.

## **2.1. Desinformação no Contexto Digital**

Com o avanço das tecnologias digitais e a popularização das redes sociais, a desinformação se espalha de forma rápida e eficaz. Plataformas como Facebook, X (antigo Twitter), WhatsApp e TikTok possibilitam a difusão de conteúdos falsos em uma escala global. A velocidade e o alcance dessas plataformas tornam difícil rastrear e corrigir informações falsas antes que elas causem impactos negativos.

Nesse cenário, o público infantil e juvenil é particularmente vulnerável. As crianças e adolescentes, muitas vezes, não têm a capacidade crítica necessária para avaliar a veracidade das informações que consomem. Eles estão em uma fase de desenvolvimento cognitivo e emocional, sendo mais suscetíveis a influências externas e a informações que correspondam às suas crenças ou expectativas.

## **2.2. A Vulnerabilidade das Crianças e Adolescentes à Desinformação**

O estado de vulnerabilidade das crianças e adolescentes no ambiente digital deve ser abordado com seriedade. Essa faixa etária ainda está em processo de desenvolvimento cognitivo, emocional e social, o que a torna especialmente sensível à manipulação digital. Fatores como a busca por pertencimento social, inseguranças e

curiosidade natural podem levar jovens a consumir conteúdo sem discernir sua veracidade.

A facilidade de compartilhamento rápido e a falta de controle sobre o que é disseminado em grupos e mensagens privadas contribuem para a difusão de fake news. Esse fenômeno é especialmente perigoso entre os jovens, que tendem a consumir e compartilhar informações sem verificar sua veracidade, alimentados pela dinâmica de consumo acelerado de conteúdos, muitas vezes sem o tempo necessário para uma análise crítica.

O excesso de informações geradas e compartilhadas, aliado ao uso de algoritmos que direcionam conteúdos, cria uma "bolha de desinformação", onde os usuários se veem rodeados por conteúdos que reafirmam suas crenças, sem questionamento.

O perigo da desinformação digital vai além de uma simples falha de discernimento. Em muitos casos, a propagação de conteúdos violentos, xenofóbicos, sexistas ou discriminatórios pode resultar em danos psicológicos e sociais significativos para os jovens, afetando seu desenvolvimento emocional e as relações interpessoais.

### **2.3. Impactos Da Desinformação No Público Infantil**

As crianças estão em processo de construção de seu entendimento do mundo, o que as torna suscetíveis à absorção de informações errôneas ou enganosas. Além disso, o uso da internet por crianças e adolescentes cresce a cada ano.

Os efeitos da desinformação digital sobre as crianças podem ser variados e profundos, como:

- **Risco de Violência Digital**

A violência digital refere-se a qualquer forma de abuso ou agressão realizada por meio de tecnologias digitais, como internet, redes sociais, aplicativos e dispositivos eletrônicos.

Informações falsas e manipuladas podem ser direcionadas para criar campanhas de difamação, assédio online, bullying ou até mesmo incitação à violência. Campanhas de ódio e discursos violentos muitas vezes são alimentados por desinformação, criando um ciclo de agressões entre os jovens.

Crianças e adolescentes que passam muito tempo na internet correm maior risco de cyberbullying, o que pode levar à depressão, ansiedade e até suicídio. Essa forma de intimidação é conceituada no art. 2º, parágrafo único da Lei n. 13.185, de 06 de novembro de 2015, sendo incluída no Código Penal, em seu art. 146-A, pela Lei 14.811, de 12 de janeiro de 2024.

**Intimidação sistemática (bullying)** [Art. 146-A](#). Intimidar sistematicamente, individualmente ou em grupo, mediante violência física ou psicológica, uma ou mais pessoas, de modo intencional e repetitivo, sem motivação evidente, por meio de atos de intimidação, de humilhação ou de discriminação ou de ações verbais, morais, sexuais, sociais, psicológicas, físicas, materiais ou virtuais: Pena - multa, se a conduta não constituir crime mais grave.

**Intimidação sistemática virtual (cyberbullying)** Parágrafo único. Se a conduta é realizada por meio da rede de computadores, de rede social, de aplicativos, de jogos **on-line** ou por qualquer outro meio ou ambiente digital, ou transmitida em tempo real: Pena - reclusão, de 2 (dois) anos a 4 (quatro) anos, e multa, se a conduta não constituir crime mais grave.

À medida que crianças e adolescentes se tornam mais visíveis nas redes sociais, podem atrair atenção indesejada de estranhos, de colegas, o que pode ser difícil para eles lidarem.

Segundo relatório da UNICEF, 2024, sobre crianças e ambientes digitais, as vítimas do cyberbullying estão mais propensas a usar álcool e drogas, não ir à escola, experienciar bullying pessoalmente, tirar notas baixas e experienciar baixa autoestima, além de outros tipos de problemas de saúde. As vítimas desse tipo de abuso expressam a severidade do problema, que, em alguns casos, leva ao suicídio ou a pensamentos de suicídio.

Além disso, quando uma criança ou adolescente se torna alvo de um conteúdo falso ou distorcido sobre si mesmo, isso pode desencadear um sofrimento emocional intenso. A violência digital, alimentada pela desinformação, pode afetar a autoestima, as relações interpessoais e, em casos extremos, levar a atitudes autodestrutivas. A facilidade com que os conteúdos se espalham nas redes sociais e a falta de controle sobre o que é compartilhado tornam o risco de ataques virtuais ainda mais presente.

Essa violência digital não se restringe apenas ao conteúdo ofensivo ou prejudicial diretamente dirigido a um indivíduo, mas também pode se manifestar em comportamentos agressivos mais amplos, como incitação ao ódio contra grupos específicos. O impacto disso pode ser devastador para o desenvolvimento social e emocional dos jovens, levando a uma cultura de intolerância e hostilidade.

- **Preocupação com *Deepfake***

A tecnologia de *deepfake*, que permite criar vídeos e áudios falsificados de forma realista, apresenta uma nova fronteira na desinformação digital. Para crianças e adolescentes, que já estão em um estágio de aprendizado e reconhecimento de mídias digitais, o risco de serem manipulados por esses conteúdos é significativo.

Os *deepfakes* representam uma forma de manipulação digital que pode distorcer a verdade e comprometer a integridade das informações.

Uma das principais preocupações éticas relacionadas aos deepfakes é o potencial de causar danos às vítimas. Quando uma pessoa é retratada de forma falsa em um vídeo ou imagem deepfake, isso pode ter sérias consequências para sua reputação e bem-estar emocional, especialmente se tratando de crianças e adolescentes, por pertencerem a um grupo vulnerável, devido a estarem ainda em desenvolvimento mental e físico. (OLIVEIRA; OLIVEIRA, 2024, p.49)

*Deepfakes* podem ser usados para criar vídeos ou áudios falsos de figuras públicas, amigos ou familiares, levando os jovens a acreditarem em informações que são completamente fabricadas.

Essa forma de manipulação pode ser particularmente prejudicial, pois as crianças muitas vezes não possuem as habilidades necessárias para identificar se um conteúdo é real ou alterado. Além disso, o uso de *deepfakes* para espalhar desinformação pode ser utilizado em campanhas de difamação, assédio online ou até para causar confusão e desinformação em grandes escalas, o que pode afetar a saúde mental e emocional dos jovens.

- **Publicidade Infantil (Consumo e Aparência)**

A publicidade dirigida a crianças nas plataformas digitais é um fenômeno que tem gerado preocupações significativas. As crianças são especialmente vulneráveis à manipulação publicitária, uma vez que ainda estão desenvolvendo suas capacidades críticas de julgamento e discernimento.

Os pressupostos da propaganda, área dedicada a conquistar determinado público – como, por exemplo, quando faz o mapeamento do gosto de um grupo de pessoas para conhecê-lo e assim por influenciá-lo – demonstra ser pertinentes para direcionar e manipular. (PRADO, 2022)

A pesquisa do Instituto Alana, 2020, sobre o programa Criança e Consumo revelou que o consumo excessivo de publicidade leva as crianças a desenvolverem um

comportamento consumista, desejando produtos que não têm necessidade real, mas que são constantemente apresentados como essenciais.

Esse tipo de publicidade é particularmente preocupante porque muitas vezes promove valores superficiais, como a busca por uma aparência idealizada ou o consumo excessivo de produtos alimentícios pouco saudáveis. O documentário *Criança, a alma do negócio*, da cineasta Estela Renner, ilustra como as crianças, influenciadas por essas mensagens publicitárias, podem se afastar de hábitos saudáveis, como a alimentação balanceada, e adotar padrões de consumo que não condizem com sua realidade socioeconômica, levando a comportamentos impulsivos e insustentáveis.

Para mencionar esse efeito da mídia e a publicidade em relação às crianças, a fala de uma menina não identificada no documentário, de aproximadamente 7 anos de idade a respeito do sentimento em relação a essas publicidades:

Mexe muito até com o coração, porque quando a pessoa vê alguma coisa ali e quer, gostou, única coisa que a pessoa quer é ir lá na loja comprar. Eu acho que é mais fácil a criança ir lá pedir pra mãe e a mãe não ter, aí sim o coração da criança fica até com vontade de chorar a criança tem. (RENNER, 2013, 7:35).

Além disso, as pressões sociais geradas por essas campanhas publicitárias podem afetar a autoestima das crianças, já que elas são constantemente expostas à ideia de que a felicidade está atrelada à aquisição de novos produtos ou à conformidade com padrões estéticos impostos pela mídia. Esse ciclo pode gerar um vazio emocional e uma sensação de inadequação, especialmente quando as crianças não conseguem alcançar esses padrões.

- **Discursos de Ódio**

O crescimento da desinformação e da violência digital está intimamente relacionado ao aumento de discursos de ódio na internet. Em um ambiente digital cada vez mais acessível e permeado por interações virtuais, as crianças e adolescentes estão particularmente expostos a mensagens discriminatórias, preconceituosas e violentas que podem incitar intolerância e discriminação.

Na Campanha #PrometoPausar: Combata a desinformação nas redes sociais lançada pela Organização das Nações Unidas, define o discurso de ódio como:

Discurso de ódio é qualquer tipo de comunicação que ataca ou usa a linguagem pejorativa ou discriminatória com referência a uma pessoa ou a

um grupo em sua religião, etnia, nacionalidade, raça, cor, descendência, gênero, ou outro fator de identificação. (ONU, 2024).

Nesse mesmo documento, o secretário-geral da ONU, António Guterres, afirmou:

O discurso de ódio é um sinal de alerta de discriminação, abuso, violência, conflito e até mesmo de crimes contra a humanidade. Já vimos isso acontecer desde a Alemanha nazista até Ruanda, Bósnia e outros lugares. Não existe um nível aceitável de discurso de ódio; todos nós devemos trabalhar para erradicá-lo completamente (ONU, 2024).

A disseminação de discursos de ódio, especialmente aqueles ligados a etnia, gênero, religião, orientação sexual e outras diferenças, pode ter consequências profundas para o público jovem, pois influencia diretamente na construção de suas identidades e visão de mundo.

O impacto dos discursos de ódio não é limitado ao mundo digital; eles podem transbordar para o mundo físico e desencadear ações violentas. No contexto dos jovens, isso é particularmente preocupante, pois os adolescentes estão em uma fase de formação de valores e crenças, e frequentemente são mais suscetíveis a adotar os discursos e comportamentos que encontram online.

A incitação ao ódio pode, em muitos casos, levar a atos infracionais, conforme o previsto no Estatuto da Criança e do Adolescente (ECA), caso os menores de 18 anos se envolvam em práticas discriminatórias ou violentas contra outros indivíduos.

É importante notar que, embora os discursos de ódio sejam prejudiciais e muitas vezes considerados como um crime, ao envolverem menores de idade, esses comportamentos são classificados como atos infracionais, não crimes propriamente ditos. Isso implica que, embora o ato de discriminar e incitar o ódio seja ilegal, a resposta legal para adolescentes se dá em um contexto educativo, com o objetivo de corrigir a conduta e não apenas punir.

No entanto, o desafio é que, ao serem expostos a esse tipo de discurso, as crianças e os adolescentes podem internalizar comportamentos discriminatórios e agressivos, perpetuando um ciclo de violência que se estende para as suas interações físicas e digitais.

- **Jogos de Azar**

A proliferação de jogos de azar online, especialmente entre jovens, é outra grande preocupação. O uso de plataformas digitais para promover jogos ilegais, como apostas e jogos de azar, tem se tornado mais comum, principalmente com o incentivo de influenciadores digitais, que têm grande poder de persuasão sobre seus seguidores. Muitos desses influenciadores promovem jogos de azar em suas redes sociais, criando campanhas que atraem especialmente crianças e adolescentes, que estão em uma fase de desenvolvimento e buscam diversão e pertencimento social.

Esses jogos, além de serem ilegais em muitos países, incluem uma série de riscos, como o envolvimento com práticas fraudulentas e a exploração econômica dos jovens. A propagação dessas atividades por influenciadores é uma forma de manipulação, onde, ao direcionar jovens para essas plataformas, eles acabam incentivando comportamentos compulsivos, viciantes e ilegais.

A busca por dinheiro fácil e a possibilidade de ganhos rápidos pode atrair os jovens a participarem desses jogos, sem compreenderem plenamente as consequências financeiras e psicológicas que podem advir dessa prática.

Para muitos jovens, o jogo se torna uma forma de buscar aceitação social ou de lidar com frustrações e inseguranças. No entanto, a dependência de jogos de azar pode levar ao desenvolvimento de sérios problemas emocionais e financeiros, incluindo o endividamento e o agravamento de quadros de ansiedade e depressão. A educação sobre os riscos desses jogos e a promoção de uma internet mais segura são essenciais para proteger as crianças e adolescentes desse tipo de exploração.

- **Exploração e Abuso (Sexual e Comercial)**

Outro ponto crítico relacionado à desinformação e ao mundo digital é a crescente preocupação com a exploração e o abuso infantil. A internet tem se tornado um espaço vulnerável para crianças e adolescentes serem manipulados e explorados, seja sexualmente ou comercialmente. A exploração sexual infantil online, por exemplo, envolve o uso de imagens, vídeos e outras formas de conteúdo manipuladas ou forjadas, muitas vezes sem o consentimento da criança, para fins ilícitos.

Além disso, a exploração comercial também se apresenta de forma crescente na internet. Crianças e adolescentes podem ser vítimas de tráfico de imagens, onde suas fotos e vídeos são compartilhados sem sua permissão, sendo vendidos ou utilizados para

lucro de criminosos cibernéticos. Esse tipo de abuso tem consequências devastadoras para o desenvolvimento psicológico das vítimas, pois a exposição a esse conteúdo pode afetar a formação de sua identidade e autoestima. As crianças podem ser manipuladas por adultos mal-intencionados a se envolverem em atividades perigosas, como a produção de material pornográfico ou até mesmo a participação em esquemas criminosos.

O impacto desse abuso é ainda mais profundo, pois muitas crianças e adolescentes não têm a maturidade necessária para compreender a gravidade da situação e a maneira de se proteger.

### **3. ANÁLISES AUTOMATIZADAS COM APRENDIZADO DE MÁQUINA E INTELIGÊNCIA ARTIFICIAL (IA)**

A Inteligência Artificial (IA) refere-se a sistemas ou máquinas que são projetados para imitar o comportamento humano, aprendendo e tomando decisões de forma autônoma, sem intervenção direta de um ser humano. Isso pode envolver tarefas como reconhecimento de voz, análise de imagens, tradução de idiomas ou até mesmo a recomendação de conteúdo online.

Já o aprendizado de máquina (machine learning) é uma subárea da IA que permite que as máquinas "aprendam" com dados. Isso significa que, em vez de serem programadas explicitamente para realizar uma tarefa, as máquinas usam grandes volumes de dados para identificar padrões e fazer previsões ou decisões. Por exemplo, quando você recebe sugestões de vídeos no YouTube ou anúncios personalizados no Facebook, isso é resultado de algoritmos de aprendizado de máquina.

O processo de aprendizado de máquina é baseado em dados. Esses dados podem ser imagens, textos, vídeos ou até interações de usuários com plataformas. O sistema de IA analisa esses dados, encontra padrões e "aprende" com eles para melhorar suas previsões ao longo do tempo. Em termos simples, é como um estudante que, ao fazer exercícios e receber correções, vai aprendendo e se aprimorando com a prática.

Diversas plataformas digitais já adotaram medidas para proteger crianças e adolescentes de conteúdos prejudiciais, incluindo desinformação, discursos de ódio, jogos de azar e abuso sexual online. Vamos ver alguns exemplos:

- Instagram e Facebook: Essas plataformas utilizam inteligência artificial para detectar e bloquear automaticamente conteúdos como discursos de ódio, pornografia infantil, bullying e disseminação de fake news. Além disso, ambas as plataformas têm filtros de segurança que permitem que os usuários definam quem pode interagir com eles e quais tipos de conteúdo podem visualizar. O Facebook, por exemplo, utiliza IA para identificar conteúdos violentos ou incitadores de ódio e removê-los. O Instagram também implementou ferramentas de moderação de comentários, com a possibilidade de ocultar comentários ofensivos automaticamente.

- WhatsApp: Embora seja mais difícil aplicar a moderação de conteúdos em mensagens privadas, o WhatsApp tem adotado mecanismos para impedir a disseminação de fake news. Além disso, a plataforma tem trabalhado para identificar e remover números suspeitos de bots que possam estar envolvidos em campanhas de desinformação. A plataforma também possui limites no reencaminhamento de mensagens, o que ajuda a reduzir a velocidade com que informações falsas podem ser compartilhadas.

- Jogos Online: Jogos como Minecraft e Roblox têm implementado sistemas de IA e moderação humana para garantir um ambiente seguro para os jogadores. O Minecraft, por exemplo, possui filtros automáticos para detectar palavras e comportamentos tóxicos, e jogadores podem denunciar abuso ou conteúdo impróprio. O Roblox também tem implementado um sistema de moderação automática que detecta linguagem inadequada e comportamento abusivo. Além disso, ambos os jogos têm configurações de controles parentais que permitem aos pais monitorarem as interações dos filhos.

A precisão dos sistemas de IA depende muito da qualidade e diversidade dos dados usados para treiná-los. Portanto, é fundamental que as plataformas e desenvolvedores adotem práticas éticas desde a criação dos algoritmos, implementando mecanismos de controle para garantir que os conteúdos propagados não sejam discriminatórios, violentos ou ilegais.

#### **4. DESAFIOS NA IMPLEMENTAÇÃO DAS ANÁLISES AUTOMATIZADAS: ÉTICA, REGULAMENTAÇÃO E EDUCAÇÃO DIGITAL**

A utilização de análises automatizadas para combater a desinformação e proteger o público infantil oferece soluções promissoras, mas também enfrenta uma série de desafios éticos, técnicos e sociais.

Estas tecnologias, embora eficazes em detectar e remover conteúdos prejudiciais, como discursos de ódio, violência e abuso sexual, não são isentas de problemas que precisam ser resolvidos, desde a sua criação e desenvolvimento até a sua aplicação prática.

A primeira grande questão envolve a ética no desenvolvimento e uso das tecnologias de monitoramento. O uso de inteligência artificial (IA) e aprendizado de máquina (AM) para analisar o comportamento online e filtrar conteúdos tem gerado preocupações sobre a privacidade e o uso indevido de dados pessoais, especialmente os de crianças e adolescentes.

A Constituição Federal, o Marco Civil da Internet, A Lei Geral de Proteção de Dados (LGPD), em conjunto estabelecem que as plataformas devem garantir a privacidade dos usuários e impedir que dados sensíveis sejam coletados sem o devido consentimento. Em se tratando de informações sobre crianças e adolescentes, deve ser observado em conjunto com o Estatuto da Criança e do Adolescente, Lei n. 8.069/1990 e a Convenção sobre os Direitos da Criança, visando sempre a proteção integral e seu melhor interesse.

Em alguns casos, a falta de transparência em relação ao funcionamento desses sistemas pode resultar em resistência à sua implementação, prejudicando a eficácia das ferramentas e a confiança do público nas plataformas.

Além disso, o uso de tecnologias de IA e AM pode ser antiético se esses sistemas forem treinados com dados enviesados. O viés discriminatório nos dados de treinamento pode fazer com que algoritmos reproduzam preconceitos históricos relacionados a raça, gênero, classe social e outras características, afetando negativamente grupos vulneráveis, como mulheres negras ou crianças de comunidades marginalizadas. Isso é um desafio contínuo para as plataformas, que precisam garantir que seus sistemas sejam justos e imparciais.

Apesar da crescente pressão para que as plataformas digitais adotem análises automatizadas, muitas empresas hesitam em aplicar essas tecnologias de forma robusta.

Essa resistência se deve, em parte, ao equilíbrio delicado entre censura e liberdade de expressão. As plataformas têm o desafio de proteger as crianças e combater conteúdos nocivos, sem infringir o direito dos usuários à liberdade de expressão e ao acesso à informação.

A responsabilidade compartilhada pela proteção de crianças e adolescentes, conforme o artigo 227 da Constituição Federal brasileira, estabelece que a proteção integral dos direitos das crianças é uma responsabilidade de todos: do poder público, das famílias, da sociedade e das plataformas digitais.

As plataformas têm a obrigação legal de garantir que seus sistemas, serviços e algoritmos não prejudique os direitos dos menores, especialmente no que se refere à exposição a conteúdos violentos, discurso de ódio, e a exposição excessiva de imagens.

A responsabilidade dos provedores de aplicações de internet, entendidos como aqueles que utilizam acesso à internet para prestar serviços, a exemplo do Instagram, Facebook, TikTok, WhatsApp, Google Chrome, Firefox, poderão ser responsabilizados na medida descrita no art. 21 da Lei n.º 12.965/2014, conhecida como Marco Civil da Internet, que aduz o seguinte:

Art. 21. O provedor de aplicações de internet que disponibilize conteúdo gerado por terceiros será responsabilizado subsidiariamente pela violação da intimidade decorrente da divulgação, sem autorização de seus participantes, de imagens, de vídeos ou de outros materiais contendo cenas de nudez ou de atos sexuais de caráter privado quando, após o recebimento de notificação pelo participante ou seu representante legal, deixar de promover, de forma diligente, no âmbito e nos limites técnicos do seu serviço, a indisponibilização desse conteúdo.

O Conselho Nacional dos Direitos da Criança e do Adolescente (CONANDA) tem regulado as plataformas digitais, exigindo que as empresas de tecnologia adotem práticas de monitoramento preventivo para identificar e mitigar riscos à saúde mental, à autoestima e ao bem-estar físico e emocional das crianças e adolescentes

A educação digital desempenha um papel importante na proteção de crianças e adolescentes contra os riscos da desinformação e da violência online. Para que as ferramentas de análise automatizada sejam eficazes, é necessário que pais, educadores e comunidade estejam preparados para educar as crianças sobre os riscos do ambiente digital e como proteger sua privacidade e segurança online.

A alfabetização midiática deve ser incorporada aos currículos escolares, ensinando desde cedo as crianças a identificar fontes confiáveis de informação, questionar o que leem online e desenvolver o pensamento crítico.

O dispositivo de suma importância para o tema é o art. 29, §1º da Resolução n. 245 do CONANDA:

Art. 29 O Poder Público e as empresas provedoras de produtos e serviços digitais devem promover ações de sensibilização sobre os direitos e riscos a crianças e adolescentes no ambiente digital, focando, particularmente, nos setores cujas práticas tenham impacto direto ou indireto nas crianças e adolescentes.

§1º Deve ser estimulado o desenvolvimento de ações educacionais para crianças e adolescentes, familiares, responsáveis e cuidadores, autoridades públicas e sociedade em geral, para ampliar o conhecimento sobre os direitos de crianças e adolescentes, especificamente quanto aos benefícios e riscos associados com produtos e serviços digitais, bem como ampliar a autonomia e o senso crítico, individual e coletivos, sobre decisões relacionadas ao desenvolvimento e gestão do ambiente digital.

Campanhas de conscientização e treinamento para educadores também são essenciais. Essas ações devem envolver professores, pais e responsáveis no processo de capacitação constante sobre como identificar desinformação, discurso de ódio e outros conteúdos prejudiciais. Além disso, deve-se estimular a participação dos pais em fóruns e comunidades online dedicadas à segurança digital, onde possam compartilhar experiências e boas práticas.

## **CONCLUSÃO**

A proteção de crianças e adolescentes no ambiente digital é uma responsabilidade compartilhada que exige um esforço conjunto entre o poder público, as plataformas digitais, as famílias e a sociedade em geral.

Embora a tecnologia, como a inteligência artificial e o aprendizado de máquina, tenha mostrado um potencial considerável para detectar e bloquear conteúdos nocivos, sua implementação não é isenta de desafios. Questões éticas, como o viés nos algoritmos e a falta de transparência no uso de dados, bem como a resistência das plataformas em adotar medidas de monitoramento mais rigorosas, ainda precisam ser enfrentadas com urgência.

Além disso, a proteção infantil no ambiente digital não pode se restringir apenas a ferramentas automatizadas. A educação digital desempenha um papel central, pois capacita crianças, pais, educadores e comunidades a desenvolverem habilidades

críticas para discernir desinformação, compreender os riscos da internet e agir de forma proativa para proteger os jovens. A alfabetização midiática e a conscientização sobre os perigos do ambiente digital são fundamentais para a construção de um espaço online mais seguro e saudável.

Portanto, a criação de políticas públicas eficazes, a regulamentação rigorosa das plataformas e a promoção de práticas de monitoramento responsáveis são essenciais para garantir que as ferramentas tecnológicas não apenas protejam as crianças, mas também respeitem seus direitos fundamentais, como a privacidade e a liberdade de expressão.

A colaboração entre os diversos setores da sociedade é imprescindível para que possamos avançar na construção de um ecossistema digital que favoreça a proteção integral da criança e do adolescente, conforme preconizado pelo Estatuto da Criança e do Adolescente (ECA) e pela Constituição Federal Brasileira.

Assim, o desafio é grande, mas não insuperável. Ao fomentar um diálogo aberto e contínuo entre reguladores, plataformas, especialistas e a sociedade, podemos criar um ambiente digital mais ético, transparente e seguro para as futuras gerações. É um compromisso que exige ação em várias frentes: ética, educação, regulação e inovação tecnológica, todas voltadas para um único objetivo: a proteção e o bem-estar das crianças e adolescentes no mundo digital.

## **Referências**

ALANA. Publicidade Infantil na TV paga. 2020. Disponível em: <https://alana.org.br/material/publicidade-infantil-na-tv-paga/>. Acesso em: 25 nov. 2024.

BRASIL. Código Penal. Decreto-Lei nº 2.848, de 07 de dezembro de 1940. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/decreto-lei/del2848compilado.htm](https://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm). Acesso em: 20 fev. 2024.

BRASIL. Constituição da República Federativa do Brasil de 1988. Brasília, DF: Presidência da República, [2023]. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm). Acesso em: 20 nov. 2024.

BRASIL. Lei nº 8.069, de 13 de julho de 1990. Dispõe sobre o Estatuto da Criança e do Adolescente e dá outras providências. Brasília, DF: Presidência da República, [1990].

Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/l8069.htm](https://www.planalto.gov.br/ccivil_03/leis/l8069.htm). Acesso em: 05 set. 2024.

BRASIL. Lei nº 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. *Diário Oficial da União*, Brasília, DF, 2014. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/l12965.htm](https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm). Acesso em: 05 nov. 2024.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais. Brasília, DF: *Diário Oficial da União*, 15 ago. 2018. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/l13709.htm](https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm). Acesso em: 20 nov. 2024.

BRASIL. Projeto Lei nº 2.628, de 2022. Dispõe sobre a proteção de crianças e adolescentes em ambientes digitais. 2022. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/154901>. Acesso em: 20 nov. 2024.

BRISOLA, A.; BEZERRA, A. Desinformação e circulação de “fake news”: distinções, diagnóstico e reação. *XIX Encontro Nacional de Pesquisa em Ciência da Informação (XIX ENANCIB)*, Brasil, out. 2018. Disponível em: [http://enancib.marilia.unesp.br/index.php/XIX\\_ENANCIB/xixenancib/paper/view/1219/1636](http://enancib.marilia.unesp.br/index.php/XIX_ENANCIB/xixenancib/paper/view/1219/1636). Acesso em: 26 nov. 2024.

CGI - COMITÊ GESTOR DA INTERNET NO BRASIL. Sistematização das contribuições à consulta sobre regulação de plataformas digitais. São Paulo: CGI.br, 2023. Disponível em: <https://www.nic.br/publicacoes/>. Acesso em: 25 nov. 2024.

CONANDA. Resolução n. 245, de 05 de abril de 2024. Dispõe sobre os direitos das crianças e adolescentes em ambiente digital. 2024. Disponível em: <https://www.gov.br/participamaisbrasil/blob/baixar/48630>. Acesso em: 24 nov. 2024.

OLIVEIRA, G.A.L.; OLIVEIRA, J.L. Ética e Responsabilidade na Criação e Disseminação de Deepfakes: um estudo de caso de cyberbullying com o uso de inteligência artificial. *Revista Científica Virtual*, São Paulo, v. 46, p. 46-56, 2024. Disponível em: [https://issuu.com/esa\\_oabsp/docs/edic\\_a\\_o\\_46](https://issuu.com/esa_oabsp/docs/edic_a_o_46). Acesso em: 19 set. 2024.

ONU. #PrometoPausar: Combata a desinformação nas redes sociais. 2024. Disponível em: <https://brasil.un.org/pt-br/240171-prometopausar-combata-desinformacao-nas-redes-sociais>. Acesso em: 28 nov. 2024.

PRADO, Magaly. Fake News e inteligência artificial: o poder dos algoritmos na guerra da desinformação. São Paulo: Edições 90, 2022.

RENNER, Estela. Documentário Criança, a alma do negócio. Exibição no YouTube, 2013. Disponível em: <https://www.youtube.com/watch?v=ur9lIf4RaZ4>. Acesso em: 27 nov. 2024.

UNICEF. Convenção sobre os Direitos da Criança. Disponível em: <https://www.unicef.org/brazil/convencao-sobre-os-direitos-da-crianca>. Acesso em: 02 fev. 2024.

UNICEF. Cyberbullying: O que é e como pará-lo. O que os adolescentes querem saber sobre cyberbullying. 2024. Disponível em: <https://www.unicef.org/brazil/cyberbullying-o-que-eh-e-como-para-lo>. Acesso em: 02 nov. 2024.