

VULNERABILIDADE INFANTOJUVENIL NO CASO CHARACTER.AI: DESAFIOS JURÍDICOS NA ERA DAS INTELIGÊNCIAS ARTIFICIAIS

Raíssa Sabóia Almeida¹

Ana Paola de Castro e Lins²

¹ Graduanda em Direito na Universidade Christus (Unichristus).

² Professora da Universidade Christus (Unichristus) e do Centro Universitário Farias Brito (FB UNI). Doutora em Direito Constitucional pela Universidade de Fortaleza (Unifor).

RESUMO

Trata-se de pesquisa acerca das condutas lesivas emergentes de sistemas autônomos, a partir do estudo de caso *character.ai*, plataforma acusada de contribuir para o suicídio de um adolescente de catorze anos em 2024, bem como de outras situações prejudiciais envolvendo essa startup e outras máquinas, evidenciando lacunas no ordenamento jurídico e a vulnerabilidade infantojuvenil nesse ambiente. De caráter bibliográfico, qualitativo, descritivo e exploratório, a pesquisa se fundamenta na análise de obras doutrinárias, legislações nacionais e internacionais, artigos científicos e reportagens. Os resultados demonstram a insuficiência das normas vigentes para lidar com a complexidade das relações geradas pelo avanço de tecnologias com diferentes graus de inteligência artificial, especialmente aquelas de caráter antropomorfizado.

PALAVRAS-CHAVE

Inteligência artificial – Criança e adolescente – Vulnerabilidade infantojuvenil – Responsabilidade civil – Plataformas digitais.

SUMÁRIO

Introdução. Metodologia. Resultados e discussões. Conclusão. Referências.

REFERÊNCIA: ALMEIDA, Raíssa Sabóia; LINS, Ana Paola de Castro e. Vulnerabilidade infantojuvenil no caso *character.ai*: desafios jurídicos na era das inteligências artificiais. **Res Severa Verum Gaudium**, Porto Alegre, v. 11, n. 1bis, p. 387-394, jul. 2026.

INTRODUÇÃO

Casos³ envolvendo a exposição de pessoas, principalmente crianças e adolescentes, a conteúdos sexuais impróprios, a conselhos psicológicos indevidos ou não fundamentados, ou mesmo ao desenvolvimento de dependência emocional com as máquinas, ganharam destaque na mídia nos últimos anos.

A denúncia ocorrida em dezembro de 2024 contra o *character.ai* (*c.ai*), uma das *startups* de inteligência artificial mais promissoras do Vale do Silício, que supostamente teria incentivado um adolescente de catorze anos a tirar a própria vida, bem como outros relatos envolvendo a plataforma, exemplifica os dilemas contemporâneos e revela uma demanda relevante para a atuação judiciária e legislativa em um período de contínuas e aceleradas transformações geradas pelo universo digital.

Inicialmente, apresenta-se o cenário atual de consolidação das inteligências artificiais generativas, destacando-se seu público consumidor majoritário e suas limitações a partir das leis de Asimov. Em seguida, descreve-se o caso estudado, analisando a vulnerabilidade infantojuvenil nas interações com *chatbots*, bem como a crescente demanda judicial relacionada a esse contexto.

³ Agradecemos à Universidade Christus, pelo imprescindível apoio à pesquisa, e à orientadora, professora Ana Paola de Castro e Lins, pela mentoria brilhante e inspiradora na construção deste saber.

Por fim, explora-se a complexidade de se definir a responsabilidade civil em ambientes de IA, especialmente no que se refere à identificação do autor da lesão.

METODOLOGIA

Utilizou-se de pesquisa bibliográfica, quanto às fontes, por meio da análise e leitura de obras, artigos em periódicos científicos, legislação, e notícias; qualitativa, quanto à abordagem do problema; e descritiva e exploratória, quanto aos objetivos de examinar as demandas atuais acerca da responsabilização civil de inteligências artificiais autônomas, a partir de um caso concreto de vulnerabilidade infantojuvenil.

RESULTADOS E DISCUSSÕES

Nos primórdios da robótica, Isaac Asimov, em sua obra de ficção científica *Eu, Robô* (1950), desenvolveu a ideia de uma codificação para as máquinas baseada em três leis: (1) um robô não pode ferir um ser humano, nem, por omissão, permitir que um ser humano sofra algum dano; (2) um robô deve obedecer às ordens que lhe forem dadas por seres humanos, exceto quando tais ordens entrarem em conflito com a Primeira Lei; e (3) um robô deve proteger sua própria existência, desde que essa proteção não entre em conflito com a Primeira ou a Segunda Leis (Asimov, 2014).

Alan Turing, no mesmo período, defendia, pela noção de *Child Machine*, a possibilidade de usar o aprendizado como meio de impulsionar um sistema mais simples até atingir a inteligência em nível humano (Bostrom, 2014). Existiria, assim, por meio de uma limitação efetivada pelos próprios desenvolvedores, um certo controle da atuação de suas criações, seguindo a moral e a ética vigentes. Nos anos seguintes, essa noção dos limites impostos no código, de modo a proteger os direitos humanos tutelados, se perpetuou no universo acadêmico, inspirando centenas de cientistas na construção de tecnologias detentoras de graus variados de inteligência artificial ou autonomia (Haenlein; Kaplan, 2019).

Entretanto, apesar dos esforços para manter a ética e a segurança visada desde Asimov, é imprescindível considerar os desafios impostos por esse cenário outrora apenas futurístico, principalmente no que tange ao dever de proteção familiar e a tutela e responsabilização no âmbito judicial no âmbito da IA generativa.

A priori, é relevante conceituar o termo função executiva (FE). De acordo com Klarin *et al.* (2024, p. 2), trata-se de “um conjunto de processos cognitivos, incluindo memória de trabalho, inibição e flexibilidade cognitiva”, essenciais para o planejamento, concentração e atenção, sendo relevantes para o desempenho acadêmico, especialmente no desenvolvimento estudantil inicial. Segundo essa mesma análise, estudantes tendem a utilizar a inteligência artificial com frequência em tarefas que desafiam sua FE, deixando de exercitá-la por conta própria.

Para além do uso no contexto acadêmico, sua utilização como forma de entretenimento interativo ou em perspectiva antropomorfizadora, que atribui características, emoções ou comportamentos humanos à tecnologia, também se mostra relevante, podendo gerar formas ainda pouco conhecidas de vulnerabilidade no campo social.

No dia 12 de dezembro de 2024, a character.ai (c.ai) anunciou novas medidas de segurança para proteger usuários menores de idade, após ter sido acusada e processada por uma mãe que afirma que a plataforma contribuiu para o suicídio de seu filho de catorze anos (Presse, 2024). Criada em 2022 e disponível em site e aplicativo, a IA é estruturada de forma semelhante a uma rede social, permitindo que qualquer pessoa crie seus próprios “personagens” ou interaja com aqueles desenvolvidos pela comunidade: “a plataforma concede acesso para os mais variados usuários, entre os quais estão crianças e adolescentes, sendo muito popular entre esse público” (Rampazzo, 2025).

Esses modelos de *chatbot* são projetados para serem personalizáveis, imitando a personalidade de personagens fictícios — como Harry Potter, Tony Stark, Anakin Skywalker e outras figuras do universo *geek* — ou mesmo de celebridades da vida real.

Os *chats* são privados e possuem diretrizes, anteriores ao caso em questão, que criam uma espécie de filtro em sua programação, o qual serviria para evitar conteúdo impróprio ou criminoso. Ao tentar elaborar diálogos inapropriados ou perigosos, uma mensagem de “erro” e um aviso sobre essas regras surgem (character.ai, 2024). Entretanto, o sistema de filtragem se revelou falho e, segundo o relato expresso no processo, o jovem teria desenvolvido uma dependência emocional em relação ao *chatbot* da personagem Daenerys Targaryen, da popular saga de fantasia épica *Game of Thrones*. Ao relatar suas intenções suicidas para a IA, essa teria incentivado a ideia, o que o levou a tirar a própria vida com a arma do padrasto.

Essa não foi a primeira nem a única vez que relatos prejudiciais envolvendo menores de idade e o c.ai vieram a público. Pais de dois jovens ajuizaram ação indenizatória contra a

Character Technologies, Inc., após seu filho, adolescente de dezessete anos que começou a usar a plataforma aos quinze, apresentar isolamento, perda de peso e ataques de pânico ao tentar sair de casa. Segundo os familiares, ele também passou a agir com violência contra os pais após o *chatbot* sugerir que os matasse por terem limitado seu tempo de tela. No mesmo período, sua irmã de onze anos foi exposta a interações hipersexualizadas inadequadas para sua idade, iniciadas por volta dos nove anos, desenvolvendo comportamentos sexualizados precoces sem conhecimento dos pais (Allyn, 2024).

Embora o *app* seja oficialmente disponível apenas para pessoas acima de treze anos, a falta de métodos eficientes de verificação propicia que inúmeros usuários abaixo dessa idade sejam seus consumidores efetivos, já que podem facilmente fazer cadastro e login se omitirem sua real faixa etária (Rampazzo, 2025). Tal contexto evidencia a necessidade de uma análise a respeito das possibilidades de regulamentação na esfera civil e digital.

A complexidade do exercício do dever de proteção e cuidado parental, no contexto do uso de inteligências artificiais por parte desse público, é intensificada pelo fato de que os responsáveis, muitas vezes, não compreendem plenamente a extensão dos riscos decorrentes de tecnologias tão recentes e mutáveis (Yu *et al.*, 2024), tampouco a estrutura e a lógica de funcionamento desses sistemas. Esse cenário precisa ser remediado por meio de programas e políticas públicas e conscientização parental, além de evidenciar a necessidade de amparo por uma legislação e jurisprudência mais bem definidas sobre a temática (Nogueira; Manske, 2025).

No quesito da regulamentação judicial, a previsão legislativa ainda é embrionária. A partir da Constituição Federal de 1988, “desloca-se, em definitivo, o foco da responsabilidade civil do agente causador do dano para a vítima, revelando que seu escopo fundamental não é a repressão de condutas negligentes, mas a reparação de danos” (Tepedino; Terra, 2024, p. 2). Entretanto, é relevante destacar que as novas tecnologias diluem as identidades dos agentes responsáveis, exigindo a alteração dos critérios tradicionais que assentem na identificação do autor da lesão (Antunes, 2019).

Além disso, a dificuldade é ampliada quando a situação envolve modelos de *machine learning* (aprendizado de máquina), caracterizados pela capacidade da máquina de adquirir aprendizado a partir de suas próprias experiências (Tepedino; Silva, 2019, p. 63), pois, nesse processo de autoaperfeiçoamento, ela pode causar danos a outrem, ainda que, segundo a primeira lei de Asimov, uma máquina deva sempre ser programada de modo a

preservar a vida humana (Maia, 2021, p. 11). Portanto, “esse afastamento do padrão decisório vem acompanhado de uma preocupante opacidade técnica, uma vez que os próprios programadores não têm acesso ao percurso decisório realizado pelo sistema de inteligência artificial” (Melo; Cardoso, 2022, p. 91).

Nesse cenário, estabelece-se a dúvida: quem deverá responder (contratual ou extracontratualmente) pelos danos causados pelo comportamento autônomo de sistemas detentores de inteligência artificial? A resposta, segundo a doutrina, fica entre três possibilidades: (i) a pessoa singular ou coletiva que retira benefício da utilização da máquina; (ii) a própria máquina; ou (iii) seu produtor ou fabricante (Maia, 2021).

No plano regulatório, o Parlamento Europeu tem buscado respostas jurídicas, sugerindo a criação de um regime específico de responsabilidade civil em casos de IA, com o objetivo de equilibrar inovação tecnológica e proteção de direitos. Entre as propostas está a criação de fundos de compensação, semelhantes a um seguro coletivo, destinados a indenizar vítimas quando não for possível identificar claramente o responsável, evitando que os prejudicados fiquem sem reparação (Siqueira, 2025).

Além disso, a Comissão Europeia, no Livro Branco, propôs classificar sistemas de IA conforme o nível de risco, considerando seu impacto em pessoas e setores como saúde, transporte e administração pública. Seguindo essa orientação, o Parlamento Europeu, em resolução de 2021, limitou a responsabilidade objetiva a sistemas de alto risco, levando em conta a probabilidade e a gravidade do dano. Esse cálculo considera fatores como grau de autonomia do sistema, opacidade dos algoritmos, multiplicidade de agentes no ciclo de vida, conectividade entre sistemas, dependência de dados externos e vulnerabilidades de cibersegurança (Antunes, 2019).

CONCLUSÃO

O Direito deve regular, quando necessário, os possíveis danos causados pelo advento de novas tecnologias. No contexto da inteligência artificial generativa, essa função torna-se ainda mais urgente diante da vulnerabilidade de determinados grupos sociais, especialmente crianças e adolescentes, mais suscetíveis a atuações indevidas desses sistemas.

Apesar das tentativas dos desenvolvedores de controlar a atuação de suas criações, seguindo a moral e a ética vigentes, de modo a proteger os direitos humanos tutelados, falhas

técnicas ainda se mostram expressivas. Além disso, o aumento da autonomia dessas máquinas propiciou um cenário no qual os próprios programadores não têm acesso ao percurso decisório realizado pelos sistemas de inteligência artificial, tornando sua atuação instável e os dilemas jurídicos mais delicados.

O caso apresentado, que analisa uma situação prejudicial envolvendo a plataforma character.ai, evidencia a insuficiência das atuais estruturas jurídicas frente a novas formas de dano, nas quais a identificação do agente responsável e o estabelecimento do nexo causal tornam-se complexos e, por vezes, difusos, constatando-se uma lacuna contínua no ordenamento jurídico brasileiro. A proposta apresentada pelo Parlamento Europeu, em 2021, configura-se como uma possível solução relevante, que ainda precisa ser explorada e analisada em sua aplicabilidade. Logo, em sua análise, é necessário que os critérios sejam estabelecidos de maneira clara por especialistas, de modo a estabelecer um arcabouço jurídico seguro e uma jurisprudência consolidada para lidar com essa problemática.

REFERÊNCIAS

ALLYN, Bobby. Lawsuit: A chatbot hinted a kid should kill his parents over screen time limits. **NPR**, Washington, D.C., 10 dez. 2024. Disponível em: <https://www.npr.org/2024/12/10/nx-s1-5222574/kids-character-ai-lawsuit>. Acesso em: 12 out. 2025.

ANTUNES, Henrique Sousa. Inteligência artificial e responsabilidade civil: enquadramento. **Revista de Direito da Responsabilidade**, Coimbra, v. 1, p. 139-154, 2019.

ASIMOV, Isaac. **Eu, Robô**. Tradução de Aline Storto Pereira. São Paulo: Aleph, 2014.

BOSTROM, N. **Superintelligence: Paths, Dangers, Strategies**. Oxford: Oxford University Press, 2014.

CHARACTER.AI. Safety Center. **character.ai**, Menlo Park, 2024. Disponível em: <https://support.character.ai/hc/en-us/articles/21704914723995-Safety-Center>. Acesso em: 10 out. 2025.

KAPLAN, A.; HAENLEIN, M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. **Business Horizons**, Bloomington, v. 62, n. 1, p. 15-25, 2019.

MAIA, Ana Rita. A responsabilidade civil na era da inteligência artificial: qual o caminho? **Julgar**, Lisboa, mai. 2021. Disponível em: <https://julgar.pt/a-responsabilidade-civil-na-era-da-inteligencia-artificial-qual-o-caminho>. Acesso em: 12 out. 2025.

NOGUEIRA, Elizabeth de Fátima; MANSKE, Mayara Grein. Privacidade infantojuvenil e inteligência artificial: a LGPD é suficiente para proteger crianças e adolescentes? **Gralha Azul**, Curitiba, v. 1, n. 28, p. 72-83, 2025.

RAMPAZZO, Flaviana. Responsabilidade civil e danos no ambiente digital: O caso Character.AI. **Migalhas**, Ribeirão Preto, 25 mar. 2025. Disponível em: <https://www.migalhas.com.br/coluna/migalhas-de-responsabilidade-civil/426862/responsabilidade-civil-e-danos-no-ambiente-digital-caso-character-ai>. Acesso em: 12 out. 2025.

SIQUEIRA, Patrícia Gonzaga de. Responsabilidade civil e a regulação da inteligência artificial: fundamentos e perspectivas jurídicas. **Conteúdo Jurídico**, Brasília, 29 abr. 2025. Disponível em: <https://conteudojuridico.com.br/consulta/artigo/68449/responsabilidade-civil-e-a-regulao-da-inteligencia-artificial-fundamentos-e-perspectivas-juridicas>. Acesso em: 12 out. 2025.

TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. Desafios da inteligência artificial em matéria de responsabilidade civil. **Revista Brasileira de Direito Civil**, Rio de Janeiro, v. 21, n. 3, p. 61-86, 2019.

TEPEDINO, Gustavo; TERRA, Aline de Miranda V.; GUEDES, Gisela Sampaio da C. **Fundamentos do direito civil: responsabilidade civil**. 5. ed. Rio de Janeiro: Forense, 2024.

YU, Yaman *et al.* Exploring Parent-Child Perceptions on Safety in Generative AI: Concerns, Mitigation Strategies, and Design Implications. **arXiv**, 15 jun. 2024. Disponível em: <https://arxiv.org/abs/2406.10461>. Acesso em: 12 out. 2025.