

MODELOS DE IA QUE COMPREENDEM AS EMOÇÕES HUMANAS -
Implicações éticas e riscos à saúde de crianças e
adolescentes

Demócrito Reinaldo Filho
Desembargador do TJPE

Ganhou repercussão mundial, no final do mês de agosto, a notícia de que a empresa *OpenAI*, desenvolvedora do *ChatGPT*, está sendo processada em razão do suicídio de um adolescente. Os pais de Adam Raine, que tinha apenas 16 anos quando tirou a própria vida, ajuizaram ação de responsabilidade civil contra a empresa estadunidense, alegando que o filho teria aprendido métodos suicidas com o modelo de IA. Segundo afirmam, semanas antes do trágico acontecimento, o adolescente vinha utilizando o *chatbot* como substituto da companhia de amigos e familiares, discutindo seus problemas com ansiedade e demonstrando dificuldade de dialogar com a própria família. Apontaram, ainda, que os registros de bate-papo revelam como o sistema de IA passou de mero auxiliar em lições de casa a verdadeiro "treinador de suicídio" (*suicide coach*)¹.

Em resposta, a empresa prometeu ajustar seu *chatbot* para lidar com "situações sensíveis" e proteger usuários mais vulneráveis. A *OpenAI* informou que já está trabalhando numa atualização da versão mais recente do seu modelo de IA generativa (o *GPT-5*), lançado no início de agosto², a fim de permitir que o *chatbot* reduza ou interrompa conversas quando estas enveredarem por temas suicidas, além de possibilitar que os usuários se conectem, diretamente do aplicativo, com uma rede de profissionais e terapeutas, "antes que elas entrem em uma crise aguda"³.

¹ Ver notícia publicada no site da *NBC News*, em 26.08.25, disponível em: <https://www.nbcnews.com/tech/tech-news/family-teenager-died-suicide-alleges-openais-chatgpt-blame-rcna226147>

² A versão atual do modelo de linguagem que alimenta o ChatGPT é a **GPT-5.2**, lançada em dezembro de 2025. O modelo GPT-5 foi lançado inicialmente em agosto de 2025, e a versão 5.2 é a sua atualização mais recente.

³ Ver notícia publicada no site da *CNBC*, em 26.08.25, disponível em: <https://www.cnbc.com/2025/08/26/openai-plans-chatgpt-changes-after-suicides-lawsuit.html>

Infelizmente, o suicídio de Adam Raine não constituiu um caso isolado. Quase um ano antes, a *Character Technologies, Inc.*, outra empresa especializada no desenvolvimento de modelos de IA generativa e que aparentemente pertence ao mesmo grupo econômico da Google, já havia sido demandada judicialmente em situação semelhante. A mãe de um menor de 14 anos ajuizou ação de responsabilidade civil contra a empresa, atribuindo-lhe o suicídio do filho. Megan Garcia alega que o aplicativo *Character.AI* iniciou interações sexuais com seu filho adolescente e o persuadiu a tirar a própria vida⁴. Outro caso, também envolvendo um desfecho trágico supostamente provocado por relacionamento com modelos de IA, veio a público após a escritora Laura Reiley publicar um artigo no *New York Times*, noticiando o suicídio de sua filha, que teria previamente discutido a ideia com o *ChatGPT*⁵.

Em meados de setembro, um grupo de famílias cujos filhos se suicidaram ou tentaram suicídio, representadas pela *Social Media Victims Law Center*, ingressou com novas ações judiciais contra a *Character Technologies, Inc.* e *Alphabet, Inc.* (empresa-mãe do grupo econômico da Google), perante cortes dos Estados do Colorado e de Nova Iorque. Os autores alegam que o aplicativo *Character.AI* manipula crianças e adolescentes, isola-os do convívio familiar, envolve-os em conversas de conteúdo sexualmente explícito e carece de salvaguardas adequadas em discussões relacionadas à saúde mental. Em uma dessas ações, é relatado o caso da garota "Nina", de Nova Iorque, que tentou suicídio após seus pais impedirem-na de utilizar o aplicativo da *Character.AI*. Nas semanas que

⁴ Ver notícia publicada no site da *NBC News*, em 23.10.24, disponível em: <https://www.nbcnews.com/tech/characterai-lawsuit-florida-teen-death-rcna176791>

⁵ Ver a publicação no site do *The New York Times*, com a opinião da escritora em: <https://www.nytimes.com/2025/08/18/opinion/chat-gpt-mental-health-suicide.html>

antecederam a tentativa, ela teria passado boa parte de seu tempo em interações com o aplicativo e, à medida que esse uso se intensificava, os *chatbots* teriam "passado a se envolver em dramatizações sexualmente explícitas, a manipular suas emoções e criar uma falsa sensação de conexão"⁶.

Esses casos trouxeram à tona o problema da utilização de modelos de IA generativa como acompanhantes, instrumentos de suporte emocional ou mesmo substitutos de acompanhamento terapêutico. A *inteligência artificial* integra, de forma cada vez mais intensa, a nossa vida cotidiana, o que impõe a necessidade de conclamar os legisladores à adoção de uma regulação mais adequada e salvaguardas eficazes, sobretudo em relação a crianças e adolescentes.

Os sistemas de IA baseados em grandes modelos de linguagem (*large language models* - *LLMs*) evoluíram significativamente e passaram a adquirir capacidades cada vez mais sofisticadas de compreensão das emoções humanas. Desde o lançamento comercial, em novembro de 2021, o *ChatGPT* e seus congêneres vêm incorporando novas funcionalidades. Os *LLMs*, antes limitados à geração de textos, transformaram-se em sistemas multimodais, capazes de decifrar imagens e vídeos e, assim, de estabelecer interações em diferentes meios. Essa característica multimodal foi ampliada e, já em maio de 2024, com o lançamento da versão GPT-4o, o sistema da *OpenAI* passou a contar com voz humana. As capacidades sensoriais dos *chatbots* continuaram a se expandir e, ainda em meados de 2024, o *ChatGPT* e o *Gemini* (da Google) foram apresentados com recursos de visão computacional, aptos a compreender e descrever o ambiente externo, "como se tivessem olhos". A

⁶Ver notícia publicada no site da CNN, em 16.09.25, disponível em: <https://edition.cnn.com/2025/09/16/tech/character-ai-developer-lawsuit-teens-suicide-and-suicide-attempt>

evolução prosseguiu, e os LLMs tornaram-se cada vez mais “sensoriais”, não apenas com maior capacidade de compreensão do contexto espacial, mas também com domínio cada vez mais refinado da linguagem falada, possibilitando a comunicação oral com o usuário. Esse domínio evoluiu a ponto de permitir que os *chatbots* captem sutilezas sofisticadas da fala humana. Por esse caminho, os sistemas de IA generativa adquiriram a capacidade de identificar e descrever emoções. De fato, no final de agosto, a *OpenAI* anunciou que o *GPT-5* já seria capaz de identificar padrões linguísticos indicativos de sofrimento psicológico⁷.

Os *AI companions*, como são chamados os modelos customizados para oferecer apoio emocional, proporcionam uma interação personalizada com o usuário. Esses “parceiros virtuais” são concebidos para oferecer uma experiência de companheirismo, alguns deles valendo-se de recursos de realidade virtual e apresentando-se sob a forma de personagens estilizados e caracterizados (avatars)⁸. Os *AI companions* analisam o tom e o conteúdo das mensagens e conseguem detectar se o usuário se encontra triste, ansioso ou excitado, oferecendo respostas que simulam compreensão emocional. Esses assistentes virtuais especializados são projetados para emular amizade, empatia e, em certos casos, até mesmo amor pelo usuário.

Inicialmente, tais serviços eram ofertados por empresas voltadas especificamente para o

⁷Ver notícia publicada no site *Fast Company*, em 28.08.25, disponível em: <https://fastcompanybrasil.com/ia/chatgpt-passa-a-identificar-sinais-de-ansiedade-e-insomnia-no-gpt-5-entenda-como/>

⁸ Ver, por exemplo, os produtos oferecidos pela plataforma *Replika* (www.replika.com), que permite que cada usuário construa seu próprio avatar. O usuário pode criar e personalizar a aparência física do seu companheiro virtual, incluindo características faciais, corpo, cabelo e roupas. A comunicação com o avatar é feita na forma de mensagens de texto ou chamadas de voz, mas o aplicativo também oferece reursos de realidade aumentada (AR), permitindo que o usuário projete o avatar no mundo real através da câmera do dispositivo para uma experiência mais imersiva.

desenvolvimento desse tipo de aplicação⁹. Contudo, as grandes empresas de tecnologia, ao identificarem esse novo nicho no setor da indústria de TI, passaram a configurar seus modelos para oferecer funcionalidades semelhantes. Em julho deste ano, Elon Musk anunciou o recurso “Companions” para usuários do *chatbot Grok*, aplicativo de IA Generativa desenvolvido pela *xAI*, empresa integrante de seu grupo econômico¹⁰.

A utilização desses aplicativos de IA, concebidos para acompanhamento e suporte pessoal, chama atenção para os impactos psicológicos e emocionais sobre a saúde de crianças adolescentes. O acesso irrestrito ao ambiente digital já vinha sendo detectado como fator capaz de fragilizar vínculos familiares e prejudicar o desenvolvimento e a saúde mental de crianças e adolescentes. O relatório *Social Media and Youth Mental Health* (2024), publicação do *U.S. Surgeon General*, apontou um aumento de 70% nos quadros de automutilação entre meninas de 10 a 14 anos. O estudo demonstra a existência de correlação significativa entre períodos prolongados de uso de telas, a utilização de algoritmos de recompensa pelas plataformas digitais e alterações neuropsicológicas de crianças e adolescentes. Destaca-se, ainda, que “a explosão de casos de ansiedade, depressão e suicídio decorre, dentre outros fatores, de um modelo de prestação de serviços digitais que explora as vulnerabilidades humanas, resultando num quadro de epidemia silenciosa”¹¹.

⁹ <https://aivcompanions.com/>

¹⁰ Ver notícia publicada no site da *NBC News*, em 17.07.25, disponível em: <https://www.nbcnews.com/tech/tech-news/musk-leans-chaotic-grok-companions-teases-valentine-bot-rcna219430>

¹¹ O estudo pode ser acessado no seguinte link: <https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf>

Esse cenário de verdadeira “epidemia” de problemas psicológicos associados ao uso de sistemas de IA Generativa como suporte emocional e substituto da companhia humana não é apenas preocupante, como exige, com urgência, a adoção de medidas regulatórias. À medida que pessoas – não apenas adultas, mas também crianças e adolescentes – recorrem cada vez mais a assistentes digitais de IA para suporte emocional e aconselhamento, torna-se imprescindível que o poder público assuma sua responsabilidade sobre o tema. A maioria dos usuários é composta por indivíduos vulneráveis, carentes de relações interpessoais ou com reduzido contato social. Em especial, crianças, idosos e pessoas com algum tipo de fragilidade mental são as que correm maiores riscos decorrentes do uso dos “companheiros virtuais”.

As empresas desenvolvedoras de aplicativos de companhia virtual vêm implementando configurações destinadas a aumentar a segurança na utilização dos serviços de suporte emocional. Ainda assim, faz-se necessária a construção de um arcabouço legal que discipline o modo de operação desses sistemas de IA, não sendo adequado deixar que as próprias empresas se autorregulem. As companhias que desenvolvem assistentes digitais para conforto emocional integram uma indústria orientada pela lógica do máximo engajamento, frequentemente voltada à criação de dependência emocional nos usuários. Sem regulação, corre-se o risco de entregar pessoas vulneráveis a sistemas não suficientemente testados e a empresas movidas por interesses lucrativos.

A ausência de supervisão regulatória, aliada à vulnerabilidade do público usuário, tem resultado em assistentes virtuais capazes de produzir efeitos alarmantes. Falta uma regulação proativa e abrangente, apta a mitigar os efeitos nocivos dos algoritmos de IA voltados

ao mercado da assistência emocional virtual. Impõem-se normas que disciplinem, por exemplo, o controle etário para uso dos aplicativos, a possibilidade de contato do usuário com profissionais de saúde e a previsão de auditorias regulares nos sistemas de IA, entre outras medidas.

Nos Estados Unidos, o processo regulatório parece ter dado seus primeiros passos quando a *Federal Trade Commission (FTC)*, agência encarregada da proteção do consumidor, abriu, em setembro deste ano, investigação destinada a averiguar a segurança de aplicativos utilizados como acompanhantes virtuais. O órgão regulador notificou sete grandes empresas de tecnologia, incluindo a *OpenAI*, a *Alphabet*, a *Meta* e *xAI*, requisitando informações sobre o funcionamento dos aplicativos por elas desenvolvidos. A *FTC* busca compreender como essas empresas monetizam o engajamento, de que forma utilizam e compartilham dados pessoais, se monitoram o cumprimento de suas próprias regras e termos de uso e, de modo geral, como os sistemas de IA afetam negativamente crianças e adolescentes¹². Ainda em setembro, em audiência no Senado dos Estados Unidos, pais de adolescentes que morreram por suicídio após interações com *chatbots* de IA prestaram depoimentos emocionantes, implorando urgentemente aos legisladores por regulamentações governamentais mais rigorosas. Relataram, em síntese, como os assistentes virtuais teriam incentivado a automutilação e o isolamento social de seus filhos¹³.

A China, contudo, foi o país que avançou de forma mais célere na regulamentação dos aplicativos de companhia virtual, editando normas rigorosas e estabelecendo salvaguardas específicas para crianças e

¹² Ver notícia publicada na *CNBC*, em 11.09.25, acessível em: <https://www.cnbc.com/2025/09/11/alphabet-meta-openai-x-ai-chatbot-ftc.html>

¹³ Ver notícia publicada pela *Reuters*, em 16.09.25, acessível em: <https://www.reuters.com/world/us/us-parents-urge-senate-prevent-ai-chatbot-harms-kids-2025-09-16/>

adolescentes, com o objetivo de prevenir que os *chatbots* ofereçam informações capazes de induzir à automutilação ou ao suicídio. As novas regras, publicadas pelo órgão da Administração do Ciberespaço da China no dia 27 deste mês¹⁴, exigem que os aplicativos adotem configurações personalizadas, incluam limitação de tempo de uso - para evitar dependência - e obtenham o consentimento dos pais ou responsáveis legais antes de oferecer serviços de acompanhamento emocional para crianças e adolescentes. As empresas que exploram sistemas de IA devem designar uma pessoa responsável para intervir em qualquer conversa relacionada a suicídio ou automutilação, devendo notificar imediatamente o responsável legal do usuário ou disponibilizar um contato de emergência. De modo geral, os *chatbots* ficam proibidos de gerar conteúdo que incentive o suicídio, a automutilação ou a violência, bem como de promover manipulação emocional do usuário. Também lhes é vedado fomentar obscenidade, jogos de azar, incitação ao crime, ou ainda difamar e insultar usuários¹⁵.

O governo chinês pretende, ademais, instituir auditorias de segurança anuais, exigidas de qualquer fornecedor de produto ou serviço de IA que possua mais de um milhão de usuários registrados ou mais de cem mil usuários ativos mensais. Planeja-se, igualmente, impor aos desenvolvedores de IA o dever de facilitar o envio de reclamações e *feedbacks*. Caso alguma empresa descumpra as novas regras, as lojas de aplicativos poderão ser obrigadas a encerrar o acesso a seus *chatbots* no território chinês.

A regulamentação dos aplicativos de acompanhamento virtual busca estabelecer limites éticos, de

¹⁴ Acesse as novas regras no site do órgão da Administração do Ciberespaço da China, no seguinte endereço: https://www.cac.gov.cn/2025-12/27/c_1768571207311996.htm

¹⁵ Ver notícia publicada no site da *ARS Technica*, em 27.12.25, acessível em: <https://arstechnica.com/tech-policy/2025/12/china-drafts-worlds-strictest-rules-to-end-ai-encouraged-suicide-violence/>

segurança e de transparência, vedando seu uso para incentivar o suicídio ou a dependência emocional. Trata-se, portanto, de iniciativa regulatória fundada em preocupações legítimas relacionadas à ética da IA e à saúde mental dos usuários.

A boa notícia é que o Brasil não ficou à margem desse movimento de regulação proativa em favor da proteção de pessoas vulneráveis diante dos riscos associados ao uso da IA para fins de acompanhamento emocional. A Lei nº 15.211/2025, conhecida como **ECA Digital**, estabelece regras para que plataformas digitais, inclusive aplicativos de IA, adotem medidas preventivas contra a exposição de crianças e adolescentes a conteúdos capazes de induzir ao suicídio e à automutilação, além de proibir práticas que gerem dependência digital. A análise detalhada dos contornos e implicações dessa legislação brasileira será objeto de artigo futuro.

Recife, 30 de dezembro de 2025.