

New Lawsuits Targeting Personalized AI Chatbots Highlight Need for AI Quality Assurance and Safety Standards | ArentFox Schiff

Dan Jasnow

Character.AI is a chatbot, similar to those offered by other artificial intelligence (AI) developers. Where it differs, however, is that it allows customers to chat with a variety of pre-trained AI agents or “characters.” These characters can be representations of celebrities, fictional characters, or custom characters made by a customer. Despite the customer’s creation of specific characters, the customer retains very little, if any, control over a created character. The plaintiffs allege that Character.AI retains complete control over the large language model (LLM) itself, as well as the characters and how they operate.

Background

Interactions between Character.AI “characters” and two Texas minors ultimately led to this lawsuit. The first user, “J.F.,” is a 17-year-old with high-functioning autism, who began using the platform in April 2023, when he was 15. The complaint alleges that, due to his engagement with Character.AI, J.F. began isolating himself, losing weight, having panic attacks when he tried to leave his home, and became violent with his parents when they attempted to reduce his screen time. Included in the complaint is a screenshot of a conversation between J.F. and a Character.AI chatbot in which the bot encouraged J.F. to push back on a reduction in screen time and suggested that killing his parents may be a reasonable solution.

The second user, “B.R.,” is an 11-year-old girl. The complaint alleges that she downloaded Character.AI when she was 9 years old and that she was consistently exposed to hypersexualized interactions that were not age appropriate, causing her to develop sexualized behaviors prematurely and without her parents’ awareness.

The lawsuit follows shortly on the heels of another high-profile incident in which a Character.AI chatbot that infringed a well-known fictional character allegedly encouraged a 14-year-old boy to commit suicide.

Allegations

The crux of the plaintiffs’ allegations is that, through its design, Character.AI is “causing serious harms to thousands of kids, including suicide, self-mutilation, sexual solicitation, isolation, depression, anxiety, and harm towards others.” They further allege that, through deceptive and addictive designs, Character.AI isolates kids from their families and communities, undermines parental authority, and thwarts parents’ efforts to restrict kids’ online activity and keep them safe.

Much of the complaint is premised on allegations that the AI software suffers from design defects and that the defendants have failed to warn consumers of the alleged dangers, harms, or injuries that may exist when the product is used in a reasonably foreseeable manner, particularly by children. The plaintiffs are also seeking damages for intentional infliction of emotional distress.

Specifically, they assert that the defendants' failure to implement sufficient safety measures in the software before launching it into the market and targeting minors was intentional and reckless.

Two additional claims are directed against only Character.AI. The first is that by collecting and sharing personal information about children under the age of 13 without obtaining parental consent, Character.AI has violated the Children's Online Privacy Protection Act. The second alleges that Character.AI failed to meet its obligations to comply with applicable law governing harmful communication with minors and sexual solicitation of minors. Specifically, it is alleged that Character.AI failed to meet these obligations by knowingly designing Character.AI as a "sexualized product that would deceive minor customers and engage in explicit and abusive acts with them."

Google's Involvement

As a defendant, Google is a noteworthy addition to this case. The specific claims asserted against Google include strict product liability and negligence for defective design, strict liability and negligence for failure to warn, aiding and abetting, intentional infliction of emotional distress, unjust enrichment, and a violation of the Texas Deceptive Trade Practices Act.

Google's inclusion stems from the allegation that it incubated the technology behind Character.AI. Character.AI was founded by Noam Shazeer and Daniel De Freitas, former Google engineers. Both men left Google to launch Character.AI before being rehired in a reportedly \$2.7 billion deal intended to purchase shares of the startup and fund its continued operations. The complaint alleges that product development for Character.AI began when both men were still employed by Google, but that the founders faced significant internal roadblocks for failing to comply with Google's AI policies. Based on this past and continuing relationship, the plaintiffs allege that Character.AI was rushed to market with Google's knowledge, participation, and financial support.

Relief

The plaintiffs are requesting that Character.AI be taken offline and not returned until the defendants can establish that the public health and safety defects alleged have been cured. In addition to being taken offline, the plaintiffs are seeking various monetary damages, an order that would require Character.AI to warn parents and minors that the product is not suitable for minors, and requirements that the platform limit the collection and processing of minors' data.

Industry Impact and Future Considerations

This case highlights the importance of implementing robust safety measures in AI platforms, especially where they are easily accessible to and highly desirable by minors. Companies utilizing AI chatbots, non-playable characters, virtual assistants, or other similar products or services should carefully review their quality assurance programs, safety standards, data collection practices, and intellectual property policies to consider whether they have adequate safeguards in place to mitigate potential harm and ensure compliance with legal and regulatory obligations.

ArentFox Schiff will continue to monitor this case and is available to answer any questions you may have. For additional information, please contact the authors or the attorney who usually handles your matters.