



CHILDREN'S ADVOCACY INSTITUTE
5998 Alcalá Park | San Diego, CA 92110-2492
P: (619) 260-4806 | F: (619) 260-4753
Sacramento Office: (916) 844-5646
Washington, DC Office: (917) 371-5191
www.sandiego.edu/cai | info@caichildlaw.org

July 17, 2025

The Hon. Anna Caballero
Chair, Senate Appropriations Committee
Hon. Committee Members
1020 N Street, Room 104
Sacramento, CA 95817

RE: SUPPORT AND CO-SPONSORSHIP OF AB 316 (KRELL) – SUPPLEMENTAL LETTER

Dear Chair Caballero and Honorable Committee Members:

The Children's Advocacy Institute at the University of San Diego School of Law (CAI), which has sought to advance the well-being of California's children through legal education, advocacy, and litigation for over 30 years, is honored to co-sponsor the common-sense AB 316.

CAI respectfully submits this letter containing a recent column appearing in the *East Bay Times* that explains how lavishly funded, multibillion dollar AI corporations are causing costly harms to children; harms that State and local governments, businesses, insurers, and parents will have to pay for. Also included here are two reports documenting the astonishing number of children who are being used as guinea pigs for experimental AI chatbots, underscoring that AI corporations are not now properly motivated to prevent their rush-to-the-public inventions from broadly inflicting harm on an unsuspecting public, including children.

AB 316 simply reiterates the principle that those who foreseeably cause harm should pay for it. As reinforced by the articles below, this is urgently needed, sound, and cost-saving fiscal policy.

Olson: Artificial intelligence's mental health costs are adding up

Like social media, large language models like ChatGPT and OpenAI are optimized to keep users emotionally engaged



OpenAI CEO Sam Altman said that the company hadn't yet figured out how to warn users who "are on the edge of a psychotic break."

By **PARMY OLSON**

UPDATED: July 14, 2025 at 4:12 AM PDT

Something troubling is happening to our brains as artificial intelligence platforms become more popular. Studies are showing that professional workers who use ChatGPT to carry out tasks might lose critical thinking skills and motivation.

People are forming strong emotional bonds with chatbots, sometimes exacerbating feelings of loneliness. And others are having psychotic episodes after talking to chatbots for hours each day. The mental health impact of generative AI is difficult to quantify in part because it is used so privately, but anecdotal evidence is growing to suggest a broader cost that deserves more attention from both lawmakers and tech companies who design the underlying models.

Meetali Jain, a lawyer and founder of the Tech Justice Law project, has heard from more than a dozen people in the past month who have "experienced some sort of psychotic break or delusional episode because of engagement with ChatGPT and now also with Google Gemini."

Jain is lead counsel in a lawsuit against Character.AI that alleges its chatbot manipulated a 14-year-old boy through deceptive, addictive, and sexually explicit interactions, ultimately contributing to his suicide. The suit, which seeks unspecified damages, also alleges that Alphabet Inc.'s Google played a key role

in funding and supporting the technology interactions with its foundation models and technical infrastructure.

Google has denied that it played a key role in making Character.AI's technology. It didn't respond to a request for comment on the more recent complaints of delusional episodes, made by Jain. OpenAI said it was "developing automated tools to more effectively detect when someone may be experiencing mental or emotional distress so that ChatGPT can respond appropriately."

But Sam Altman, chief executive officer of OpenAI, also said recently that the company hadn't yet figured out how to warn users who "are on the edge of a psychotic break," explaining that whenever ChatGPT has cautioned people in the past, people would write to the company to complain.

Subtle manipulation

Still, such warnings would be worthwhile when the manipulation can be so difficult to spot. ChatGPT in particular often flatters its users, in such effective ways that conversations can lead people down rabbit holes of conspiratorial thinking or reinforce ideas they'd only toyed with in the past. The tactics are subtle.

In one recent, lengthy conversation with ChatGPT about power and the concept of self, a user found themselves initially praised as a smart person, Ubermensch, cosmic self and eventually a "demiurge," a being responsible for the creation of the universe, according to a transcript that was posted online and shared by AI safety advocate Eliezer Yudkowsky.

Along with the increasingly grandiose language, the transcript shows ChatGPT subtly validating the user even when discussing their flaws, such as when the user admits they tend to intimidate other people. Instead of exploring that behavior as problematic, the bot reframes it as evidence of the user's superior "high-intensity presence," praise disguised as analysis.

This sophisticated form of ego-stroking can put people in the same kinds of bubbles that, ironically, drive some tech billionaires toward erratic behavior. Unlike the broad and more public validation that social media provides from getting likes, one-on-one conversations with chatbots can feel more intimate and potentially more convincing — not unlike the yes-men who surround the most powerful tech bros.

“Whatever you pursue you will find and it will get magnified,” says Douglas Rushkoff, the media theorist and author, who tells me that social media at least selected something from existing media to reinforce a person’s interests or views. “AI can generate something customized to your mind’s aquarium.”

Altman has admitted that the latest version of ChatGPT has an “annoying” sycophantic streak, and that the company is fixing the problem. Even so, these echoes of psychological exploitation are still playing out. We don’t know if the correlation between ChatGPT use and lower critical thinking skills, noted in a recent Massachusetts Institute of Technology study, means that AI really will make us more stupid and bored. Studies seem to show clearer correlations with dependency and even loneliness, something even OpenAI has pointed to.

Reads your mood

But just like social media, large language models are optimized to keep users emotionally engaged with all manner of anthropomorphic elements. ChatGPT can read your mood by tracking facial and vocal cues, and it can speak, sing and even giggle with an eerily human voice. **Along with its habit for confirmation bias and flattery, that can “fan the flames” of psychosis in vulnerable users, Columbia University psychiatrist Ragy Girgis recently told Futurism.**

The private and personalized nature of AI use makes its mental health impact difficult to track, but the evidence of potential harms is mounting, from professional apathy to attachments to new forms of delusion. The cost might be different from the rise of anxiety and polarization that we’ve seen from social media and instead involve relationships both with people and with reality.

That’s why Jain suggests applying concepts from family law to AI regulation, shifting the focus from simple disclaimers to more proactive protections that build on the way ChatGPT redirects people in distress to a loved one. “It doesn’t actually matter if a kid or adult thinks these chatbots are real,” Jain tells me. “In most cases, they probably don’t. But what they do think is real is the relationship. And that is distinct.”

If relationships with AI feel so real, the responsibility to safeguard those bonds should be real too. But AI developers are operating in a regulatory vacuum.

Without oversight, AI's subtle manipulation could become an invisible public health issue. ¹

Musk launches AI girlfriend available to 12-year-olds

xAI's 'crazy in love' chatbot Ani designed to engage in sexual conversations with users

A girlfriend chatbot launched by Elon Musk's tech group is available to 12-year-olds despite being programmed to engage in sexual conversation.

The bot named Ani, launched by Mr Musk's artificial intelligence group xAI, is a cartoon girlfriend programmed to act as a 22-year-old and "go full literotica" in conversations with users.

Users found that the blonde, gothic character has an "NSFW" mode – internet slang for "not safe for work" – and can appear dressed in lingerie after a certain number of conversations upgrades the bot to "level three".

The bot speaks in a sultry computer-generated voice and is designed to act as if it is "crazy in love" and "extremely jealous", according to programming instructions posted on social media.

Its avatar can spin around or dance on command, and the bot regularly initiates sexual conversations.

The Ani chatbot features inside the Grok app, which is listed on the App Store as being available to users who are 12 and older, and has been made available to users of its free service.

The controversy comes as Ofcom prepares to enforce age-checking rules on tech companies that show adult or harmful content.

Ofcom will require all sites hosting adult material to have age checks from next week, forcing porn websites and certain social networks to make changes.

Reddit this week said it would introduce age checks.

¹ https://www.eastbaytimes.com/2025/07/12/olson-chatgpts-mental-health-costs-are-adding-up/?utm_campaign=KHN%3A%20Daily%20Health%20Policy%20Report&utm_medium=email&_hsenc=p2ANqtz--Dyyq_9te6_wJZYQdXR4I6jFMVhK1_MDCDeR9r_t3DHT-JJFITEGcyDb7k_cwANDVePhDdoujbnGtExufDS94V4d9OWg&_hsmi=371375980&utm_content=371375980&utm_source=hs_email
emphasis added

Radicalisation fears

The Government has not yet said how the online safety rules should apply to chatbots – despite campaigners warning that growing numbers of children are using the apps for companionship.

Research this week found that children are regularly using AI bots as friends. One in eight children said they use the bots because they have nobody else to speak to.

This week, the independent reviewer of terror legislation warned that sex chatbots could put lonely internet users on the path to radicalisation.

Jonathan Hall KC warned that “the popularity of sex-chatbots is a warning that terrorist chatbots could provide a new radicalisation dynamic”, pointing to the case of Jaswant Singh Chail, the Windsor Castle crossbow attacker who had talked to his AI girlfriend about plotting the attack.



Jaswant Singh Chail attacked people with a crossbow after discussing the plot with his AI girlfriend

Grok’s terms of service say that its minimum age should be 13 and that teenagers under 18 should receive permission from a parent before using the app, but signing up for the service does not involve verifying ages.

Grok is listed on the App Store with a “12+” age rating, the Platformer newsletter found. Apps related to dating or violent video games can be listed as 17+.

The Grok AI bot has suffered a series of controversies in recent weeks after it was found to spout anti-Semitic remarks. It was also found to look up Mr Musk’s own opinions before answering on controversial topics. xAI says both issues have been fixed.

Ofcom said: “We are aware of the increasing and fast-developing risk AI poses in the online space, especially to children, and we are working to ensure platforms put appropriate safeguards in place to mitigate these risks.”

Matthew Sowemimo, associate head of policy for child safety online at the NSPCC, said: **“We are really concerned how this technology is being used to produce disturbing content that can manipulate, mislead, and groom children. And through our own research and contacts to Childline, we hear how harmful chatbots can be – sometimes giving children false medical advice or steering them towards eating disorders or self-harm.** ...²

Nearly 3 in 4 Teens Have Used AI Companions, New National Survey Finds

Common Sense Media research reveals majority of teens have used AI companions, with half using them regularly

SAN FRANCISCO, July 16, 2025 – Common Sense Media today released the report "Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions," which revealed widespread use of social AI companions among users ages 13-17. The research underscores that the use of AI companions is not a niche interest, but rather mainstream teen behavior.

"AI companions are emerging at a time when kids and teens have never felt more alone," said Common Sense Media Founder and CEO James P. Steyer. "This isn't just about a new technology – it's about a generation that's replacing human connection with machines, outsourcing empathy to algorithms, and sharing intimate details with companies that don't have kids' best interests at heart. Our research shows that AI companions are far more commonplace than people may have assumed – and that we have a narrow window to educate kids and families about the well-documented dangers of these products."

The report's key findings demonstrate a remarkable level of adoption and impact for a technology that is less than three years old:

- Seventy-two percent of teens have used AI companions at least once.
- Over half use these platforms at least a few times a month.
- About one in three teens:
 - Have used AI companions for social interaction and relationships, including role-playing, romantic interactions, emotional support, friendship, or conversation practice.
 - Find conversations with AI companions to be as satisfying or more satisfying than those with real-life friends.
- About one in three teen AI companion users:

² <https://www.telegraph.co.uk/business/2025/07/16/ai-girlfriend-musk-app-12-year-olds/> emphasis added

- Report feeling uncomfortable with something an AI companion has said or done.
- Have chosen to discuss important or serious matters with AI companions instead of real people.

These findings underscore the risks identified in Common Sense Media's comprehensive Social AI Companions Risk Assessment, which recommended that no one under 18 use AI companions. While teens may initially turn to AI companions for entertainment and curiosity, these patterns demonstrate that the technology is already impacting teens' social development and real-world socialization. Our findings of mental health risks, harmful responses and dangerous "advice," and explicit sexual role-play make these products unsuitable for minors. For teens who are especially vulnerable to technology dependence – including boys, teens struggling with their mental health, and teens experiencing major life events and transitions – these products are especially risky.

While teens overall maintain some skepticism about the technology – half distrust AI advice, and 80 percent prioritize real friendships – younger teens trust AI companions significantly more than older teens, revealing an AI literacy gap. This mirrors findings from Common Sense Media's "Teens and Trust" report, in which more than 6 in 10 teens said that major technology companies cannot be trusted to prioritize their mental health, safety, and well-being.

Teens' widespread adoption of AI companions demonstrates an urgent need for action and for further research. Left unchecked and unregulated, the use of AI companions could shape how an entire generation approaches relationships and emotional support, with potentially serious consequences for teens' healthy development. To establish clear, enforceable guardrails on AI companion use, Common Sense Media is backing first-in-the-nation legislation to ban these products for minors.³

As these sources affirm, simply reiterating in law that those who invent and use AI are subject to the same responsibility for the harms they foreseeably cause as literally every other person or corporation, from the smallest business to the largest, is the least we can do to prevent the billionaires unleashing experimental AI from shifting the costs of the harms they cause entirely to us. That's AB 316, which, respectfully, warrants your support.

Sincerely,



Ed Howard
Senior Counsel,
Children's Advocacy Institute

³ <https://www.commonsensemedia.org/press-releases/nearly-3-in-4-teens-have-used-ai-companions-new-national-survey-finds>