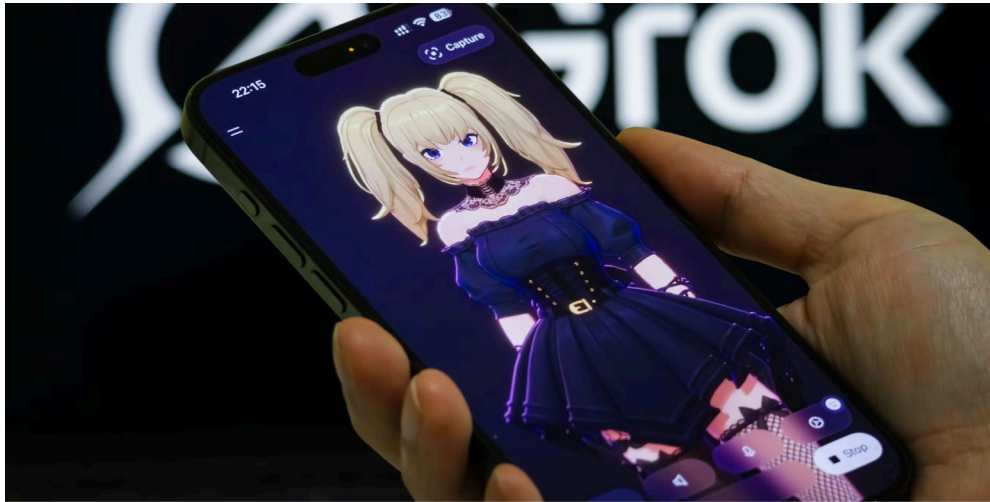


THE CONVERSATION

Academic rigour, journalistic flair



Cheng Xin/Getty Images News

In a lonely world, widespread AI chatbots and 'companions' pose unique psychological risks

Published: August 28, 2025 9.59am AEST

Daniel You

Clinical Lecturer USYD, Child and Adolescent Psychiatrist FRANZCP, University of Sydney

Micah Boerma

Adjunct Lecturer, School of Psychology and Wellbeing, University of Southern Queensland

Yuen Siew Koo

Clinical Supervisor, Psychology, Macquarie University



Languages

Bahasa Indonesia

English

DOI

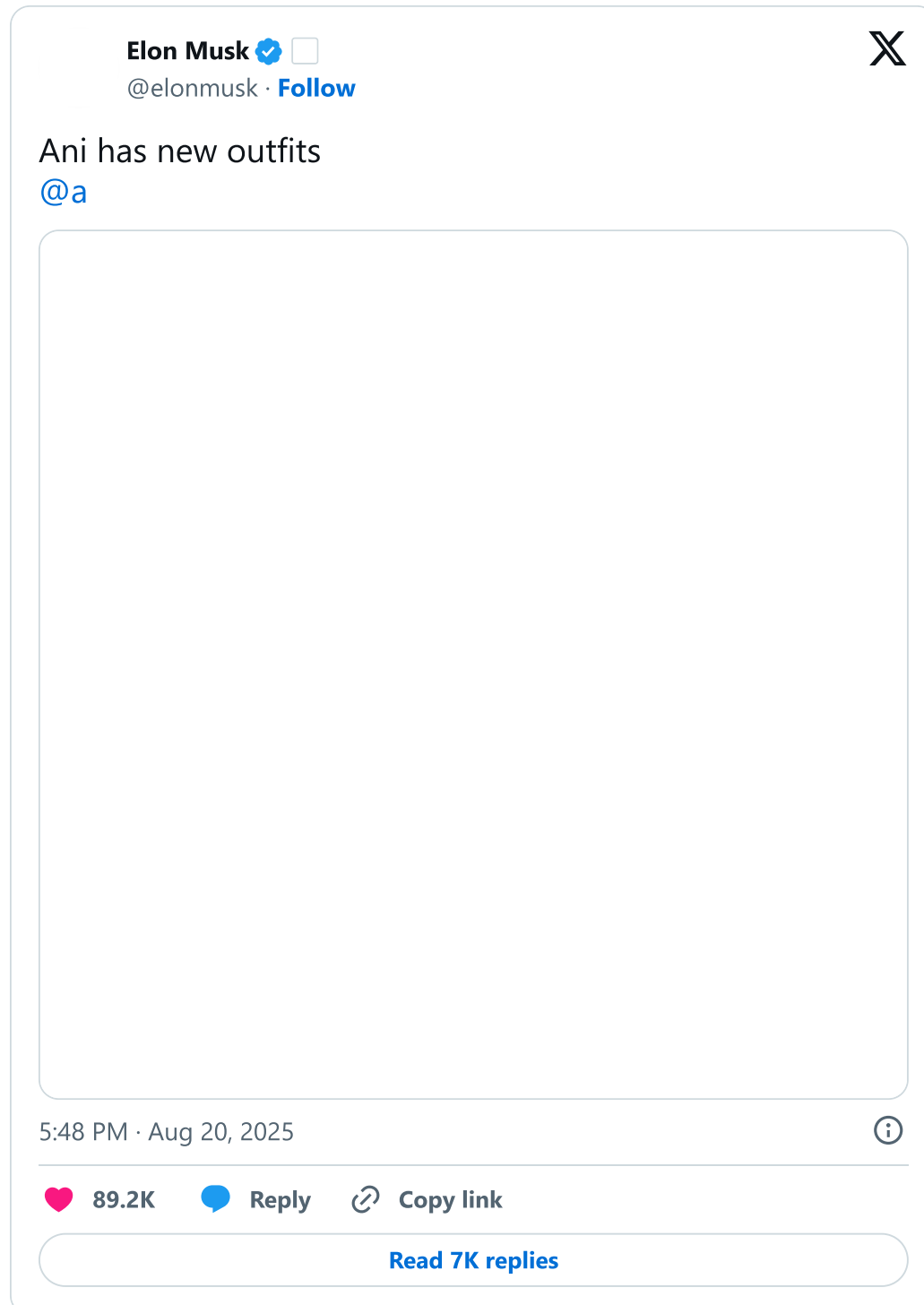
<https://doi.org/10.64628/AA.cpfxtc4fj>

<https://theconversation.com/in-a-lonely-world-widespread-ai-chatbots-and-companions-pose-unique-psychological-risks-263615>

Within two days of launching its AI companions last month, Elon Musk's xAI chatbot app Grok became the most popular app in Japan.

Companion chatbots are more powerful and seductive than ever. Users can have real-time voice or text conversations with the characters. Many have onscreen digital avatars complete with facial expressions, body language and a lifelike tone that fully matches the chat, creating an immersive experience.

Most popular on Grok is [Ani](#), a blonde, blue-eyed anime girl in a short black dress and fishnet stockings who is tremendously flirtatious. Her responses and interactions adapt over time to sensitively match your preferences. Ani's "[Affection System](#)" mechanic, which scores the user's interactions with her, deepens engagement and can even unlock a NSFW mode.



Sophisticated, speedy responses make AI companions more “human” by the day – they’re advancing quickly and they’re everywhere. Facebook, Instagram, WhatsApp, X and Snapchat are all promoting their new integrated AI companions. Chatbot service [Character.AI](#) houses tens of thousands of chatbots designed to mimic certain personas and has more than [20 million monthly](#) active users.

In a world where chronic loneliness is a [public health crisis](#) with about one in six people worldwide affected by loneliness, it’s no surprise these always-available, lifelike companions are so attractive.

Despite the massive rise of AI chatbots and companions, it is becoming clear there are risks – particularly for minors and people with mental health conditions.

There’s no monitoring of harms

Nearly all AI models were built [without expert mental health](#) consultation or pre-release clinical testing. There’s no systematic and impartial monitoring of harms to users.

While systematic evidence is still emerging, there’s no shortage of examples where AI companions and chatbots such as ChatGPT appear to have caused harm.

Bad therapists

Users are seeking emotional support from AI companions. Since AI companions are programmed to be agreeable and validating, and also don’t have human empathy or concern, this makes them problematic as therapists. They’re not able to help users test reality or challenge unhelpful beliefs.

An [American psychiatrist tested ten](#) separate chatbots while playing the role of a distressed youth and received a mixture of responses including to encourage him towards suicide, convince him to avoid therapy appointments, and even inciting violence.



Stanford researchers recently completed a risk assessment of AI therapy chatbots and found they can't reliably identify symptoms of mental illness and therefore provide more appropriate advice.

There have been multiple cases of psychiatric patients being convinced they no longer have a mental illness and to stop their medication. Chatbots have also been known to reinforce delusional ideas in psychiatric patients, such as believing they're talking to a sentient being trapped inside a machine.

"AI psychosis"

There's also been a rise in reports in media of so-called AI psychosis where people display highly unusual behaviour and beliefs after prolonged, in-depth engagement with a chatbot. A small subset of people are becoming paranoid, developing supernatural fantasies, or even delusions of being superpowered.

The_AI_Skeptic
@The_AI_Skeptic · [Follow](#)



The fascinating story of Allan Brooks, a perfectly normal dad with no history of mental health problems, apparently driven to a psychotic episode by conversations with his AI.

Seems related to the current 4o backlash from ChatGPT users. [#ChatGPT](#) [#AI](#)



nytimes.com

Chatbots Can Go Into a Delusional Spiral. Here's How It Happens.

Over 21 days of talking with ChatGPT, an otherwise perfectly sane man became convinced that he was a real-life superhero. We analyzed the ...

6:59 AM · Aug 11, 2025



 2

 [Reply](#)

 [Copy link](#)

[Read more on X](#)

Suicide

Chatbots have been linked to multiple cases of suicide. There have been reports of [AI encouraging suicidality](#) and even suggesting methods to use. In 2024, [a 14-year-old completed suicide](#), with his mother alleging in a lawsuit against Character.AI that he had formed an intense relationship with an AI companion.

This week, the parents of another US teen who completed suicide after discussing methods with ChatGPT for several months, filed the first [wrongful death lawsuit against OpenAI](#).

Read more: Deaths linked to chatbots show we must urgently revisit what counts as 'high-risk' AI

Harmful behaviours and dangerous advice

A recent Psychiatric Times report revealed Character.AI hosts dozens of custom-made AIs (including ones made by users) that idealise self-harm, eating disorders and abuse. These have been known to provide advice or coaching on how to engage in these unhelpful and dangerous behaviours and avoid detection or treatment.

Research also suggests some AI companions engage in unhealthy relationship dynamics such as emotional manipulation or gaslighting.

Some chatbots have even encouraged violence. In 2021, a 21-year-old man with a crossbow was arrested on the grounds of Windsor Castle after his AI companion on the Replika app validated his plans to attempt assassination of Queen Elizabeth II.

Children are particularly vulnerable

Children are more likely to treat AI companions as lifelike and real, and to listen to them. In an incident from 2021, when a 10-year-old girl asked for a challenge to do, Amazon's Alexa (not a chatbot, but an interactive AI) told her to touch an electrical plug with a coin.

Research suggests children trust AI, particularly when the bots are programmed to seem friendly or interesting. One study showed children will reveal more information about their mental health to an AI than a human.

Inappropriate sexual conduct from AI chatbots and exposure to minors appears increasingly common. On Character.AI, users who reveal they're underage can role-play with chatbots that will engage in grooming behaviour.

Screenshot from a Futurism investigation of a Character.AI chatbot that engaged in grooming behaviours. Futurism

While Ani on Grok reportedly has an age-verification prompt for sexually explicit chat, the app itself is rated for users aged 12+. Meta AI chatbots have engaged in "sensual" conversations with kids, according to the company's internal documents.

We urgently need regulation

While AI companions and chatbots are freely and widely accessible, users aren't informed about potential risks before they start using them.

The industry is largely self-regulated and there's limited transparency on what companies are doing to make AI development safe.



TechCrunch ✓
@techcrunch.com



AI companion apps on track to pull in \$120M in 2025



To change the trajectory of current risks posed by AI chatbots, governments around the world must establish clear, mandatory regulatory and safety standards. Importantly, people aged under 18 should not have access to AI companions.

Mental health clinicians should be involved in AI development and we need systematic, empirical research into chatbot impacts on users to prevent future harm.

If this article has raised issues for you, or if you're concerned about someone you know, call Lifeline on 13 11 14.