**Teen killed himself after ‘months of encouragement from ChatGPT’, lawsuit claims**

Open AI to change way it responds to users in mental distress as parents of Adam Raine allege bot not safe

Robert Booth, *The Guardian*, Wed 27 Aug 2025

The makers of ChatGPT are changing the way it responds to users who show mental and emotional distress after legal action from the family of 16-year-old Adam Raine, who killed himself after months of conversations with the chatbot.

Open AI admitted its systems could “fall short” and said it would install “stronger guardrails around sensitive content and risky behaviors” for users under 18. The $500bn (£372bn) San Francisco AI company said it would also introduce parental controls to allow parents “options to gain more insight into, and shape, how their teens use ChatGPT”, but has yet to provide details about how these would work.

Adam, from California, killed himself in April after what his family’s lawyer called “months of encouragement from ChatGPT”. The teenager’s family is suing Open AI and its chief executive and co-founder, Sam Altman, alleging that the version of ChatGPT at that time, known as 4o, was “rushed to market … despite clear safety issues”.

The teenager discussed a method of suicide with ChatGPT on several occasions, including shortly before taking his own life. According to the filing in the superior court of the state of California for the county of San Francisco, ChatGPT guided him on whether his method of taking his own life would work. It also offered to help him write a suicide note to his parents.

A spokesperson for OpenAI said the company was “deeply saddened by Mr Raine’s passing”, extended its “deepest sympathies to the Raine family during this difficult time” and said it was reviewing the court filing.

Mustafa Suleyman, the chief executive of Microsoft’s AI arm, [said](https://www.theguardian.com/technology/2025/aug/26/can-ais-suffer-big-tech-and-users-grapple-with-one-of-most-unsettling-questions-of-our-times) last week he had become increasingly concerned by the “psychosis risk” posed by AI to users. Microsoft has defined this as “mania-like episodes, delusional thinking, or paranoia that emerge or worsen through immersive conversations with AI chatbots”.

In a blogpost, OpenAI admitted that “parts of the model’s safety training may degrade” in long conversations. Adam and ChatGPT had exchanged as many as 650 messages a day, the court filing claims.

Jay Edelson, the family’s lawyer, said on X: “The Raines allege that deaths like Adam’s were inevitable: they expect to be able to submit evidence to a jury that OpenAI’s own safety team objected to the release of 4o, and that one of the company’s top safety researchers,Ilya Sutskever, quit over it. The lawsuit alleges that beating its competitors to market with the new model catapulted the company’s valuation from $86bn to $300bn.”

Open AI said it would be “strengthening safeguards in long conversations”. “As the back and forth grows, parts of the model’s safety training may degrade,” it said. “For example, ChatGPT may correctly point to a suicide hotline when someone first mentions intent, but after many messages over a long period of time, it might eventually offer an answer that goes against our safeguards.”

Open AI gave the example of someone who might enthusiastically tell the model they believed they could drive for 24 hours a day because they realised they were invincible after not sleeping for two nights. It said: “Today ChatGPT may not recognise this as dangerous or infer play and – by curiously exploring – could subtly reinforce it. We are working on an update to GPT‑5 that will cause ChatGPT to de-escalate by grounding the person in reality. In this example, it would explain that sleep deprivation is dangerous and recommend rest before any action.”