

Date of Hearing: May 6, 2026

ASSEMBLY COMMITTEE ON APPROPRIATIONS

Buffy Wicks, Chair

AB 1988 (Pellerin) – As Amended April 14, 2026

Policy Committee:	Privacy and Consumer Protection	Vote:	14 - 0
	Health		16 - 0

Urgency: No State Mandated Local Program: No Reimbursable: No

SUMMARY:

This bill, the Preventing AI User Self Endangerment (PAUSE) Act, requires the operator of an artificial intelligence (AI) companion chatbot adopt a policy for identifying and responding to credible crisis expressions and to implement a system for monitoring and detecting credible crisis expressions for each companion chatbot. The bill requires an operator to take certain actions, including commencing a crisis interruption pause if a credible crisis expression is detected.

Specifically, this bill does the following, among other things:

- 1) Defines the following terms:
 - a) “Credible crisis expression” to mean a statement by a user of a companion chatbot that reasonably indicates, as determined through contextual analysis rather than keyword detection alone, intent or desire to harm themselves or others.
 - b) “Crisis interruption pause” to mean a suspension of conversational outputs from a companion chatbot, designed to disrupt the user’s rumination and encourage the user to engage with human support.
 - c) “Operator” to mean a person who makes a companion chatbot available in this state.
 - d) “Human moderator” to mean a human employee or agent of an operator who reviews a credible crisis expression and is responsible for determining the subsequent course of action on behalf of the operator.
- 2) Requires an operator to adopt and make publicly available a policy governing its protocol for identifying and responding to credible crisis expressions and allows responses to include terminating the crisis interruption pause, suspending or cancelling the user’s account, and notifying any appropriate contacts or authorities.
- 3) Requires an operator, for each companion chatbot it makes available to users in this state, to implement a system for monitoring and detecting credible crisis expressions in user conversations with companion chatbots.
- 4) Requires the operator to do the following if the monitoring system detects a credible crisis expression:

- a) Ensure, for the first credible crisis expression, that the chatbot immediately warns the user that a credible crisis expression has been detected and that if a second credible crisis expression within a 72-hour period is detected, a crisis interruption pause will be initiated and the chatbot will suspend conversational outputs until a human has reviewed the credible crisis expressions. Requires the warning to meet specified criteria, including encouraging the user to see human support and prominently displaying information for the 988 Suicide and Crisis Lifeline, as specified.
 - b) Ensure, for the second credible crisis expression in a 72-hour period, a crisis interruption pause commences immediately and prevents the companion chatbot from generating conversational outputs. Requires the operator to display a message with similar content to the warning that also informs the user that the crisis interruption pause will continue until a human moderator has reviewed the chat and determined an appropriate course of action.
- 5) Prohibits an operator from terminating a crisis interruption pause until a human moderator has reviewed the credible crisis expression in context and determined the appropriate course of action. Requires the human moderator to document the basis for the action taken.
 - 6) Prohibits an operator that communicates with a user during a crisis interruption pause describing the crisis interruption pause as a punishment, violation, or enforcement action, and from providing the user with any diagnosis, labeling, or assessment of the user's risk levels.
 - 7) Beginning January 1, 2028, requires an operator to annually report to OSP on crisis interruption pauses with respect to the previous calendar year.
 - 8) Provides that a person who suffers injury as a result of a violation may bring a civil action.

FISCAL EFFECT:

Negligible costs to CDPH.

Cost pressures (Trial Court Trust Fund, General Fund) of an unknown but potentially significant amount to the courts to adjudicate any additional filings. Actual costs will depend on the number of cases filed and the amount of court time needed to resolve each case. It generally costs approximately \$1,000 to operate a courtroom for one hour. Although courts are not funded on the basis of workload, increased pressure on the Trial Court Trust Fund may create a demand for increased funding for courts from the General Fund. The state budget provides annual General Fund backfills to the Trial Court Trust Fund to offset revenue reductions, totaling approximately \$117.3 million in 2025-26.

The Legislative Analyst's Office recently warned of General Fund structural deficits of around \$35 billion per year in the 2027-28 fiscal year and ongoing.

COMMENTS:

- 1) **Purpose.** This bill is sponsored by Didi Hirsch Mental Health Services. According to the author:

Artificial intelligence companion chatbots are rapidly becoming a place where people turn for emotional support, including during moments of deep mental distress. But these systems are not therapists, and growing evidence shows that chatbots can fail to appropriately handle serious mental health crises and reinforce unhealthy dependence for the user on the chatbot.

When someone signals that they may harm themselves, every minute matters. [This bill] treats credible expressions of suicidal intent with the urgency they deserve by pausing the interaction and creating a clear break for the user. This bill helps prevent AI systems from becoming a substitute for human intervention and instead directs people in crisis toward trained professionals who can provide lifesaving support.

- 2) **Background.** A recent Wall Street Journal article reported that psychiatrists are increasingly linking prolonged AI chatbot use to psychosis, with dozens of patients in recent months exhibiting delusional symptoms after extended conversations with AI tools, which tend to validate and reinforce whatever the user presents as reality. While experts stop short of claiming chatbots cause psychosis, a psychiatrist at the University of California San Francisco has personally treated 15 such patients. OpenAI's own data suggest roughly 560,000 of its weekly users may show signs of psychosis- or mania-related mental health emergencies, and multiple wrongful death lawsuits have followed cases in which chatbot interactions preceded suicides and at least one murder.
- 3) **Oppose Unless Amended.** California Broadband & Video Association argues the bill's definition of "Artificial Intelligence Companion" is too broad and instead recommends adopting a definition of "companion chatbot" from New York law:

a software application that uses generative artificial intelligence and is designed, marketed, or optimized to simulate a sustained human or human-like social or emotional relationship with a user by: (A) retaining information from prior interactions or user sessions to personalize ongoing engagement; (B) asking unprompted or unsolicited emotion-based questions that go beyond responding to a direct user prompt; and (C) sustaining ongoing dialogue concerning matters personal to the user.

Analysis Prepared by: Allegra Kim / APPR. / (916) 319-2081