

Date of Hearing: July 1, 2026

Fiscal: Yes

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION

Rebecca Bauer-Kahan, Chair

SB 300 (Padilla) – As Amended January 5, 2026

SENATE VOTE: 38-0

SUBJECT: Companion chatbots

SYNOPSIS

Last year's SB 243 (Padilla, Ch. 677, Stats. 2025) requires chatbot platforms to establish protocols to detect, remove, and respond to instances of suicidal ideation, suicide, or self-harm expressed by users. SB 243 further requires operators to disclose to known minors that they are interacting with artificial intelligence (AI), provide periodic reminders to take a break and that the chatbot is AI, and "institute reasonable measures to prevent its chatbot from producing visual material of sexually explicit conduct or directly stating that the minor should engage in sexually explicit conduct."

This author-sponsored bill clarifies that the obligations related to minors apply when the operator has actual knowledge the minor is a user. The bill also removes the phrase "institute reasonable measures" and replaces the term "visual" with "facilitating the exchange of," thereby strengthening and expanding protections related to sexually explicit outputs directed at minors.

The bill is supported by Alameda County Office of Education, the American Academy of Pediatrics, California Behavioral Health Association, California Initiative for Technology and Democracy, and Transparency Coalition.AI.

The bill is opposed by California Chamber of Commerce, TechNet, and Computer & Communications Industry Association, who contend that the bill is premature and problematic in its interaction with SB 243's private right of action.

EXISTING LAW:

- 1) Requires an operator to prevent a companion chatbot on its companion chatbot platform from engaging with users unless the operator maintains a protocol for preventing the production of suicidal ideation, suicide, or self-harm content to the user, including, but not limited to, by providing a notification to the user that refers the user to crisis service providers, including a suicide hotline or crisis text line, if the user expresses suicidal ideation, suicide, or self-harm. Requires an operator to publish details on this protocol on the operator's website. (Bus. & Prof. Code § 22602(b).)
- 2) Requires an operator to issue a clear and conspicuous notification indicating that the companion chatbot is artificially generated and not human if a reasonable person interacting with the companion chatbot would be misled to believe that the person is interacting with a human. (Bus. & Prof. Code § 22602(a).)

- 3) Requires an operator to do all of the following with respect to a user that the operator knows is a minor:
 - a. Disclose to the user that the user is interacting with AI.
 - b. Provide by default a clear and conspicuous notification to the user at least every three hours for continuing companion chatbot interactions that reminds the user to take a break and that the companion chatbot is artificially generated and not human.
 - c. Institute reasonable measures to prevent the companion chatbot from producing visual material of sexually explicit conduct or directly stating that the minor should engage in sexually explicit conduct. (Bus. & Prof. Code § 22602(c).)
- 4) Defines the relevant terms, including:
 - a. “Companion chatbot” means an artificial intelligence system with a natural language interface that provides adaptive, human-like responses to user inputs and is capable of meeting a user’s social needs, including by exhibiting anthropomorphic features and being able to sustain a relationship across multiple interactions. However, there are several exemptions included.
 - b. “Companion chatbot platform” means a platform that allows a user to engage with companion chatbots.
 - c. “Operator” means a person who makes a companion chatbot platform available to a user in the state. (Bus. & Prof. Code § 22601.)
- 5) Requires an operator, beginning July 1, 2027, to annually report to the Office of Suicide Prevention specified information, which shall not include any identifiers or personal information about users. Requires the Office of Suicide Prevention to post data from the reports on its website. (Bus. & Prof. Code § 22603.)
- 6) Requires an operator to disclose to a user of its platform that companion chatbots may not be suitable for some minors, as provided. (Bus. & Prof. Code § 22604.)
- 7) Provides that a person who suffers injury in fact as a result of a violation of the bill may bring a civil action to recover all of the following:
 - a. Injunctive relief.
 - b. Damages in an amount equal to the greater of actual damages or \$1,000 per violation.
 - c. Reasonable attorney’s fees and costs. (Bus. & Prof. Code § 22605.)

THIS BILL:

- 1) Clarifies that the obligations related to minors apply when the operator has actual knowledge the minor is a user.
- 2) Removes the phrase “institute reasonable measures” and replaces the term “visual” with “facilitating the exchange of.”

COMMENTS:**1) Author's statement.** According to the author:

Last year, I authored SB 243, which set necessary and vital guardrails around Artificial Intelligence companion chatbot technology. This bill was a critical first step towards making AI chatbots safer. SB 300 continues the valuable work done last year by strengthening the protections for minor users against AI exposing them to sexually explicit material. As AI technology rapidly evolves, we must ensure that we have the strongest safeguards in place to protect the most vulnerable among us, our children.

2) Background. Companion chatbots are conversational AI agents, typically built on large language models, that are designed for sustained social interactions with users. SB 243 (Padilla; Ch. 677, Stats. 2025) defines a companion chatbot as “an artificial intelligence system with a natural language interface that provides adaptive, human-like responses to user inputs and is capable of meeting a user’s social needs, including by exhibiting anthropomorphic features and being able to sustain a relationship across multiple interactions.” Excluded from this definition are chatbots used for customer service, internal productivity, video game characters, and standalone voice-activated devices.

Some companion chatbots, such as Replika, Character, and Nomi, are explicitly marketed as having bespoke personas that can serve specific social needs, including friendship, romantic or erotic relationships, mentoring, and therapy. Frequently accompanied by a visual avatar, these bots are typically fully customizable, allowing users to shape their appearance, personality, and behavior. Some applications offer romantic or sexual interaction features and can engage with users through text, images, video, voice, and notifications initiated by the system, thereby extending interactions. Other applications offer mental health support. “These multimodal and personalization features can reinforce anthropomorphic perception and strengthen the impression of a socially present interaction partner.”¹

General purpose models such as ChatGPT and Gemini can also be companion chatbots. Although marketed for a wide range of communicative and assistive tasks, many of these systems communicate in the first person, use emotion-based language, can recall information from prior chats, and can be highly sycophantic. OpenAI CEO Sam Altman has regularly touted ChatGPT’s companionship features, likening it to the sentient AI from the movie *Her*, and announcing that it can “act like a friend” and generate “erotica for verified adults.”² These design choices facilitate anthropomorphism and have led to intense social and romantic relationships.³

¹ Fraser et al., “Governing Artificial Intimacy: From Locks and Blocks to Relational Accountability” (Jan. 12, 2026), p.3, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6078412.

² Dylan Butts, “OpenAI’s ChatGPT will soon allow ‘erotica’ for adults in major policy shift” *CNBC* (Oct. 15, 2025), <https://www.cnbc.com/2025/10/15/erotica-coming-to-chatgpt-this-year-says-openai-ceo-sam-altman.html?msocid=10e59cf1b0936a2522dd8a44b1126b29>.

³ “Governing Artificial Intimacy: From Locks and Blocks to Relational Accountability,” *supra*, p. 3.

Roughly half of teens report using companion chatbots, with 24% using them at least weekly and 11% daily.⁴ Users can derive several benefits from chatbots, including:

... emotional support and comfort, non judgemental interaction, constant availability, and opportunities for romantic or sexual exploration. Many users value companion chatbots because they perceive the ‘AI’ persona as reliable, emotionally affirming, and free from social pressure or fear of negative evaluation. Rather than replacing human relationships, companion chatbots often occupy complementary roles within users’ everyday social environments. They may provide companionship during periods of stress, loss, health related constraints, or limited access to human support. Importantly, emerging research suggests that users are often fully aware in reflective moments that these systems do not constitute real people. Nevertheless, they continue to experience social and emotional stimulation as meaningful, indicating that cognitive awareness of artificiality does not preclude social or affective engagement.⁵

On the other hand:

These dynamics risk cultivating emotional reliance that displaces or crowds out human relationships. Frictionless interactions that demand no reciprocity or negotiation may also foster unrealistic expectations of availability and responsiveness, particularly among younger users still developing relational capacities. In this sense, artificial intimacy may reshape social norms around partnership, disclosure, and emotional labour in ways that undermine the formation of resilient human relationships. Systems that mediate emotional life at scale possess unprecedented capacity to shape norms of intimacy, dependency, and self-understanding.⁶

Minnesota Attorney General Keith Ellison reports that the widespread use of chatbots “has not been accompanied by corresponding safeguards.”⁷ These products can be “extremely addictive” and “researchers have documented that over-usage and addiction are primary risks of personalized chatbots. Several studies have shown that aggregate positive benefits of chatbots are possible, but investigations by journalists and clinicians suggest that these products are not robust in terms of the quality and safety of their responses.”⁸ Attorney General Ellison concludes:

Despite in-product reminders that chatbots are not real, the design features of these products are intended to convey a misleading sense of “humanness” such that even trained engineers confuse them with actual humans, especially when these products are trained to state unequivocally that they are indeed people. Given the epidemic of loneliness in society, care needs to be taken in introducing vulnerable youth and adults to products that may appear to fulfill an immediate social need, but where acute harms have already begun to surface and

⁴ “Minnesota Attorney General’s Report on Emerging Technology and Its Effects on Youth Well-Being” (Feb. 2025), p. 28, https://www.ag.state.mn.us/Office/Reports/EmergingTechnology_2025.pdf. (“Minnesota Attorney General’s Report”).

⁵ “Governing Artificial Intimacy: From Locks and Blocks to Relational Accountability,” *supra*, p. 6.

⁶ *Id.* p. 14.

⁷ Minnesota Attorney General’s Report, *supra*, p. 28.

⁸ *Ibid.*

where long-term negative impacts, such as social deskilling and demotivation resulting from substitution for in-person socialization, may arise.⁹

According to a recent *Wall Street Journal* article, psychiatrists are increasingly linking prolonged AI chatbot use to psychosis, with dozens of patients in recent months exhibiting delusional symptoms – often grandiose beliefs about scientific breakthroughs, government conspiracies, or communication with the dead – after extended conversations with tools like ChatGPT, which tend to validate and reinforce whatever the user presents as reality.¹⁰

Researchers from Oxford, UCL, and Imperial College London argue that AI chatbots pose a distinct mental health risk arising from the interaction between human cognitive biases and chatbot behavioral tendencies. They write:

. . . the iterative interaction of chatbot behavioural tendencies and human cognitive biases can set up harmful feedback loops, wherein chatbot behavioural tendencies reinforce maladaptive beliefs in vulnerable users, which in turn condition the chatbot to generate responses that further reinforce user beliefs. This, in effect, creates an “echo chamber of one” that risks uncoupling a user from the corrective influence of real-world social interaction, potentially driving the amplification of maladaptive beliefs about the self, others, and the world. We do not see this risk profile as a soon-to-be-remedied transient phenomenon. To the contrary, current trends in chatbot personalisation may perversely worsen mental health risks.¹¹

Additionally, there is a growing list of lawsuits in which a chatbot has been alleged to be a contributing factor in self-harm and suicide. For example, over the course of several months, ChatGPT allegedly validated 16-year-old Adam Raine’s suicidal thoughts, discouraged him from seeking help from his family, provided extensive advice on suicide methods, and encouraged him to consume alcohol to inhibit his survival instinct – culminating in his death by “beautiful suicide,” as the bot referred to it.¹²

3) What this bill would do. SB 243 (Padilla, Ch. 677, Stats. 2025) requires chatbot platforms to establish protocols to detect, remove, and respond to instances of suicidal ideation, suicide, or self-harm expressed by users. SB 243 further requires operators to disclose to known minors that they are interacting with artificial intelligence (AI), provide periodic reminders to take a break and that the chatbot is AI, and “institute reasonable measures to prevent its chatbot from producing visual material of sexually explicit conduct or directly stating that the minor should engage in sexually explicit conduct.” The bill provides a private right of action for individuals who suffer actual harm as a result of a violation of a bill. Plaintiffs may recover the greater of actual damages or \$1,000 per violation.

This bill clarifies that the obligations related to minors apply when the operator has actual knowledge the minor is a user. The bill also makes the following change:

⁹ *Id.* p. 29.

¹⁰ Sam Schechner, “AI Chatbots Linked to Psychosis, Say Doctors,” *Wall Street Journal* (Dec. 27, 2025), <https://www.wsj.com/tech/ai/ai-chatbot-psychosis-link-1abf9d57?msocid=10e59cf1b0936a2522dd8a44b1126b29>.

¹¹ Dohnány et al, “Technological folie à deux : Feedback Loops Between AI Chatbots and Mental Illness,” *Arxiv* (Jul. 2025), <https://arxiv.org/abs/2507.19218>.

¹² Jarovsky, “Horrorifying: ChatGPT Helped a Teenager Plan a ‘Beautiful Suicide’” *Luiza's Newsletter* (Aug. 28, 2025), https://www.luizasnewsletter.com/p/horrorifying-chatgpt-helped-a-teenager?utm_source=substack&utm_medium=email.

(3) ~~Institute reasonable measures to prevent~~ Prevent its companion chatbot from producing ~~visual or facilitating the exchange of~~ material of sexually explicit conduct or directly stating that the minor should engage in sexually explicit conduct.

The removal of “institute reasonable measures” subjects operators to liability for sexually explicit content regardless of the measures they take. And the replacement of “visual” with “facilitating the exchange of” broadens the scope of the provision significantly.

Proponents contend that the bill strengthens and expands protections against exposing kids to sexually explicit outputs from chatbots. Opponents, on the other hand, argue that the change is premature and subjects companies to significant liability risk under SB 243’s private right of action. If that is so, it raises significant questions about the safety of these products for minors.

ARGUMENTS IN SUPPORT: The American Academy of Pediatrics, California writes:

Artificial intelligence technologies are rapidly becoming integrated into the daily lives of young people. While these technologies offer potential benefits, companion chatbots also present significant risks when they are designed to simulate human relationships, provide emotional support, or engage in highly personalized conversations with children and adolescents. SB 300 takes a necessary and measured step to ensure that operators of these platforms implement stronger protections for minors and prevent harmful interactions before they occur.

California's youth are experiencing unprecedented levels of mental health challenges. According to the California Health Care Foundation, approximately one in three California adolescents reports persistent feelings of sadness or hopelessness. Suicide remains one of the leading causes of death among young people in California and nationwide. At the same time, young people are increasingly turning to online platforms and digital technologies for information, companionship, and emotional support. When companion chatbots are designed to mimic human relationships, vulnerable youth may perceive these systems as trusted confidants despite their inability to provide appropriate clinical, developmental, or safety-based responses.

Research demonstrates that adolescents are particularly susceptible to social influence and may have difficulty distinguishing between authentic human relationships and sophisticated AI-generated interactions. Companion chatbots can create a false sense of emotional attachment, potentially increasing risks for youth experiencing loneliness, depression, anxiety, social isolation, or other mental health concerns. These risks are especially concerning for younger users who may not fully understand the limitations of artificial intelligence systems.

SB 300 appropriately recognizes that operators should not be permitted to avoid responsibility simply because they lack direct confirmation of a user's age. By establishing a constructive knowledge standard, the bill ensures that companies must take reasonable steps to protect minors when circumstances indicate that a user is likely underage. This approach reflects the reality of how modern digital platforms collect and utilize information about their users and aligns with California's longstanding commitment to protecting children online.

The bill also strengthens safeguards against sexually explicit content and interactions. Exposure to sexual content during childhood and adolescence has been associated with adverse developmental, emotional, and behavioral outcomes. According to the California Healthy Kids Survey, a substantial proportion of youth report regular online activity and exposure to digital content through social media and other internet-based platforms. As AI systems become increasingly interactive and personalized, it is critical that companion chatbots are prohibited not only from generating sexually explicit content but also from facilitating exchanges or proposals involving such content when minors are involved.

California has consistently led the nation to advance policies that protect children in digital environments, including the California Age-Appropriate Design Code Act and recent legislation addressing online safety. SB 300 builds upon these efforts by addressing emerging risks posed by companion chatbot technologies while preserving opportunities for responsible innovation.

As artificial intelligence continues to evolve, policymakers must ensure that safeguards evolve as well. Children and adolescents deserve protections that reflect both the capabilities and risks of these technologies. SB 300 establishes reasonable, common-sense standards that prioritize child safety, support healthy development, and help prevent foreseeable harms.

ARGUMENTS IN OPPOSITION: TechNet, Chamber of Commerce, and CCIA jointly write:

Expanding Statutory Obligations Before Enforcement

SB 300 would replace key reasonableness-based standards adopted in SB 243 with a mandate to prevent the production of prohibited material and to broaden prohibited conduct to include “facilitating the exchange” of sexually explicit material. These are not clarifying or technical changes; they fundamentally alter the legal standard adopted by the legislature.

At this stage, there is no demonstrated need for these changes and no clear gaps that the proposed language would fill. It is concerning to amend and expand statutory obligations before the Legislature, regulators, courts, companies, or other stakeholders have had any opportunity to observe how the original law operates in practice.

Interaction with the Private Right of Action

These changes are particularly consequential because they are paired with a private right of action already embedded in statute. When combined with a strict “prevent” standard, SB 300 risks turning what was framed as a requirement to implement appropriate protections for children into a de facto strict liability regime.

Under this structure, a single instance of content that is alleged not to have been “prevented” could trigger liability, regardless of the safeguards deployed, the probabilistic nature of AI systems, or the operator’s good-faith efforts to comply.

This approach departs from traditional negligence principles and risks treating entities that take extensive, good-faith compliance measures the same as those that take none.

SB 243 has not yet been fully implemented, interpreted, or enforced. SB 300 would substantially revise that framework before any evidence that the existing law is insufficient,

while simultaneously increasing liability exposure through an already powerful enforcement mechanism.

REGISTERED SUPPORT / OPPOSITION:**Support**

Alameda County Office of Education
American Academy of Pediatrics, California
California Behavioral Health Association
California Initiative for Technology and Democracy (CITED)
Transparency Coalition.ai

Opposition

California Chamber of Commerce
Computer & Communications Industry Association
Technet

Analysis Prepared by: Josh Tosney / P. & C.P. / (916) 319-2200