

California and New York launch AI companion safety laws

October 30, 2025 | Client Update | 6-minute read

Starting November 5, operators of AI companions available in New York must implement a reasonable protocol to detect potential self-harm and disclose that interactions do not involve a human. Operators of AI companions available in California must follow similar requirements starting in January 2026, while also providing special notice to minor users and reporting to California regulators.

Background

New York will become the first state to require certain safety guardrails for AI companions when its [Artificial Intelligence \(AI\) Companion Models](#) law comes into force on November 5, 2025. The law requires operators of AI companions—a form of AI chatbot that simulate human interaction—available to New York residents to take reasonable measures to detect and address expressions of self-harm or suicidal ideation and to provide reminders that a user is not communicating with a human. California will follow quickly on January 1, 2026, when its recently signed [Companion Chatbot](#) law goes into effect. The new California law similarly targets companion chatbots and requires disclosure and protocols for responding to users in crisis. But it also requires additional notifications to known minor users and establishes a reporting regime to the California Department of Public Health's Office of Suicide Prevention.

Although other states, such as Maine, Texas and Utah, have enacted transparency laws requiring AI chatbots to disclose that users are not interacting with a human, New York and California are the first states to impose special obligations on providers of AI chatbots designed to simulate human relationships.

Applicability

New York's AI Companion Models law defines "AI companions" to include any system using AI, generative AI, or emotional recognition algorithms to simulate human relationships with users by remembering interactions, asking users unsolicited, emotion-based questions, and sustaining ongoing dialogue about users' personal matters. The law defines "human relationships" broadly, to include intimate, romantic, platonic or other forms of interaction or companionship. Notably, the law exempts several, common AI chatbot applications such as systems used by a business only for customer service purposes, employee productivity or other internal purposes. It also exempts AI chatbots that are designed and marketed for providing efficiency improvements, research or technical assistance.

Similarly, California's Companion Chatbot law applies to natural language interfaces that provide humanlike responses and sustain relationships across multiple interactions. As with New York, California exempts AI chatbots used only for certain tasks such as consumer service, internal productivity purposes, research or technical assistance. California also carves out chatbots within video games that can discuss only game-related topics—aside from certain high-risk topics related to health and sexuality—and voice-activated virtual assistants on consumer devices that do not sustain relationships across interactions and are not likely to elicit emotional responses from users.

Requirements

New York's AI Companion Models law has two primary requirements:

- First, it requires operators of AI companions to ensure their systems contain a reasonable protocol to detect and address an expression of suicidal ideation or self-harm by users. Although the law defines “self-harm” to include all forms of intentional self-injury, it does not explain what constitutes an “expression” of self-harm. Nor does the law explain what is “reasonable,” aside from referring users to crisis services like the 998 Lifeline or a crisis text line.
- Second, it requires operators to clearly and conspicuously notify users—verbally or in writing—that they are not communicating with a human. These notifications are required both at the beginning of a user interaction, and then again every three hours during “continuing” interactions. The law does not specify how to determine when one interaction ends and another begins, for purposes of calculating three-hour periods.

California's Companion Chatbot law also requires user notice and safety protocols, but also requires annual reporting to regulators and publication of related information.

- As with New York, the California law requires that companion chatbots employ protocols for responding to expressions of suicidal ideation or self-harm by users that include, but are not limited to, referring users to crisis service providers. California requires operators to measure for “suicidal ideations” using evidence-based methods, but goes no further in specifying what appropriate methods might be.
- California also imposes two separate notification requirements. First, for all users the operator knows are minors, it must disclose that the user is interacting with AI, provide recurring notifications every three hours, and institute measures to prevent the companion chatbot from producing sexually explicit materials or encouraging the minor to engage in such conduct. Second, the operator must provide clear and conspicuous notice that the companion chatbot is not human, but only if a reasonable person would be misled to believe otherwise.
- Finally, California's Companion Chatbot law establishes a new reporting regime for operators of AI companion chatbots. Beginning July 1, 2027, operators must submit annual reports to the Office of Suicide Prevention detailing their protocols for responding to suicidal ideation by users and for preventing the chatbot from engaging with users about suicidal ideation, as well as the number of times the operator referred users to a crisis service provider in the preceding calendar year. Additionally, operators must publish data from these reports, as well as details about their chatbots' protocols, on their websites.

Enforcement

New York's AI Companion Models law allows the New York Attorney General to seek civil penalties up to \$15,000 per day. The New York Attorney General may also seek injunctive relief where the Attorney General believes that operators of AI companions have violated or are about to violate the law.

Although the California Companion Chatbot law does not provide for civil penalties, it does allow a private right of action for any person harmed by a violation of the law for the greater of actual damages or \$1,000 per violation, injunctive relief, and reasonable attorney's fees and costs.

Key takeaways

- Companies should assess potential applicability of the New York AI Companion Models and California Companion Chatbot laws by comparing any AI companion offerings against the laws' exempted use cases.
- If applicable, companies should consider what constitutes a “reasonable” protocol given multiple ambiguities in the New York AI Companion Models law, including with respect to what constitutes an “expression” of self-harm.
- Companies should document the “evidence” their protocols rely on in assessing user expressions of self-harm and the standards they employ to evaluate explicit or similarly covered content.
- Companies subject to California's Companion Chatbot law should monitor the required reporting metrics on a regular basis as part of their oversight and governance functions.

More broadly, companies should review existing and planned chatbot deployments against requirements in other states, including Maine, Texas and Utah, as well as potential federal legislation, including the bipartisan legislation introduced in the Senate on October 28, 2025 that would ban minors from accessing AI companions. Chatbot offerings for minors are

also the focus of ongoing regulatory and enforcement activities—for example, the FTC’s 6(b) [study](#) of the potentially negative impacts of the technology on children and teens and the Texas Attorney General’s ongoing [investigation](#) into AI chatbots offered as mental health or counseling tools to children and other vulnerable individuals.

If you have any questions regarding the matters covered in this publication, please reach out to any of the lawyers listed below or your usual Davis Polk contact.

Matthew J. Bacal

+1 212 450 4790
matthew.bacal@davispolk.com

David I. Feinstein

+1 212 450 3293
david.feinstein@davispolk.com

James W. Haldin

+1 212 450 4059
james.haldin@davispolk.com

David Lisson

+1 650 752 2013
david.lisson@davispolk.com

Howard Shelanski

+1 202 962 7060
howard.shelanski@davispolk.com

This communication, which we believe may be of interest to our clients and friends of the firm, is for general information only. It is not a full analysis of the matters presented and should not be relied upon as legal advice. This may be considered attorney advertising in some jurisdictions. Please refer to the firm's privacy notice for further details.