

Publications

July 18, 2025 • Updates

Regulatory Trend: Safeguarding Mental Health in an AI-Enabled World

Key Takeaways:

- New York is the first state to enact mental health-focused statutory provisions for “AI Companions,” requiring user disclosures and suicide prevention measures for emotionally interactive AI systems.
- Other states are exploring similar approaches, with laws targeting compulsive use, requiring suicide prevention protocols or mandating user awareness of AI-human distinctions.
- Organizations must assess their AI risk to ensure compliance with the myriad laws and statutory provisions governing AI systems.

New York, as part of its state budget process, enacted in May 2025 new statutory provisions for “AI Companions” that highlight an emerging desire to monitor and safeguard the mental health of AI tool or system users. It aligns with a broader regulatory awareness of the mental health risks involved in AI interactions and the desire to safeguard vulnerable AI users, particularly minors or those experiencing mental health crises like suicidal ideation.

An Emerging Desire to Safeguard Mental Health in an AI-Enabled World

Regulators are increasingly aware of the mental health risks involved in AI interactions and seeking ways to safeguard vulnerable users. These risks were brought into sharp focus with the death of a 14-year-old Florida teenager, Sewell Setzer, who committed suicide after forming a romantic and emotional relationship with an AI chatbot and allegedly informing the chatbot that he was thinking about suicide, which has resulted in a closely watched lawsuit regarding the chatbot’s role in his death.

States have considered a variety of techniques to regulate this space, ranging from user disclosures to safety measures. Utah’s law on mental health chatbots (H.B. 452), for example, imposes advertisement restrictions and requires certain disclosures to ensure users are aware they are interacting with an AI rather than a human being. Other states, like California (via SB 243), are considering design mandates like banning reward systems

Related People

- Paul D. Squire
- Jennifer Bauer
- Ashleigh Bickford

Related Capabilities

- Artificial Intelligence & Machine Learning
- Health Care
- Digital Health

that encourage compulsive use and requiring suicide prevention measures within any AI chatbots that are being marketed as emotional buddies. Currently, NY is the only state that has enacted safety-focused measures (like suicide prevention) around AI companionship.

NY's Approach to Embedding Mental Health Safeguards in AI

NY's new statutory provisions (which go into effect on November 5, 2025) focus on AI systems that retain user information and preferences from prior interactions to engage in human-like conversation with their users.

These systems, termed "AI Companions," are characterized by their ability to sustain ongoing conversations about personal matters, including topics typically found in friendships or emotionally supportive interactions. That means chatbots, digital wellness tools, mental health apps or even productivity assistants with emotionally aware features could fall within the scope of AI Companions depending on how they interact with users, although interactive AI systems used strictly for customer service, international operations, research and/or productivity optimization are excluded.

The law seeks to drive consumer awareness and prevent suicide and other forms of self-harm by mandating such AI systems (1) affirmatively notify users they are not interacting with a human and (2) take measures to prevent self-harm. Operators must provide clear and conspicuous notifications at the start of any interaction (and every three hours for long and ongoing interactions) to ensure users are aware they're not interacting with a human. Operators must also ensure the AI system has reasonable protocols to detect suicidal ideation or expressions of self-harm expressed by a user and refer them to crisis service providers like the 988 Suicide Prevention and Behavioral Health Crisis Hotline whenever such expressions are detected.

Assessing AI Regulatory Risk

Whether in the context of chatbots, wellness apps, education platforms or AI-driven social tools, regulators are increasingly focused on systems that engage deeply with users. Because these systems may be uniquely positioned to detect warning signs like expressions of hopelessness, isolation or suicidal ideation, it's likely that other states will follow NY in requiring certain AI systems to identify, respond to or otherwise escalate signals of mental health distress to protect vulnerable populations like minors.

NY's new AI-related mental health provisions also showcase how U.S. laws and statutory provisions around AI heavily focus on how the technology is being used. In other words, your use case determines your risk. To effectively navigate the patchwork of AI-related laws and statutory provisions in the U.S. — of which there are over 100 state laws currently — organizations must evaluate each AI use case to identify their compliance risks and obligations.

Polsinelli offers an AI risk assessment that enables organizations to do exactly that. Understanding your AI risks is your first line of defense — and a powerful business enabler. Let us help you evaluate whether your AI use case falls within use case or industry-specific laws like NY's "AI Companion" law or industry-agnostic ones like Colorado's AI Act, so you can deploy innovative business tools and solutions with confidence.