

New York AI Companion Models Law

What Companies Need to Know

Redefining Responsible AI: What the New York Law Signals

New York's Artificial Intelligence (AI) Companion Models law takes effect November 5, 2025, establishing the first state-level safety standards for AI "companions" — systems that simulate human interaction and sustain personal dialogue with users. The law requires safety protocols and recurring disclosures to reduce risks of self-harm and other harms. Companies offering AI companions to New York users should begin compliance preparations now. California will follow on January 1, 2026, with a similar law that adds requirements for minors and reporting.

Scope and Applicability in NY

The New York law applies to "operators" that provide AI companions to users in the state. An "AI companion" is defined as a system using artificial intelligence, generative AI, and/or emotional recognition algorithms to simulate social human interaction by retaining information from prior interactions, asking questions, offering advice, and engaging in simulated conversation on matters of personal well-being. The definition excludes systems used solely for customer service, customer account information, or other information related to a user's relationship with a business, as well as internal productivity or technical assistance tools.

Companies should carefully map each AI product or feature against the law's definition and exemptions. If an AI system sustains relationship-like interactions across sessions, remembers user preferences, engages in personal well-being dialogue, and exhibits anthropomorphic dynamics, it is likely to be captured.

New York's Core Requirements

The law imposes two categories of obligations on operators:

- 1. Implement Safety Protocols**
 - Protocols should address user-expressed risk including possible suicidal ideation or self-harm, physical harm to others, and financial harm to others.
 - Protocols must include a notification referring the user to crisis service providers.
 - Ensure protocols are reasonable, evidence-informed, and operationalized across models, interfaces, and channels where companion interactions occur.
- 2. Provide Clear Disclosures**
 - Operators must provide clear disclosures at the start of every AI companion interaction and again at least every three hours during continuing interactions
 - The disclosure must state verbally or in bold, capital letters of at least 16-point type: "THE AI COMPANION IS A COMPUTER PROGRAM AND NOT A HUMAN BEING. IT IS UNABLE TO FEEL HUMAN EMOTION."

New York AI Companion Models Law

What Companies Need to Know

Enforcement and Liability in New York

The law allows individuals harmed through self-harm or by another's physical or financial actions to sue for violations of safety or disclosure rules. Plaintiffs can seek damages or equitable relief, and operators should expect challenges around causation – making thorough records of triggers, responses, and essential disclosures.

Immediate Compliance Steps

- ➔ **Assess applicability.** Inventory AI features accessible in NY. Identify systems that mimic human interaction, retain memory, or discuss well-being, and document exemptions for customer-service or internal tools.
- ➔ **Design and implement risk detection.** Implement model- and rule-based tools to detect self-harm, violence, or financial-harm risks, defining triggers, thresholds, and fallback rules to balance accuracy and ambiguity.
- ➔ **Define response protocols.** For each risk type, script empathetic responses that avoid reinforcement, direct users to crisis services, and escalate when needed. Ensure referrals are clear, accessible, and localized.
- ➔ **Instrument disclosures.** Embed the required disclosure at session start and resurface it every three hours. Use the statute's required language and ensure verbal and written notices are clear and conspicuous.
- ➔ **Clarify session boundaries.** Define session start and end points for disclosure timing. Account for idle timeouts, re-authentication, or user re-entry, and apply logic consistently across platforms.
- ➔ **Establish auditability.** Log detection events, responses, disclosures, and session data with appropriate retention and access controls.
- ➔ **Tune model policies and train teams.** Enhance model safety frameworks and ensure teams are trained to address escalations, interpret ambiguous content, and support accessibility and multilingual scenarios in line with protocol standards.

California's Companion Chatbot Law - Key Differences to Plan For

California's law, effective January 1, 2026, targets "companion chatbots" that provide adaptive, human-like responses and sustain relationships across interactions. Like New York, it requires crisis-response protocols and disclosures. It adds several obligations:

01. Minor-specific requirements

For users known to be minors, operators must disclose that the user is interacting with AI, provide recurring notifications at least every three hours, and implement reasonable measures to prevent the chatbot from producing sexually explicit material or encouraging such conduct.

02. Protocol transparency and reporting

Operators must publish details of their suicidal ideation prevention protocol on their website. Beginning July 1, 2027, operators must annually report to the Office of Suicide Prevention on crisis referral counts and protocols for detecting, removing, and responding to suicidal ideation, and the office will post data from those reports online.

03. Conditional disclosure for all users

Operators must issue a clear and conspicuous notice that the chatbot is artificially generated and not human if a reasonable person could be misled to believe they are interacting with a human.

04. Private right of action

California authorizes individuals harmed by noncompliance to seek injunctive relief, damages (the greater of actual damages or \$1,000 per violation), and reasonable attorneys' fees and costs.

Companies operating nationwide should harmonize New York and California implementations, with California's minor-focused requirements and reporting regime layered on top.



Practical Takeaways

New York's law demands immediate, operational readiness for companion interactions, particularly around detection and responsive protocols for self-harm, threats of harm to others, and possible financial harm. The recurring, conspicuous "not human" disclosure, including precise language and formatting, is not optional and must be engineered into product flows and audio experiences. California's forthcoming law will extend these obligations and introduce minor protections and reporting. Companies should approach compliance as a cross-functional program with robust detection, response, disclosure instrumentation, logging, and governance.