

California se convirtió en el primer estado en regular los chatbots de acompañamiento de inteligencia artificial

14/10/2025



Ante el avance exponencial de la Inteligencia Artificial (IA) y la aparición desenfrenada de **herramientas y aplicaciones**, el gobernador del Estado de California – en EEUU -, Gavin Newsom, **promulgó una nueva** ley para regular los chatbots acompañantes de IA. La misma pone el foco sobre la responsabilidad de los operadores de estas tecnologías y les **exigirá la implementación de protocolos de seguridad**.

En detalle, la ley fue presentada por los senadores estatales Steve Padilla y Josh Becker en enero de este año. Conocida como SB245, el texto legislativo busca **proteger a niños y usuarios vulnerables de algunos daños asociados con el uso de chatbots de IA**.

El debate alrededor de estas herramientas estuvo presente desde un inicio, pero se intensificó en Estados Unidos durante los últimos meses, luego de la [muerte del adolescente, Adam Reine](#), quien se suicidó tras una larga serie de conversaciones con ChatGPT de OpenAI. Recientemente, una familia del estado de Colorado presentó una demanda contra la startup de juegos

de rol Character AI, luego de que su hija de 13 años se quitara la vida tras una serie de conversaciones problemáticas y sexualizadas con los chatbots desarrollados por la empresa.

Chatbots de IA: la regulación pionera en California

En pocas palabras, la nueva ley exige responsabilidad legal a las empresas que desarrollen Chatbots IA, ya sean grandes laboratorios – como **Meta u OpenAI** – como también startups como la mencionada Character AI. Al respecto, el gobernador Newsom detalló: «Las tecnologías emergentes, como los chatbots y las redes sociales, pueden inspirar, educar y conectar; pero sin unas medidas de seguridad reales, la tecnología también puede explotar, engañar y poner en peligro a nuestros hijos. Hemos visto ejemplos verdaderamente horribles y trágicos de jóvenes perjudicados por tecnología no regulada, y no nos quedaremos de brazos cruzados mientras las empresas sigan actuando sin los límites ni la rendición de cuentas necesarios».



El gobernador firmó el acta de la ley presentada en enero de este año.

La nueva medida entrará en vigencia recién el 1 de enero de 2026 y exigirá a las empresas implementar ciertas funciones, como la verificación de edad y advertencias sobre redes sociales y chatbots acompañantes. La ley también impone sanciones más severas para quienes se lucren con deepfakes ilegales, incluyendo **hasta u\$s250.000 por infracción**.

Más en detalle, las empresas desarrolladoras también deberán establecer protocolos para abordar el **suicidio y las autolesiones**, que será compartido con el Departamento de Salud Pública del estado.

El texto también detalla otras medidas, entre las que se encuentran:

- Si una persona razonable que interactúa con un chatbot acompañante es engañada y creyera que está interactuando con un humano, el desarrollador deberá emitir una notificación clara y visible que indique que la herramienta acompañante «es generada artificialmente y no es humana».
- Se deberá evitar que el chatbot interactúe con los usuarios a menos que el desarrollador **mantenga un protocolo para prevenir la producción de contenido de ideación suicida, suicidio o autolesión para el usuario**. Esto deberá incluir proporcionar una notificación al usuario que lo remita a **proveedores de servicios de crisis**, como líneas directas de ayuda de prevención del suicidio, en caso de que la persona exprese ideación suicida, suicidio o autolesión.
- Para los usuarios menores de edad, el desarrollador deberá notificar a la persona que está interactuando con una IA. En caso de interacciones continuadas en el tiempo, cada tres horas se deberá notificar al ciudadano que el «debe tomar un descanso y que el chatbot complementario se genera artificialmente y no es humano».

- Establecer medidas razonables para evitar que el chatbot que lo acompaña produzca material visual de conducta sexualmente explícita o indique directamente que el menor debe participar en una conducta sexualmente explícita.

Varias compañías tecnológicas comenzaron a aplicar **protocolos de seguridad orientados a proteger a los menores de edad** en el uso de inteligencia artificial. Entre ellas, **OpenAI** puso en marcha recientemente **controles parentales, filtros de contenido y un sistema de detección de conductas autolesivas** dentro de ChatGPT, en un intento por reducir riesgos asociados al uso de la herramienta por parte de adolescentes.

En paralelo, **California avanza con una batería de regulaciones sobre IA**. La **SB 243**, aprobada esta semana, representa la **segunda norma relevante** en pocas semanas. Previamente, el **29 de septiembre**, el gobernador **Gavin Newsom** había promulgado la **SB 53**, una ley que **impone nuevas exigencias de transparencia a las grandes empresas del sector**.

La normativa obliga a **laboratorios como OpenAI, Anthropic, Meta y Google DeepMind** a **publicar sus protocolos de seguridad** y a **proteger a los empleados que actúen como denunciantes** dentro de esas compañías.

Fuente: **Ámbito Financiero**