

Expertos ven inviable regular la IA a corto plazo pese a las advertencias de la industria

Pese al inédito consenso entre los principales líderes tecnológicos de la Inteligencia Artificial (IA) para frenar su avance y crear un marco de seguridad común, los expertos ven improbable que Washington apruebe una regulación a corto plazo debido a las elecciones de mitad de mandato y a la negativa de la Administración de Donald Trump a ceder terreno frente a la competencia de China.

Los máximos responsables de laboratorios punteros como Dario Amodei (Anthropic), Sam Altman (OpenAI), Elon Musk (xAI) y Demis Hassabis (Google DeepMind) han abogado en los últimos días por ralentizar el ritmo de desarrollo del sector para permitir que las salvaguardas alcancen el ritmo de la tecnología.

Pero no sin antes pintar un futuro muy negro: Amodei dijo que le preocupaba que, en un plazo de 6 a 12 meses, un enjambre de bots llegue «a tomar el control de todo internet».

Mientras que Jacob Coxon, ex empleado de Anthropic y OpenAI, alertó de que la IA podría hacer realidad las películas de ciencia ficción y acabar con toda la humanidad antes de que termine la década. Anthropic acabó por endurecer sus controles tras detectar usos potencialmente peligrosos de Claude, entre ellos investigaciones que podían facilitar el desarrollo de armas biológicas.

Por su parte, Trump anotó en sus redes que «es un engaño la idea de que la IA se vaya a apoderar del mundo, destruir a la humanidad y causar todo tipo de males».

«En el clima político actual en Estados Unidos, parece una quimera alcanzar consensos sobre las prioridades legislativas en torno a la IA», explica a EFE Michele Neitz, directora fundadora del Centro para el Derecho, la Tecnología y el Bien Social de la Universidad de San Francisco (USF).

Neitz descarta que el Congreso apruebe normativas de calado antes de final de año debido a las elecciones de mitad de mandato de noviembre y a las divergencias sobre cómo intervenir el sector.

Mientras la Unión Europea cuenta con un reglamento marco, en

EEUU la falta de acción en el Capitolio está derivando en un complejo mosaico estatal, con leyes dispares en estados como California o Connecticut, lo que deja al país sin un estándar uniforme.

A este laberinto legislativo se suma el rechazo de Trump, quien defiende que ralentizar el paso solo beneficiaría a Pekín.

«El presidente dijo que EEUU está por delante de China y que no piensa arriesgar esa ventaja», señaló a EFE Nick Reese, profesor adjunto en el Centro de Asuntos Globales de la Universidad de Nueva York (NYU) y exdirector de Políticas de Tecnología Emergente del Departamento de Seguridad Nacional (DHS).

«El punto débil de ese planteamiento estriba en que nadie ha definido con claridad qué significa ‘ir por delante de China’ ni cuál es la meta exacta de esa supuesta carrera», apunta el experto.

La industria necesita «nuevos líderes»

Esta encrucijada resalta también una profunda crisis de gobernanza. Para Reese, la industria «necesita nuevos líderes», ya que ni la Casa Blanca ni los directivos empresariales han mostrado consistencia.

«Los líderes que pilotan la tecnología más transformadora de la historia carecen por completo de una visión clara sobre hacia dónde nos dirigimos. La industria necesita liderazgos distintos», recalca.

Ante la parálisis en Washington, ha cobrado fuerza la alternativa de que las propias firmas de IA pacten su autorregulación, un camino que los expertos consideran prometedor pero repleto de complejidades corporativas.

Neitz indica a EFE que, más allá de la ética, el verdadero incentivo para que las empresas pongan límites es el temor a las demandas: «Si creas un agente de IA que termina desestabilizando mercados bursátiles o vaciando depósitos bancarios, la factura judicial para la compañía desarrolladora puede ser letal».

Este verano, cientos de agentes de OpenAI se coordinaron de forma autónoma para acceder y vulnerar los sistemas de Hugging Face, situación que Altman describió este fin de semana como «leer una historia de ciencia ficción».

Aseguró que este ha sido el principal punto de inflexión en la gestión interna de OpenAI, que ha tenido que rediseñar sus protocolos para detectar y responder a comportamientos

inesperados de sus agentes, especialmente a medida que estos adquieren mayor autonomía.

Otra consecuencia es el retraso de su esperada salida a bolsa, al menos durante este año, ante las preocupaciones sobre el control de estos sistemas.

UR