

China y EEUU buscan un 'teléfono rojo' frente a los riesgos de la IA

Los riesgos de la inteligencia artificial (IA), desde su uso con fines maliciosos hasta el temor de que tome decisiones militares de forma autónoma, están llevando a China y Estados Unidos a estudiar vías de diálogo para evitar incidentes.

El alcance de este diálogo, que sugirió el domingo sin muchos detalles el secretario del Tesoro estadounidense, Scott Bessent, debería ser limitado.

Por el momento, el secretario estadounidense se limitó a mencionar la necesidad de pasar «de una situación de opacidad a una mayor transparencia».

La propuesta refleja la preocupación por el potencial destructivo de la IA, que podría actuar de forma imprevista o acelerar sin control la toma de decisiones militares.

«Por ambas partes se observa una toma de conciencia de que existe la posibilidad de incidentes y usos indebidos en detrimento de los intereses de los dos países», señala a la AFP Jean-Marc Rickli, director de Riesgos Globales y Emergentes del Centro de Política de Ginebra.

«El 'teléfono rojo' se estableció a raíz de la crisis de Cuba» de 1962, recuerda el experto, en referencia al sistema de comunicación directa entre Estados Unidos y Rusia en plena Guerra Fría para evitar incidentes militares graves.

Rickli recuerda un incidente reciente en el que, según la cadena CNN, Estados Unidos estuvo a punto de abordar un barco chino basándose en información falsa elaborada con IA.

Distintas prioridades

Desde 2023 existe un diálogo bilateral entre expertos y centros de investigación, pero sin intervención del gobierno, sobre los modelos de IA más avanzados de compañías como OpenAI, Anthropic, DeepSeek o Alibaba.

Según la investigadora Mathilde Viellet, del Instituto Francés de Relaciones Internacionales (Ifri), ese diálogo ha puesto de relieve que las prioridades de las dos potencias son distintas.

#photo2

Mientras China lleva años preocupada por la seguridad o luchando contra los 'deepfakes', Estados Unidos «se preocupa sobre todo por la pérdida de control y la 'desalineación' de la IA, cuando actúa de forma autónoma», explica la experta a la AFP.

Por el momento, según la consultora estadounidense Teneo, «ninguno de los dos Gobiernos está dispuesto a asumir compromisos vinculantes sobre el desarrollo de la IA».

Además, Estados Unidos apuesta por modelos de IA cerrados y de propiedad privada para preservar su ventaja comercial, mientras que China opta por el código abierto para acelerar los avances tecnológicos.

Según Jean-Marc Rickli, tendría sentido crear un sistema de comunicación entre China y Estados Unidos en caso de una decisión militar en la que la IA tome decisiones demasiado rápidas o basadas en información errónea.

Por su parte, Mathilde Viellet aboga por un 'teléfono rojo' que sirva en caso de pérdida de control de la IA y para evitar su utilización por parte de actores maliciosos.

En septiembre, Anthropic reveló cinco casos de utilización de su modelo de IA Claude de una forma «que podría servir para el desarrollo de armas biológicas».

Pero aunque llegue a establecerse «cierta coordinación flexible», el investigador estadounidense Matt Sheehan, cree que «las medidas más importantes en materia de gobernanza y seguridad de la IA serán las que adopte cada país a nivel nacional».

AFP