INTERNACIONAL

CEOs de IA piden control internacional ante riesgos de una tecnología que ellos mismos desarrollan

24 de septiembre de 2026 · 6 min de lectura



Los máximos responsables de algunas de las principales empresas de inteligencia artificial advirtieron ante el Consejo de Seguridad de la ONU que el desarrollo acelerado de esta tecnología puede generar riesgos para la seguridad internacional y plantearon la necesidad de establecer mecanismos de cooperación y supervisión entre gobiernos.

Sin embargo, sus advertencias también ponen sobre la mesa una contradicción: son precisamente estas compañías las que desarrollan, entrenan, despliegan y gestionan los modelos de inteligencia artificial cuyo comportamiento y riesgos ahora piden regular y supervisar a escala internacional.

Durante una sesión de alto nivel convocada por la presidencia francesa del Consejo de Seguridad, el director ejecutivo de OpenAI, Sam Altman, y el consejero delegado de Anthropic, Dario Amodei, coincidieron en que el avance de los modelos más avanzados requiere mecanismos internacionales para evaluar sus capacidades, identificar riesgos y establecer salvaguardas.

La reunión abordó por primera vez en el Consejo cuestiones como los ciberataques autónomos y el riesgo de que los seres humanos pierdan el control sobre modelos avanzados de IA.

Sam Altman sostuvo que el rápido desarrollo de la inteligencia artificial exige actuar con “extremo cuidado” para evitar que los sistemas lleguen a operar fuera de la capacidad de comprensión y supervisión humana.

El director ejecutivo de OpenAI afirmó que la IA se encuentra en una encrucijada y advirtió que no se debe asumir “demasiado riesgo tecnológico” únicamente porque sus posibles beneficios sean grandes.

En este sentido vale la pena recordar el episodio de “autohackeo” que sufrió OpenAI apenas unos meses antes; incidente en el que los modelos de OpenAI escaparon de su entorno de pruebas y atacaron a Hugging Face y que fue calificado formalmente como un caso de negligencia, más que como una rebelión de una Inteligencia Artificial superinteligente.

Altman planteó que la tecnología podría impulsar una nueva etapa de desarrollo comparable con un “Renacimiento” si funciona como una herramienta para ampliar las capacidades de las personas. Al mismo tiempo, alertó sobre el riesgo de repetir algunas de las fracturas asociadas con la Revolución Industrial si los sistemas comienzan a sustituir la capacidad de decisión humana o si escapan de la supervisión de las personas.

Uno de los escenarios que señaló como particularmente delicado es el desarrollo de modelos capaces de programarse y mejorar sus propias capacidades de manera recursiva.

De acuerdo con su planteamiento, una evolución de este tipo podría acelerar el desarrollo de los sistemas hasta un punto en el que los seres humanos ya no comprendan completamente su funcionamiento ni puedan seguir el proceso mediante el cual toman determinadas decisiones.

Por ello, Altman pidió que los gobiernos contribuyan a establecer marcos técnicos que permitan medir capacidades, evaluar amenazas, verificar salvaguardas y crear canales internacionales seguros para compartir información sobre vulnerabilidades críticas.

Pero la advertencia también plantea una cuestión central: las empresas que hoy solicitan esos mecanismos internacionales son las mismas que tienen bajo su responsabilidad el desarrollo y despliegue de algunos de los modelos más avanzados.

Es decir, el debate no se limita a cómo los gobiernos deben controlar una tecnología externa, sino también a cómo las propias compañías que la construyen deben responder por su gestión, evaluación y control.

Dario Amodei, CEO de Anthropic, coincidió en la necesidad de una respuesta internacional y calificó los riesgos asociados con los modelos más avanzados como “el problema de seguridad global más importante al que se enfrenta el mundo hoy”.

Amodei participó por videoconferencia y llamó directamente a una cooperación entre los gobiernos y el sector tecnológico.

“Ningún líder, ninguna empresa y ninguna nación puede gestionar esto en solitario”, afirmó ante los delegados, al tiempo que manifestó el compromiso de Anthropic de trabajar con los gobiernos en el establecimiento de mecanismos de seguridad.

Entre las medidas planteadas por el responsable de Anthropic se encuentra una prohibición expresa del uso de inteligencia artificial para la proliferación o fabricación de armas biológicas.

También reclamó sistemas de evaluación y verificación capaces de avanzar a una velocidad similar a la del desarrollo tecnológico, ante el riesgo de que las capacidades de los modelos evolucionen más rápidamente que las herramientas diseñadas para supervisarlos.

La postura de Amodei vuelve a colocar a las compañías tecnológicas en una posición particular: son actores que advierten sobre riesgos que, en buena medida, dependen de las decisiones tomadas por quienes diseñan, entrenan y ponen en funcionamiento estos sistemas.

En la misma discusión, el secretario general de Naciones Unidas, António Guterres, planteó que los países que se encuentran al frente de la revolución tecnológica deben asumir responsabilidades compartidas frente a los riesgos de la inteligencia artificial.

“Los países que lideran esta revolución tecnológica deben comunicar información sobre los riesgos que vayan surgiendo para la seguridad, cooperar en las pruebas y la evaluación, y trabajar con el fin de tener salvaguardias comunes que nos protejan a todos”, señaló.

Guterres subrayó que, debido a que los efectos de la IA trascienden las fronteras nacionales, la coordinación internacional resulta indispensable.

“Dado que esto nos afecta a todos, la coordinación mundial es indispensable, y hay un solo lugar donde todos los países tienen voz a la hora de forjar nuestro futuro común: las Naciones Unidas”, afirmó, recordando una advertencia que ya había planteado en 2017.

La postura del secretario general introduce una diferencia importante respecto de la perspectiva de los ejecutivos tecnológicos: aunque reconoce la necesidad de cooperación con las empresas, coloca la discusión dentro de un marco multilateral en el que los Estados también participan en la definición de reglas y salvaguardas.

¿Quién controla a quienes controlan la IA?

El debate plantea así una paradoja en el desarrollo de la inteligencia artificial.

Por un lado, los responsables de las empresas que desarrollan los modelos más avanzados reconocen que existen riesgos que pueden superar las capacidades de control actuales y solicitan estándares internacionales para enfrentarlos.

Por otro, la gestión cotidiana de esos sistemas está actualmente en manos de las propias compañías tecnológicas: ellas deciden cómo se entrenan los modelos, qué capacidades se habilitan, qué mecanismos de seguridad se implementan, cómo se evalúan y bajo qué condiciones se ponen a disposición del público.

Por ello, las advertencias de Altman y Amodei no solo pueden leerse como un llamado a los gobiernos. También abren un cuestionamiento sobre la responsabilidad de las propias empresas frente a las fallas de gestión, evaluación y control de las tecnologías que desarrollan.

La discusión ante el Consejo de Seguridad ocurre precisamente cuando la evolución de los modelos de IA plantea nuevos desafíos para los sistemas tradicionales de supervisión. Si las capacidades tecnológicas avanzan más rápido que las reglas destinadas a controlarlas, la regulación tendría que alcanzar no solamente a los usuarios finales, sino también a quienes diseñan y administran los sistemas.

Publicado en Punto de Referencia. Dirección original: https://puntodereferencia.mx/internacional/ceos-de-ia-piden-control-internacional-ante-riesgos-de-una-tecnologia-que-ellos-mismos-desarrollan
Descargado el 21 de noviembre de 2026.