



[NATE SOARES, INFORMÁTICO Y COAUTOR DEL LIBRO “SI ALGUIEN LA CREA, TODOS MORIREMOS”:]

“Lo importante es si (las IA) pueden escapar, replicarse o construir sus infraestructuras”

Científicos destacados a nivel mundial advierten de la creación de una superinteligencia artificial (SIA), debido a las dudas sobre las capacidades de las actuales, sumado a que “nadie sabe dónde está el punto de no retorno”.

V.B.V. / Agencia EFE

Las actuales inteligencias artificiales (IA) ¿son un juego de niños comparados con las amenazantes super-IA que podrían llegar? Eliezer Yudkowsky, considerado una de las 100 personas más influyentes en IA por la revista TIME; y Nate Soares, presidente del Instituto de Investigación de Inteligencia Artificial (MIRI), señalaron que sí y alertaron de los peligros de que una máquina sea más inteligente que cualquier humano, o que la humanidad en su conjunto.

Los investigadores escribieron el libro “Si alguien la crea, todos moriremos”, que se ha situado entre los más vendidos de no ficción en Estados Unidos, desatando un intenso debate sobre si los riesgos que describen son reales o se trata sólo de argumentaciones teóricas.

Esto ha desencadenado reacciones polarizadas por presentar escenarios de “todo o nada”, como la supervivencia o la extinción, además de aplausos entre quienes abogan por regulación y gobernanza tecnológica.

Soares dijo a EFE que detener la carrera “absolutamente imprudente” hacia la “superinteligencia” es posible y no tiene por qué afectar a la vida cotidiana.

El autor no habló de cerrar ChatGPT, de renunciar a los autos sin conductor o abandonar la IA aplicada al descubrimiento de nuevos fármacos, sino de desistir de la carrera por una inteligencia “que nadie, ni siquiera los que las



LA SUPER-IA PODRÍA ADQUIRIR CONSCIENCIA DE SÍ MISMA Y CREAR SU PROPIA INFRAESTRUCTURA O ARRASAR CON LO QUE YA ESTÁ.

construyen, comprende”.

Yudkowsky y Soares son el presidente fundador y el director de MIRI, que nació hace 25 años para investigar el desarrollo de la IA y ahora centra su labor en desentrañar el peligro potencial de estas tecnologías, en la necesidad de asegurar que se sustenten en programas “amigables”.

SISTEMAS “TONTOS”

Los dos firmaron en 2023, junto a miles de científicos de todos el mundo -entre ellos varios premios Nobel- una carta para solicitar pausar durante un tiempo los experimentos con la IA más potente y avanzada, junto con ralentizar su desarrollo.

Soares señaló que, a su juicio, la misiva se quedó “muy corta” porque “su-



NATE SOARES TRABAJÓ EN GOOGLE Y MICROSOFT, ENTRE OTRAS.

bestimaba terriblemente” el problema, ante la posibilidad real de que existan “super-IA” (SIA) capaces de desarrollar objetivos propios en conflicto con los humanos.

Por lo que se sabe, esta SIA aún no ha llegado y el libro puede interpretarse

como un relato de ciencia ficción o como un alegato a favor de una legislación responsable y estricta, junto a un control absoluto de la investigación en estas tecnologías, a fin de evitar que se terminen convirtiendo en la mayor amenaza para la humanidad.

Soares indicó que las IA “siguen siendo tontas” y no son gran cosa comparadas con las “super-IA” que las empresas intentan crear. No obstante, señaló que las diferencias radicales que se pueden producir llevan al ejemplo del hombre y el chimpancé, ya que ambos tienen un cerebro similar, pero los humanos cruzaron un umbral y son capaces de diseñar cohetes y pisar la Luna.

“No sabemos cuándo cruzará ese umbral la IA, pero sí sabemos que es posible”, agregó el investigador, para quien desarrollar una superinteligencia que “nadie” comprende es “una receta para el desastre”, y el primer paso debería ser que los líderes mundiales entendieran y asumieran que esta tecnología “puede ser letal”.

Según él, las advertencias, las amenazas y los peligros de esa superinteligencia la han descrito ya algunos de los directores de laboratorios, investigadores o empresarios, entre ellos Sam Altman (OpenAI), Elon Musk (X) o Geoffrey Hinton, considerado uno de los padres de la IA.

Sobre los sentimientos, el exingeniero de Google y Microsoft dijo que es “irrelevante” si estos sistemas pueden llegar a desarrollarlos, “lo importante es si pueden escapar, replicarse o construir sus propias infraestructuras”.

“Nadie sabe dónde está el punto de no retorno”, afirmó. La SIA es “como un tren que avanza hacia nosotros”, y los académicos, investigadores y directores de muchos laboratorios están gritando: “cuidado con el tren”, pero los líderes mundiales no han escuchado el mensaje y no se sabe si cuando lo hagan podrán detener “una carrera suicida”.

En Chile, el director de Ingeniería en Automatización y Robótica de la Universidad Andrés Bello (UNAB), Miguel Solís, indicó que “persiste una brecha entre adopción y comprensión” de las IA, ya que se habla de ellas, existe una estrategia país, aunque “seguimos trabajando con automatización estática avanzada”.

“El punto clave está en comprender que estos modelos de lenguaje no son aún un nuevo actor cognitivo, sino espejos sofisticados del lenguaje humano. El riesgo no es que piensen por nosotros, sino que nosotros pensemos que lo hacen”, destacó el académico.