¿Frenos para un Cohete?



Publicado el 16 de abril de 2023 por [**Editor**](https://www.xbrl.org/news/brakes-on-a-rocketship/)

Con los nuevos sistemas y aplicaciones de IA generativa que se lanzan, anuncian o demuestran a diario, no sorprende que esta nueva generación de herramientas también esté llamando la atención de los reguladores y legisladores de todo el mundo.

Chat GPT y sus relaciones (competidoras) están brindando una visión cada vez más clara de lo que se avecina. La combinación de algoritmos de aprendizaje automático, un poder de cómputo previamente inimaginable y una "base de datos" de aprendizaje que comprende todo Internet significa que todo tipo de esfuerzos que antes hubieran sido imposibles, están al alcance. Las empresas, las nuevas empresas, los gobiernos y las universidades de todo el mundo se apresuran a explorar lo que es posible hoy y lo que podría ser posible en el futuro.

Si bien las herramientas de hoy pueden sorprendernos con sus capacidades, también pueden impactar con su capacidad para disimular o, técnicamente, para “alucinar”. Más claramente: ¡mienten, y mienten con convicción!

Hace dos semanas, unas 25 000 personas *{Ed, bueno, y probablemente algunos robots}* , incluido Elon Musk,  firmaron  una [carta abierta](https://futureoflife.org/open-letter/pause-giant-ai-experiments/) en la que pedían una moratoria en el trabajo de los laboratorios de IA para entrenar herramientas más potentes que GPT-4 durante al menos 6 meses.

Esta semana, una agencia clave de los EE. UU. está  [buscando comentarios](https://ntia.gov/press-release/2023/ntia-seeks-public-input-boost-ai-accountability)  sobre si las herramientas de IA deben auditarse o licenciarse y, de ser así, cómo. La  [UE](https://www.euronews.com/2023/03/17/eus-ai-regulation-vote-looms-were-still-not-sure-how-unrestrained-ai-should-be)  y el  [Reino Unido](https://cdei.blog.gov.uk/category/artificial-intelligence/)  se han centrado en la ética en la IA durante algún tiempo.

¿Es esta un área donde la regulación puede moverse tan rápido como la tecnología? ¿Debería? ¿Ha alcanzado ya el cohete la velocidad de escape?

*Si bien podemos, y debemos, preocuparnos por la forma en que se desarrollan estas herramientas y la forma en que su uso puede restringirse para un uso positivo, sería una tontería no pensar en las formas en que las herramientas avanzadas de IA afectarán nuestros esfuerzos en informes digitales. Nuestra propia opinión es que muchos aspectos de estas tecnologías serán profundamente importantes para XBRL, ¡y de manera positiva! Una nueva área de enfoque para este boletín, así como para muchos de ustedes. ¡Manténganse al tanto!*

[Ética](https://www.xbrl.org/tag/ethics/)[de AI](https://www.xbrl.org/tag/ai/)[CHATGPT](https://www.xbrl.org/tag/chatgpt/)

Se avecina la votación del reglamento de IA de la UE. Todavía no estamos seguros de cuán desenfrenada debería ser la IA



*Logotipo de ChatGPT en un teléfono inteligente en Washington, DC, marzo de 2023*  -   Copyright   AFP/Euronews

*Por Oliver Rolofs, experto en seguridad estratégica*

*• Actualizado:  17/03/2023*

*Las opiniones expresadas en este artículo son las del autor y no representan de ninguna manera la posición editorial de Euronews.*

Se espera que la tan esperada ley de la Unión Europea sobre inteligencia artificial (IA) se someta a votación en el Parlamento Europeo a fines de este mes.

Pero los esfuerzos de Europa para regular la IA podrían verse truncados de raíz, ya que los legisladores se esfuerzan por ponerse de acuerdo sobre cuestiones críticas relacionadas con la definición, el alcance y las prácticas prohibidas de la IA.

Mientras tanto, la decisión de Microsoft esta semana de eliminar todo su equipo de ética de IA a pesar de invertir $ 11 mil millones (€ 10,3 mil millones) en OpenAI plantea dudas sobre si las empresas tecnológicas están realmente comprometidas con la creación de salvaguardas responsables para sus productos de IA.

En el centro de la disputa en torno a la Ley de IA de la UE se encuentra la necesidad de proporcionar derechos fundamentales, como la privacidad de los datos y la participación democrática, sin restringir la innovación.

**¿Qué tan cerca estamos de la algocracia?**

La llegada de plataformas sofisticadas de IA, incluido el lanzamiento de ChatGPT en noviembre del año pasado, ha provocado un debate mundial sobre los sistemas de IA.

También ha obligado a gobiernos, corporaciones y ciudadanos comunes a abordar algunas cuestiones existenciales y filosóficas incómodas.

¿Qué tan cerca estamos de convertirnos en una -algocracia- una sociedad gobernada por algoritmos? ¿A qué derechos nos veremos obligados a renunciar? ¿Y cómo protegemos a la sociedad de un futuro en el que estas tecnologías se utilicen para causar daño?

Cuanto antes podamos responder a estas y otras preguntas similares, mejor preparados estaremos para cosechar los beneficios de estas tecnologías disruptivas, pero también para protegernos de los peligros que las acompañan.

***La promesa de innovación tecnológica ha dado un gran salto con la llegada de nuevas plataformas de IA generativa.***

*Robin Li, CEO del gigante de las búsquedas Baidu, habla sobre IA en Beijing en 2018*AP Photo/Ng Han Guan

La promesa de la innovación tecnológica ha dado un gran paso adelante con la llegada de nuevas plataformas generativas de IA, como ChatGPT y DALL-E 2, que pueden crear palabras, arte y música con un conjunto de instrucciones simples y brindar respuestas similares a las de los humanos. preguntas complejas

Estas herramientas podrían aprovecharse como un poder para el bien, pero la noticia reciente de que ChatGPT aprobó un examen de licencia médica de EE. UU. y un examen de MBA de Wharton Business School es un recordatorio de los desafíos operativos y éticos que se avecinan.

Las instituciones académicas, los encargados de formular políticas y la sociedad en general todavía luchan por ponerse al día.

**ChatGPT pasó la prueba de Turing, y todavía está en su adolescencia**

Desarrollado en la década de 1950, el llamado Test de Turing ha sido durante mucho tiempo la línea en la arena para la IA.

La prueba se utilizó para determinar si una computadora es capaz de pensar como un ser humano.

El matemático y descifrador de códigos Alan Turing estaba convencido de que algún día un ser humano sería incapaz de distinguir entre las respuestas dadas por una persona real y una máquina.

Tenía razón: ese día ha llegado. En los últimos años, las tecnologías disruptivas han avanzado más allá de todo reconocimiento.

Las tecnologías de inteligencia artificial y los chatbots avanzados de aprendizaje automático todavía están en su adolescencia, necesitan más tiempo para florecer.

***Los optimistas entre nosotros se apresuran a señalar el enorme potencial para el bien que presentan estas tecnologías.***

*Robin Li, CEO del gigante de las búsquedas Baidu, habla sobre IA en Beijing en 2018*AP Photo/Ng Han Guan

Pero nos dan una valiosa visión del futuro, incluso si estos destellos a veces son un poco borrosos.

Los optimistas entre nosotros se apresuran a señalar el enorme potencial para el bien que presentan estas tecnologías: desde mejorar la investigación médica y desarrollar nuevos medicamentos y vacunas hasta revolucionar los campos de la educación, la defensa, el cumplimiento de la ley, la logística, la fabricación y más.

Sin embargo, organizaciones internacionales como la Agencia de Derechos Fundamentales de la UE y el Alto Comisionado de las Naciones Unidas para los Derechos Humanos han tenido razón al advertir que estos sistemas a menudo no funcionan según lo previsto.

Un ejemplo de ello es el sistema SyRI de la autoridad fiscal holandesa que utilizó un algoritmo para detectar sospechas de fraude de beneficios en violación de la Convención Europea de Derechos Humanos.

**¿Cómo regular sin frenar la innovación?**

En un momento en que la IA está cambiando fundamentalmente la sociedad, carecemos de una comprensión integral de lo que significa ser humano.

Mirando hacia el futuro, tampoco hay consenso sobre cómo experimentaremos, y deberíamos, experimentar la realidad en la era de la inteligencia artificial avanzada.

Necesitamos entender las implicaciones de las herramientas sofisticadas de IA que no tienen un concepto de lo correcto o lo incorrecto, herramientas que los actores maliciosos pueden usar mal fácilmente.

* [¿Por qué ChatGPT y otros programas generativos de IA de repente son tan emocionantes?](https://www.euronews.com/2023/03/10/why-are-we-suddenly-so-excited-by-chatgpt-and-other-generative-ai)
* [¿Los programas basados ​​en IA como ChatGPT traen cambios útiles o caos desconocido?](https://www.euronews.com/2023/03/13/are-ai-based-programmes-like-chatgpt-bringing-useful-change-or-unknown-chaos)

*Robin Li, CEO del gigante de las búsquedas Baidu, habla sobre IA en Beijing en 2018*AP Photo/Ng Han Guan

Entonces, ¿cómo hacemos para gobernar el uso de la IA para que esté alineado con los valores humanos? Creo que parte de la respuesta radica en crear regulaciones claras para los desarrolladores, implementadores y usuarios de IA.

Todas las partes deben estar en sintonía en lo que respecta a los requisitos y límites para el uso de la IA, y empresas como OpenAI y DeepMind tienen la responsabilidad de llevar sus productos a la conciencia pública de una manera controlada y responsable.

Incluso Mira Murati, directora de tecnología de OpenAI y creadora de ChatGPT, ha pedido una mayor regulación de la IA.

Si se gestiona correctamente, el diálogo directo entre los responsables de la formulación de políticas, los reguladores y las empresas de IA proporcionará salvaguardias éticas sin ralentizar la innovación.

Una cosa es segura: el futuro de la IA no debe dejarse solo en manos de programadores e ingenieros de software.

**En nuestra búsqueda de respuestas, necesitamos una alianza de expertos de todos los campos**

El profesor Nayef Al-Rodhan, filósofo, neurocientífico y experto en ética de la IA, presenta un caso convincente a favor de un tipo pionero de investigación transdisciplinaria: la neuro-tecno-filosofía (NTP).

NTP defiende la creación de una alianza de neurocientíficos, filósofos, científicos sociales, expertos en inteligencia artificial y otros para ayudar a comprender cómo las tecnologías disruptivas afectarán a la sociedad y al sistema global.

Sería prudente tomar nota.

***Necesitamos asegurarnos de que existan marcos éticos y legales para protegernos de los lados oscuros de la IA.***

*Robin Li, CEO del gigante de las búsquedas Baidu, habla sobre IA en Beijing en 2018*AP Photo/Ng Han Guan

Al-Rodhan y otros académicos que conectan los puntos entre la (neuro)ciencia, la tecnología y la filosofía serán cada vez más útiles para ayudar a la humanidad a navegar los desafíos éticos y existenciales creados por estas innovaciones que cambian el juego y sus impactos potenciales en los riesgos fronterizos consecuentes y los futuros de la humanidad.

En un futuro no muy lejano, veremos robots realizar tareas que van mucho más allá de procesar datos y responder a instrucciones: una nueva generación de humanoides autónomos con niveles de sensibilidad sin precedentes.

Antes de que esto suceda, debemos asegurarnos de que existan marcos éticos y legales para protegernos de los lados oscuros de la IA.

**Encrucijada de civilizaciones llama**

En la actualidad, sobreestimamos nuestra capacidad de control y muchas veces subestimamos los riesgos. Este es un enfoque peligroso, especialmente en una era de dependencia digital.

Nos encontramos en un momento único en el tiempo, una encrucijada de civilizaciones, donde todavía tenemos la agencia para dar forma a la sociedad y nuestro futuro colectivo.

Tenemos una pequeña ventana de oportunidad para las tecnologías emergentes preparadas para el futuro, asegurándonos de que, en última instancia, se utilicen al servicio de la humanidad.

No desperdiciemos esta oportunidad.

*Oliver Rolofs es un experto en seguridad alemán y cofundador de la Conferencia de seguridad cibernética de Munich (MCSC). Anteriormente fue jefe de Comunicaciones en la Conferencia de Seguridad de Munich, donde estableció el Programa de Ciberseguridad y Seguridad Energética.*

*En Euronews, creemos que todas las opiniones son importantes. Contáctenos en view@euronews.com para enviar propuestas o propuestas y ser parte de la conversación.*

TAMBIÉN PODRÍA GUSTARTE

[¿Es el riesgo para la privacidad causado por ChatGPT tan malo como parece?](https://www.euronews.com/2023/04/17/is-the-risk-to-privacy-caused-by-the-likes-of-chatgpt-as-bad-as-it-seems)

[La vigilancia retrospectiva de reconocimiento facial oculta los abusos de los derechos humanos a simple vista](https://www.euronews.com/2023/04/14/retrospective-facial-recognition-surveillance-conceals-human-rights-abuses-in-plain-sight)

[El Parlamento Europeo se reúne con Ucrania para discutir los pasos para la adhesión a la UE](https://www.euronews.com/2023/04/12/european-parliament-meets-with-ukraine-to-discuss-steps-for-eu-membership)

**Edwina Dunn: Reflexiones del Consejo Asesor del CDEI**

****

Publicado por: [Edwina Dunn](https://cdei.blog.gov.uk/author/edwina-dunn/) ,Publicado en:24 enero 2023 - Categorías: [Algoritmos](https://cdei.blog.gov.uk/category/algorithms/) , [Inteligencia artificial](https://cdei.blog.gov.uk/category/artificial-intelligence/) , [Aseguramiento](https://cdei.blog.gov.uk/category/assurance/) , [Datos](https://cdei.blog.gov.uk/category/data/) , [Innovación responsable](https://cdei.blog.gov.uk/category/responsible-innovation/) , [Confianza](https://cdei.blog.gov.uk/category/trust/) , [Innovación confiable](https://cdei.blog.gov.uk/category/trustworthy-innovation/)

Durante mucho tiempo he creído en el poder de los datos para marcar una diferencia significativa en la vida de las personas. Trabajando en el campo durante más de 30 años, he visto las formas en que los datos se pueden usar como una fuerza poderosa para el bien, como el ahorro de dinero relevante y la información con solo tocar un botón, para transformar vidas con la promesa de inteligencia artificial habilitada. descubrimientos de fármacos para curar enfermedades que amenazan la vida.

También he visto que solo podemos obtener estos beneficios usando los datos de manera inteligente y responsable, ganándonos así la confianza del público. Fue este compromiso con la innovación confiable en datos lo que me motivó a unirme al Consejo Asesor del Centro de Ética e Innovación de Datos en noviembre de 2018, y por qué me encantó que me pidieran asumir el cargo de presidente del Consejo Asesor en septiembre de 2021.

Ahora que mi tiempo en el CDEI está llegando a su fin, quiero celebrar lo que la organización ha logrado en ese tiempo, así como mirar hacia el futuro del Centro. Desde sus inicios, brindando revisiones de políticas que dieron forma al pensamiento del gobierno sobre el sesgo algorítmico y la orientación en línea, hasta nuestra posición actual como parte clave de DCMS, el CDEI continúa brindando innovación confiable en datos e IA.

El año pasado fue particularmente emocionante, con el lanzamiento de nuestros tres programas: garantía de IA, transparencia algorítmica y acceso responsable a los datos, que abordan algunos de los problemas más apremiantes en el campo. Mis prioridades para el trabajo del CDEI en estos programas han sido: un mayor enfoque en la innovación del sector privado, junto con el trabajo vital del CDEI en el sector público; permitir el acceso a datos de mejor calidad; y garantizar que la diversidad y la inclusión estén siempre integradas en nuestro enfoque.

**Nuestro programa AI Assurance es un área en particular en la que hemos estado trabajando en estrecha colaboración con el sector privado**para ayudar a ofrecer un entorno favorable a la innovación para la IA. Los proyectos en esta área incluyen nuestro trabajo maravillosamente futurista con el Centro de Vehículos Autónomos y Conectados en vehículos autónomos y nuestra asociación con la Confederación de Reclutamiento y Empleo para desarrollar una guía sobre el uso responsable de la IA en el reclutamiento.

El proyecto de reclutamiento es un ejemplo de dónde el uso de la IA tiene el potencial de ser transformador, pero debemos asegurarnos de limitar cualquier sesgo potencial. Hace tiempo que me comprometo a mejorar la diversidad, centrándome particularmente en el empoderamiento de las mujeres y el acceso equitativo a carreras en el sector tecnológico, y me complació especialmente apoyar este trabajo. Como fundadora de Female Lead, es vital para mí asegurarme de que las tecnologías del futuro no estén moldeadas por los sesgos de hoy.

En el próximo año, estoy emocionado de ver cómo el CDEI trabaja en estrecha colaboración con la industria para examinar las barreras y los habilitadores de AI Assurance, para que pueda continuar siendo un habilitador en este ecosistema (¡que podría valer miles de millones de libras para la economía del Reino Unido!).

**El enfoque del CDEI para el programa de Transparencia Algorítmica ha demostrado nuestro compromiso con la participación inclusiva del público en general.** El programa está ayudando a generar confianza en el uso de herramientas algorítmicas, datos e IA por parte del sector público. Trabajando con la Oficina del Gabinete, el CDEI ha desarrollado uno de los primeros estándares del mundo para la transparencia algorítmica en el sector público. El estándar ayuda a las organizaciones a informar a los ciudadanos sobre las herramientas algorítmicas que utilizan y por qué las utilizan. Este trabajo ha sido moldeado por un compromiso profundo con el público, ayudándonos a comprender lo que se necesita para garantizar que la transparencia sea significativa para los ciudadanos, y diseñando el Estándar basado en estos conocimientos. El programa está explorando cómo se podría usar el Estándar en otros contextos y me complace ver que algunas organizaciones del sector privado ya están considerando cómo replicarlo para sus propios fines.

**Finalmente, nuestro programa de Acceso responsable a los datos ayudará a garantizar que los innovadores tengan acceso a los datos de alta calidad que necesitan, al tiempo que genera confianza pública.**Muchos innovadores (que podrían hacer un gran trabajo con los datos) simplemente no tienen la confianza de poder acceder a los datos y usarlos de una manera confiable y legal. El programa Acceso responsable a los datos del CDEI está ayudando a desarrollar soluciones prácticas para abordar las barreras al uso de datos.

Por ejemplo, el CDEI lanzó desafíos de premios de tecnologías de mejora de la privacidad (PET) con el gobierno de los EE. UU. para impulsar la innovación en estas técnicas prometedoras. Los PET permiten a los innovadores obtener el beneficio de entrenar modelos de IA en datos confidenciales mientras preservan la privacidad de las personas, brindándoles acceso a datos de alta calidad. Estas técnicas emergentes ofrecen la promesa de usar datos para abordar algunos de nuestros mayores desafíos sociales, y ver al CDEI trabajar de cerca con algunos de los más grandes del sector privado, incluidos SWIFT y BNY Mellon, es algo que espero que continúe. Espero con interés el anuncio de los ganadores del concurso de premios a finales de este año.

El compromiso del CDEI con la colaboración y la incorporación de los mejores conocimientos a la organización es una de las cosas que más me ha impresionado del Centro. Estoy orgulloso del papel que ha desempeñado el Consejo Asesor en el desarrollo de los proyectos de CDEI, ya sea vinculando al equipo de CDEI con contactos que han desbloqueado nuevas oportunidades o brindando (¡a veces bastante sólido!) retroalimentación de su propia experiencia. Neil Lawrence, uno de los miembros existentes del Consejo Asesor y profesor de aprendizaje automático de DeepMind en la Universidad de Cambridge, asumirá la presidencia del Consejo Asesor de manera interina hasta septiembre de 2023. Le deseo a Neil y a todos nuestros colegas expertos en el Consejo Asesor, lo mejor para 2023 y más allá.

Aunque mi compromiso formal con el CDEI está llegando a su fin, seguiré su trabajo con interés, seguro de que el equipo continuará comprometiéndose y colaborando con el sector privado y explorando y adoptando algunas de las tecnologías y aplicaciones más novedosas utilizando un rico flujo de datos. El ritmo de la capacidad y el cambio no hace más que aumentar, y la misión del CDEI es más importante que nunca. ¿Se trata menos de ‘podemos’ y más de ‘deberíamos’. . . y si es así ‘cómo’?

Edwina Dunn

Etiquetas: [IA](https://cdei.blog.gov.uk/tag/ai/), [algoritmos](https://cdei.blog.gov.uk/tag/algorithms/), [datos](https://cdei.blog.gov.uk/tag/data/), [ética](https://cdei.blog.gov.uk/tag/ethics/), [gobernanza](https://cdei.blog.gov.uk/tag/governance/), [innovación](https://cdei.blog.gov.uk/tag/innovation/), [responsable](https://cdei.blog.gov.uk/tag/responsible/), [transparencia](https://cdei.blog.gov.uk/tag/transparency/), [confianza](https://cdei.blog.gov.uk/tag/trust/)