

PRIMER PLANO

EL GRAN DESAFÍO DE LA IA

El bloqueo de ChatGPT en Italia y la carta de miles de expertos para frenar la investigación durante seis meses avivan el temor. Otros científicos, en cambio, reniegan de tanto catastrofismo

CÓMO REGULAR EL AUUGE DE LA IA: "FRENAR EL DESARROLLO NO ES LA SOLUCIÓN"

RODRIGO TERRASA MADRID

«Crear una inteligencia superior podría ser el mayor acontecimiento en la historia de la humanidad... Por desgracia, también podría ser el último». Han pasado justo tres años de la advertencia del profesor Russell y sus temores, lejos de disiparse, se han disparado.

Stuart J. Russell tiene 61 años, es profesor de Ciencias de la Computación en la Universidad de California y uno de los primeros teóricos de la inteligencia artificial. Dos de sus libros están considerados los manuales de referencia sobre la materia. El primero, *Inteligencia artificial: un enfoque moderno*, publicado en 1995, sentó las bases de la disciplina y se usa hoy en más de 1.300 universidades de 116 países. El segundo, *Compatible con el ser humano: la IA y el problema del control*, publicado 25 años después, es la obra fundamental sobre el nuevo escenario. En él, Russell explica por qué ha llegado a considerar su propia disciplina como una amenaza existencial para nuestra especie.

«Parece que la marcha hacia la inteligencia sobrehumana es imparable», escribe. «Pero su éxito podría ser la perdición de la raza humana».

Bajo esa premisa con aroma a apocalipsis ciborg, el propio Russell encabeza ahora una larga lista de expertos que han pedido un tiempo muerto en el sector... antes de que sea demasiado tarde. Cerca de 3.000 líderes, entre ellos Elon Musk, CEO de Tesla y Twitter; Steve Wozniak, cofundador de Apple; o el historiador Yuval Noah Harari, han firmado una carta abierta que reclama a todos los laboratorios de inteligencia artificial que detengan al menos durante seis meses los entrenamientos de sus sistemas más potentes para establecer antes unos protocolos de seguridad.

«La IA avanzada podría representar un cambio profundo en la historia de

la vida en la Tierra, y debe planificarse y administrarse con el cuidado y los recursos correspondientes», dice la carta, publicada la semana pasada en la web de la ONG Future of Life Institute.

«¿Deberíamos dejar que las máquinas inunden nuestros canales de información con propaganda y falsedad?», se preguntan los expertos. «¿Deberíamos desarrollar mentes no humanas que podrían superarnos en número, ser más inteligentes, dejarnos obsoletos y reemplazarnos? ¿Deberíamos arriesgarnos a perder el control de nuestra civilización?».

Llevábamos años hablando de ciber guerras, imaginando un planeta gobernado por robots y de repente se nos apareció el Papa Francisco vestido con

«LA IA AVANZADA PODRÍA REPRESENTAR UN CAMBIO PROFUNDO EN LA HISTORIA DE LA VIDA EN LA TIERRA»

«SERÁ IMPOSIBLE SABER QUÉ ES VERDAD Y QUÉ ES MENTIRA... PERO HOY A QUIÉN LE IMPORTA LA VERDAD»

un abrigo acolchado de Balenciaga para cambiarlo todo. Las imágenes (falsas y ultravirales) de Su Santidad paseando por el Vaticano disfrazado como un rapero o las de Trump siendo arrestado por la Policía en plena calle —creadas ambas con herramientas de inteligencia artificial—, han demostrado los peligros de una realidad paralela cada vez más asequible.

«Pronto será imposible saber qué

es verdad o qué es mentira, pero a quién le importa hoy la verdad...», se pregunta Ulises Cortés, director del grupo de investigación de Inteligencia Artificial de Alto Rendimiento en el Centro Nacional de Supercomputación de Barcelona, y uno de los firmantes de la carta. «La calidad de la verdad va a decrecer de manera rapidísima y eso eliminará de forma sistemática el pensamiento crítico y supondrá un daño enorme para la democracia».

En nuestra portada tienen a Yolanda Díaz y a Pablo Iglesias posando ante las cámaras sólo unas horas después de que la primera lanzara su candidatura a la presidencia del Gobierno al margen de Podemos. Aquí a Sánchez y a Feijóo de risas en la barra de un bar. ¿Y qué me dicen del paseo de Iglesias y Abascal cogidos del brazo? En sólo unas horas todo puede ser (aparentemente) real.

«Corremos un riesgo tremendo de instalarnos en una desconfianza total», explica Guillermo Simari, catedrático de Inteligencia Artificial por la Universidad Nacional del Sur de Bahía Blanca, en Argentina. «Hay gente que cobra por estar sentado delante de un ordenador generando mentiras. Antes esos trolls eran bastante estúpidos, pero ahora pueden contar con un asistente capaz de hacer un trabajo perfecto que, en el ambiente adecuado, puede provocar cualquier barbaridad».

El lanzamiento el año pasado de DALL-E 2 o ChatGPT, los modelos desarrollados por la empresa OpenAI capaces de reproducir imágenes y obras de arte o conversar y escribir como un humano, han agudizado el debate en los últimos meses.

A un lado están quienes recelan del exceso de alarmismo. «Nadie podía

ASÍ SE HIZO...

48 HORAS Y 600 FOTOS PARA FABRICAR TRES IMÁGENES DE FICCIÓN

Detrás de las dos imágenes que ilustran este reportaje y de la fotografía que aparece en la portada del periódico de hoy hay muchísimas horas de trabajo, varios programas de inteligencia artificial y un grupo de creadores de vídeos e imágenes dedicado a la sátira política desde 2010. Se llaman United Unknown y sus disparatados montajes son ya célebres en las redes sociales, sobre todo desde que Andreu Buenafuente los reclutara para la televisión. «Venimos todos de diferentes campos relacionados con la fotografía, la animación o el diseño, pero empezamos a trabajar con la sátira casi como terapia personal, para poder reírnos de una cosa tan seria como la política», explica Sergey, uno de sus miembros. La noche del pasado viernes contactamos con ellos con el encargo de ilustrar este reportaje con tres retratos políticos imposibles. «Ha sido un reto porque nunca habíamos creado dos sujetos en una misma imagen, pero hemos aprendido mucho de este trabajo», cuentan. Durante todo el fin de semana y utilizando programas como Dall-e, Midjourney y Stable Diffusion, crearon las tres imágenes que aparecen hoy en nuestras páginas. «Es un trabajo muy intenso porque había que reentrenar modelos entrenados ya y especializarlos con los datos de los sujetos que queríamos tratar». Cada una de las fotos se logró tras generar entre 100 y 200 imágenes por cada retrato, procesadas, reajustadas y retocadas después hasta conseguir el resultado más parecido a la realidad. «Pensábamos que Yolanda sería más fácil, pero es la que más nos ha costado». ¿Hay algún encargo que nunca aceptarían? «Nos gusta divertirnos pero nunca haríamos nada ofensivo. Bordeamos el límite, pero hasta la fecha no hemos tenido queja».



PRIMER PLANO

EL GRAN DESAFÍO DE LA IA

IMAGEN
CREADA POR
INTELIGENCIA
ARTIFICIAL



imaginar una disrupción tan grande, pero preocuparse en este momento por el riesgo de una superinteligencia es como preocuparse por la superpoblación en Marte», ironiza Óscar Cordón, catedrático de IA de la Universidad de Granada.

¿Es mínimamente factible entonces darle a *pause* en la investigación, al mismo tiempo y en todo el mundo? ¿Sirve de algo hacerlo durante apenas medio año?

«La solución no puede ser frenar el desarrollo científico», sostiene José Ignacio Latorre, autor de *Ética para máquinas*.

«Yo comparto la preocupación ante un avance tecnológico brutal que llega a una sociedad no educada y con un sistema jurídico no preparado, pero lo que hay que hacer no es parar, sino acelerar también la creación de leyes, mantener una supervisión ética e ir contra todo lo que no sea correcto».

Al otro lado, figuran quienes creen que las advertencias de los expertos se quedan cortas. «Pausar los desarrollos de IA no es suficiente, necesitamos cerrarlo todo», reclamaba hace sólo unos días desde la revista *Time* Eliezer Yudkowsky, el controvertido fundador del Instituto de Investigación de la Inteligencia de las Máquinas de Berkeley. «Si alguien construye una IA demasiado poderosa, en las condiciones actuales, lo esperable es que todos los miembros de la especie humana y toda la vida de la Tierra mueran poco después».

Y en medio queda una larga nómina de expertos en inteligencia artificial preocupados, como mínimo, por el vertiginoso salto de los últimos años. «Estamos en un momento clave porque se ha acumulado la cantidad de energía suficiente como para

la ciudad sin ningún control. Ninguno de esos coches ha pasado ninguna certificación y, además, los coches no requieren carné de conducir. Cualquiera puede usarlos, da igual que sea menor. Parece peligroso, ¿verdad? Pues justo eso ocurre en internet».

La regulación hasta ahora es casi inexistente. Ni EEUU ni China parecen dispuestos a ceder un milímetro en su batalla particular y sólo Europa ha dado pasos para legislar. «Es muy difícil ser ético cuando los demás no lo son», apunta Simari.

La UE propuso en 2021 una primera ley que, tras dos años de debate, sigue pendiente.

«La ley va muy por detrás de la tecnología, así que necesitamos herramientas de transparencia y explicabilidad de cada nuevo sistema, al menos un sello de certificación. Lo tienen los plátanos de Canarias, pero no estos sistemas, que pueden ser mucho más peligrosos», lamenta Cristina Urdiales, investigadora en robótica e inteligencia artificial en la Universidad de Málaga.

La semana pasada, Italia bloqueó con efecto inmediato la herramienta ChatGPT por no respetar la ley de protección de datos del país. Sólo unas horas después se supo que un joven se había quitado la vida en Bélgica tras semanas de «conversaciones frenéticas» con otro *chatbot* que le había creado la ilusión de tener una respuesta a todas sus inquietudes.

«El problema fundamental es que no tenemos ningún conocimiento científico sobre el comportamiento de los sistemas, son impredecibles», subraya Guillermo Simari. «Quienes los promocionan nos dicen que son capaces de superar sin problemas un examen de Medicina o de Derecho, pero no nos dicen que lo que sale de ese examen no es un médico capaz, ni tampoco un abogado. Los sistemas sólo saben pasar el examen, pero no tienen una inteligencia real, son sólo un simulacro».

Una noticia más: el pasado verano una empresa china especializada en aplicaciones para móviles decidió nombrar como director ejecutivo a un sistema de inteligencia artificial. Seis meses después, el valor de la compañía creció un 10%.

«A corto plazo el mayor peligro es la instrumentalización política para aumentar la desinformación», asegura Latorre. «A medio plazo, el riesgo es la pérdida inevitable de puestos de trabajo».

Según un estudio de Goldman Sachs, los últimos avances en inteligencia artificial podrían llevar en los próximos años a la automatización de alrededor de 300 millones de empleos, pero también podrían disparar la productividad laboral y elevar el PIB mundial en un 7% durante la década que viene.

EEUU Y CHINA NO QUIEREN CEDER NI UN MILÍMETRO, SÓLO EUROPA DA PASOS PARA LEGISLAR

«ESTAMOS EN UN MOMENTO CLAVE PORQUE SE ACUMULA ENERGÍA PARA QUE TODO EXPLOTE»

que todo explote», asegura Cortés. «La tecnología actual en manos de ciertas personas y en un escenario de falta de legislación, o al menos de legislación laxa, se convierte en algo infinitamente tético y peligroso».

Cortés pone un ejemplo muy gráfico: «Imagínate que mañana aparece un Elon Musk español y pone mil coches en Madrid y los deja para que la gente los utilice y los comparta por

PRIMER PLANO

EL GRAN DESAFÍO DE LA IA

IMAGEN
CREADA POR
INTELIGENCIA
ARTIFICIAL

VIENE DE PÁGINA 3

En su blog personal, el mismísimo Bill Gates, fundador de Microsoft, celebraba hace sólo unas semanas la nueva era de la IA como la mayor revolución tecnológica de los últimos 40 años. «Va a cambiar la forma en que las personas trabajan, aprenden, viajan, obtienen atención médica y se comunican entre sí. Las empresas se distinguirán por lo bien que lo utilizan», aseguraba Gates.

Sólo unos días antes, Microsoft había despedido a todo el departamento de ética de su equipo de IA. Eran sólo siete personas, pero en 2020 llegaron a ser más de 30. «Eran líderes en ética hasta que se dieron cuenta de que podían competir con Google», apunta Óscar Cordón, que pone el foco en otra de las claves del debate: la competencia empresarial.

OpenAI, la compañía que está detrás de los últimos grandes avances, nació en 2015 como un laboratorio «sin ánimo de lucro». Sólo cuatro años después, Microsoft invirtió 1.000 millones de dólares en la empresa. UBS pronostica que el mercado de servicios de inteligencia artificial alcanzará los 90.000 millones en 2025.

«Quizás algunas de las corporaciones que están detrás de esa carta son las que van perdiendo en el campo de la investigación», insiste Cordón. O quizás no hay algoritmo capaz de combinar tecnología, negocio y moral. «A las grandes empresas, mientras su actividad sea legal, sólo les importa el dinero», recuerda Guillermo Simari.

¿Qué espacio tiene entonces la ética en este universo? Para José Ignacio Latorre, el mayor dilema que plantea ahora mismo la inteligencia artificial es la «delegación de decisiones». «Tener a mano herramientas tan potentes como las actuales hace que tus propias decisiones queden condicionadas y eso es un problema ético profundo. El corazón del problema. Produce vértigo porque empobrece el razonamiento humano y nos empuja a ceder decisiones de forma sistemática».

Latorre es, aun así, optimista, pero diríamos que su visión es minoritaria. «Reto a toda esa gente que cree que los humanos somos una barbaridad a que viaje a 1960 y tenga una úlcera de estómago. ¿Cuánta gente le daría al botón para viajar a ese pasado? Sí, yo soy optimista, pero creo que debemos reaccionar ya, no podemos perder medio siglo para legislar algo que avanza tan rápido».

Stuart Russell decía que la inteligencia artificial es hoy como el genio dentro de la lámpara, ese que en los cuentos nos concedía tres deseos. La inteligencia artificial —como el genio— hará lo que pidamos que haga, pero eso no significa que el resultado final sea el que esperábamos.

«¿Sabes siempre cuál es el tercer deseo?», pregunta Russell en su libro. «Por favor, deshaz los dos primeros deseos, porque lo he arruinado todo».



Imagen de Pablo Iglesias y Santiago Abascal generada por inteligencia artificial. UNITED UNKNOWN

MARK COECKELBERGH | PROFESOR DE FILOSOFÍA DE LOS MEDIOS Y LA TECNOLOGÍA

«UNA MORATORIA CONTRA LA IA NO TIENE SENTIDO»

RODRIGO TERRASA
El pensador belga Mark Coeckelbergh (Lovaina, 1975), profesor de Filosofía de la Tecnología en la Universidad de Viena, acaba de publicar *Filosofía política de la inteligencia artificial* (Cátedra), un manual para hacer frente a todos los desafíos que plantea el imparable desarrollo tecnológico.

Pregunta. ¿Qué opina de la carta de los expertos que piden una pausa de seis meses para la IA?
Respuesta. No estoy de acuerdo con su contenido. Para mí, esta carta encaja con una exageración transhumanista de la IA y una narrativa a largo plazo que no comparto. No creo que la IA vaya a evolucionar hasta convertirse en una supe-

rinteligencia ni que debamos centrarnos en un futuro lejano. Tenemos que fijarnos en la ética y centrarnos en los riesgos de ahora. No necesitamos ciencia ficción. También es cuestionable que una moratoria tenga sentido porque la gente sabrá cómo burlarla.
P. ¿Cuáles son los riesgos más inmediatos?

R. Intelligencias como ChatGPT recogen información de internet y esa información no siempre es exacta. El funcionamiento del programa también puede dar lugar a mezclas de información falsa o engañosa. La IA también puede utilizarse para manipular y a la vez puede llegar a ser muy manipuladora: a través de sus respuestas puede persuadir a la gente para que haga cosas. En Bélgica ha habido un caso de un *chatbot* que persuadió a alguien para que acabara con su vida. Se trata de riesgos graves e inminentes.

P. En términos de ética, ¿cuál es el mayor dilema que plantean los avances en IA?

R. Hay que elegir entre funcionalidad y poner límites éticos. Es bueno regular y restringir este tipo de IA. Pero, ¿hasta qué punto sigue siendo funcional si, por ejemplo, se filtran determinadas opiniones en nombre de la corrección política? ¿Hasta qué punto está justificado ese tipo de censura y quién debe tomar las decisiones al respecto? Actualmente, un gran problema es que el desarrollo de la IA está en manos de unas pocas grandes empresas y eso no es democrático.

P. La carta de los expertos dice que la IA puede cambiar la historia de la vida en la Tierra...

R. Como otras grandes tecnologías, la IA tendrá un enorme impacto en la sociedad, pero la carta exagera.

P. ¿Es realmente posible que los humanos perdamos el control de nuestra civilización?

R. El miedo a la pérdida de control y al caos ha ocurrido siempre. No hay pérdida de control cuando se regula adecuadamente esta tecnología. La pérdida de control que debería preocuparnos es que la IA siga en manos de las grandes tecnológicas y que su futuro esté influido por las personas que firmaron la carta. Son parte del problema, no de la solución... Necesitamos una IA democrática y una moratoria no es la mejor forma de abordar estos problemas. Necesitamos una regulación y más investigación sobre la ética y asegurarnos de que el uso de la inteligencia artificial por parte del Estado no conduzca a más autoritarismo y opresión.

P. ¿Quién debería auditar los avances del sector y qué papel deberían desempeñar entonces los Gobiernos en esta situación?

R. Necesitamos auditorías independientes y los Gobiernos deberían tomar medidas: no una moratoria, sino un marco sólido de regulación y más inversión en investigación independiente sobre el impacto ético y político de la IA. Actualmente hay mucho trabajo en el mundo académico, pero los medios de comunicación están más interesados en lo que dice gente como Elon Musk que en mostrar los resultados de personas que están pensando en estos temas desde hace años.